**"Enhancing Crime Detection: A CCTV Video-based Surveillance System"**

**Mrs.Rasika Rewatkar**

*Asst. Professor Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

Nagpur, India

**Ms. Gayatri Thakre**

*Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

**Mr. Dnyaneshwar Iptekar**

*Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

**Mr. Kapil Biradar**

*Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

**Ms. Geethashree Devi Patha Pati**

*Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

**Mr. Mahesh Parsutkar**

*Information Technology*

Kavikulguru Institute of Technology and Science, Ramtek

*Abstract—* CCTV cameras are commonly used for surveillance worldwide. However, despite their widespread use, there remain a significant number of crimes due to the lack of an effective system for detecting and controlling illegal activities. Our project not only focuses on real-time crime detection using video data from CCTV cameras but also addresses broader issues. We detect crimes by analyzing the captions generated from the video data, specifically by identifying keywords related to criminal activities such as "knife," "thief," "fire," and "assault." These captions are stored in text files along with timestamps, enabling us to search for specific events in the log. This approach is more secure than storing entire CCTV-captured videos since we store only text files, which can be further secured through robust encryption algorithms.

*Keywords- CCTV, Video Surveillance, Video Analytic, Data Analysis, Criminal Identification.*

# Introduction

Today, the amount of public violence has increased dramatically. This has resulted in the ubiquitous use of surveillance cameras. But almost all systems today require human inspection of these videos. which is virtually inefficient. It is therefore necessary to have a practical system that can automatically monitor and identify the surveillance videos. The development of various deep learning techniques, especially violence and fighting, and that's what we want to apply. In this project, we have explored one strategy from different strategies to find out the saliency of the features from pretrained models in detecting violence in videos. A dataset has been I collected raw surveillance videos from YouTube, sliced them into clips within 5 seconds at 20 fps, and labeled each clip as Violent or Non-Violent Behavior. We got 1600 clips and 160,000 frames as a new data set for real-world violent behavior detection under surveillance cameras. Unfortunately, I couldn't use more than 1,600 videos. because to train this model on a larger amount of data than this needs a super machine, and when training the model, I noticed that the more data, the more accurate the model. That is why it would be great to use a larger amount of data with a supercomputer. The method consists of extracting a set of frames belonging to the video, sending them to a pretrained network called VGG16, obtaining the output of one of its final layers, and from these outputs training another network.

architecture with a type of special neuron called LSTM. These neurons have memory and are able to analyze the temporal information of the video; if at any time they detect violence, it will be classified as a violent video. In the end, this strategy provides an accuracy of 74.58% on the training set and 76.25% on the validation set. And I have used regularization techniques to avoid the risk of overfitting.

# Ease of Use

## Purpose

The primary purpose is to enhance public safety by identifying and preventing criminal activities. These systems can help law enforcement respond more effectively to crimes in progress. These systems can help law enforcement allocate their resources more efficiently. By focusing on areas or times with higher crime rates, police can be more effective in their duties. These systems can help law enforcement allocate their resources more efficiently. By focusing on areas or times with higher crime rates, police can be more effective in their duties. They can provide real-time alerts to law enforcement or security personnel when suspicious activity is detected, allowing for a rapid response.

## Problem Statement

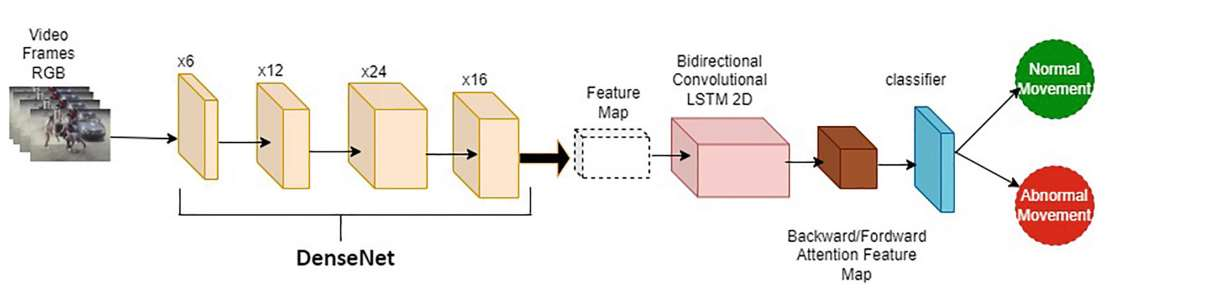
Crime activities like chain snatching and trying to kill people or rare events in a city even though we have servlets cameras, it is difficult for a human to monitor such an event. So, a developing a system that will classify the event as a normal or abnormal event using Machine learning technique. If it is an abnormal event, it will send an alert message to concern department.

# Literature Review

Criminal Detection by Analyzing CCTV Footages (2023)

In May 2023, the International Journal of Engineering Research and Technology (IJERT) developed a model of crime detection by analyzing CCTV footage to identify the crime. The crime detection method's importance resides in its capacity to foresee and stop illegal activity. While conventional techniques are useful, they frequently work in isolation. Therefore, a device that could combine the advantages of these traditional procedures would be quite helpful. Machine learning (ML) techniques were used in a study that used criminal data from Vancouver collected over a 15-year period to predict, detect, and analyze criminal activity. Data gathering, classification, pattern recognition for criminal activity, forecasting, and visualization was all part of the analysis. The criminal dataset was examined using K-nearest neighbor (KNN) models and enhanced decision trees. A total of 560,000 criminal datasets were examined from 2003 to 2018, and criminal activity was predicted with an accuracy range of 39% to 44%. The Chicago crime dataset was used to construct machine learning and data science-based models to predict and identify criminal activity. To find the most precise model for training, many combinations of ML models, including logistic regression, SVM/KNN classification, decision trees, random forests, and Bayesian models, were looked at. The classification accuracy that was highest was 78.7% for the KNN algorithm. This study's major goal was to persuade law enforcement organizations to use ML-based techniques to anticipate, identify, and settle illegal activity more successfully, ultimately lowering crime rates in society. In a different strategy, a deep neural network (DNN)-based feature-level data fusion method was suggested to accurately predict the occurrence of crimes. This approach is required. fusing environmental background knowledge with multimodal data from many fields. Data from an online crime statistics database (Chicago) as well as meteorological and demographic information was included in the database used for crime prediction. For the purpose of predicting crime, various ML models, including SVM, regression analysis, and

kernel density estimation (KDE), were used. The accuracy of the SVM and KDE models was 67.01% and 66.33%, respectively, whereas the accuracy of the suggested ML model was 84.25%.



## Process Model Adopted

Image Acquisition and Preprocessing: In this module, we will get the data from the Online Image Acquisition and Preprocessing In this module, we will get the data from an online source. Further, we will resize the image for future use. Image resizing, or image scaling, is a geometric image transformation that modifies the image size based on an image interpolation algorithm. This image scaling process can increase or decrease the resolution of a target image so that the absolute size of the image data is adjusted.

## Design

* Use Casse Diagram:

In the Unified Modeling Language (UML), a use case diagram can summarize the details of your system's users (also known as actors) and their interactions with the system. To build one, you'll use a set of specialized symbols and connectors. An effective use case diagram can help your team discuss and represent:

• Scenarios in which your system or application interacts with people, organizations, or

external systems

• Goals that your system or application helps those entities (known as actors) achieve

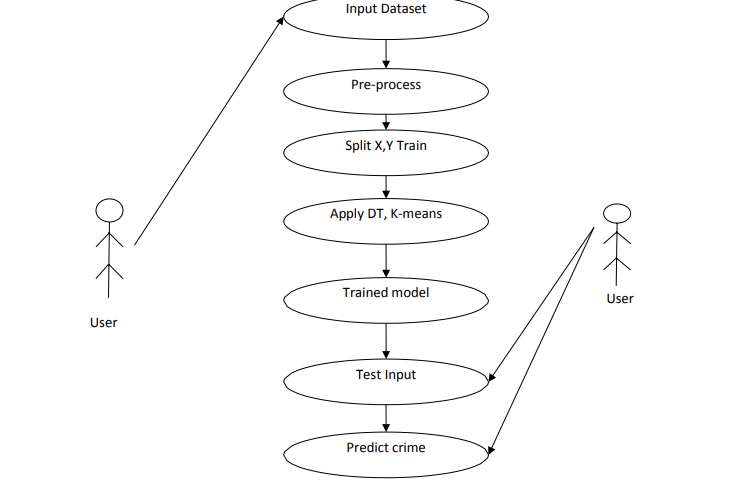
• The scope of your system

A use case diagram doesn't go into a lot of detail—for example, don't expect it to model the

order in which steps are performed. Instead, a proper use-case diagram depicts a high-level

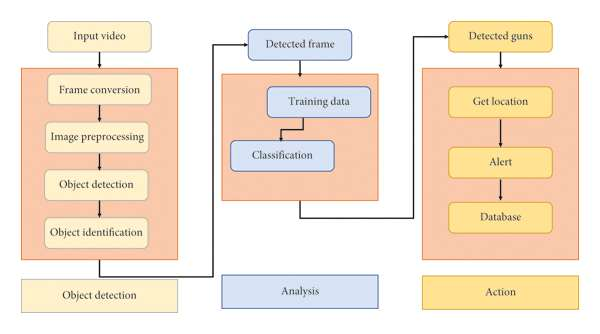
overview of the relationship between use cases, actors, and systems. Experts recommend that

Use case diagrams can be used to supplement a more descriptive textual use case.



## Data Flow Daigram

In this project, we have attempted to develop an integrated framework for reconnaissance security that distinguishes the weapons progressively. If identification is positively true, it will caution or brief the security personnel to handle the circumstances by arriving at the location of the incident through IP cameras. We propose a model that provides a visionary sense to a machine to identify an unsafe weapon and can also alert the human administrator when a gun or firearm is obvious on the edge. Moreover, we have programmed entryways 12 locking framework when the shooter seems to carry an appalling weapon. On the off chance conceivable, through IP webcams, we can likewise share the live photo to approach security personnel to make the move in the meantime. Also, we have constructed an information system for recording all the exercises to convey impact activities in the metropolitan territories for a future crisis. This further ends up in designing the database for recording all the activities in order to take prompt action in future emergencies. Figure 1 presents the overall generalized approach of our research work, divided into three parts.

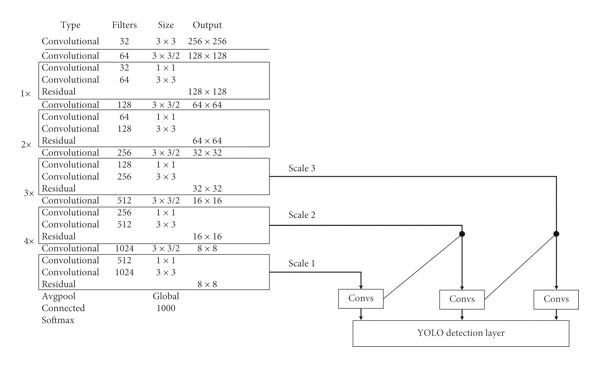


## Software Requirements

YOLO Library

As mentioned earlier, YOLO is a pretrained object detector; a pretrained model simply means that another dataset has been trained on it. It is extremely time-consuming to train a model from scratch; it can take weeks or a month to complete the training step. A pretrained model has already seen tons of objects and knows how each of them must be classified. The weights in the abovementioned pretrained model have been obtained by training the network on the COCO and ImageNet datasets. Thus, it can only detect objects belonging to the classes present in the dataset used to train the network. It uses Darknet-53 as the backbone network for feature extraction and uses three-scale predictions. The DarkNet-53 is again a convolutional neural network that has 53 layers, as elucidated in Figure 4. DarkNet-53 is a fully convolutional neural network. The pooling layer is replaced with a convolution operation with stride 2. Furthermore, residual units are applied to avoid gradient dispersion. Architectural details of DARKNET-53 layouts [10]

fig:Architectural description of YOLO V3



# SYSTEM DESCRIPTION

* + 1. Project Resources
* Hardware Requirements

All the physical equipment’s, i.e., input devices, processors, output devices, and interconnecting processors of the computers, are called hardware. The hardware requirements are the requirements of a hardware device. Most hardware only has operating system requirements or compatibility.

|  |  |
| --- | --- |
| **System** | Pentium IV 2.4GH |
| **Processor** | Intel core i3-6100 or above min 1.8 GHz |
| **RAM** | 12GB |
| **Hard-disk** | 1TB |
| **Internet Connection** | Min 1mbps or more effective |

Table 1: Hardware Requirements

## Software Requirements

The software requirements are a description of the features and functionalities of the target system. requirements convey the expectations of users from the software product.

|  |  |
| --- | --- |
| **Operating Environment** | Windows 8/10/11 |
| **Model building** | Python, CNN, Anaconda, Jupiter notebook,  Visual studio code, Android Studio. |
| **Backend Server** | XAMPP |
| **Frontend** | HTML, CSS, JS, Android Studio |

Table 2: Software Requirements

##### Conclusion

In this study, a model that helps in human-violence detection and also the detection of faces and objects is developed. These two different models were converted into a pickle file, which is imported into the local website that has been created using CSS/HTML as the front end. The FAST API web framework helps in combining these deep learning models and implementing them on the website. On this website, there are two buttons provided, one for the human-violence detection model and another for the object and face detection model. Using these, the uploaded video results in either human-violence or non-human-violence, and another button option is provided that helps to detect the objects and faces in the video. Project management requires that all scheduled work or tasks are tracked. A project tracking spreadsheet is a useful visual tool to track each task's progress against the original plan.

*Future Scope*

The future scope for crime detection systems using CCTV video projects is promising. Advancements in artificial intelligence, particularly computer vision, can enhance video analytics for more accurate identification and tracking of individuals. Integrating machine learning algorithms can improve anomaly detection, making these systems more effective in identifying suspicious behavior. Additionally, the incorporation of real-time data analysis and the development of smarter surveillance technologies could further enhance crime prevention and response capabilities. Privacy considerations and ethical implementation will be crucial aspects to address as these systems evolve.

##### References

1. Kaya V, Tuncer S, and Baran A (2021), Detection and Classification of Different Weapon Types Using Deep Learning Applied Sciences, 11 (16), 7535.
2. Singh P. and Pankajakshan V. (2018) A Deep Learning-Based Technique for Anomaly Detection in Surveillance Videos Proc. of the 24th National Conf. on Communications, pp. 1-6
3. Dandage V., Gautam H., Ghavale A., Mahore R., and Sonewar P. A 2019 Review of Violence Detection Systems Using Deep Learning Int. Research Journal of Engineering and Technology 6(12), pp. 1899–1902.
4. Wang K., Liu M. 2022 YOLOv3-MT: A YOLOv3 Using Multi-Target Tracking for Vehicle Visual Detection Appl. Intell. 52, pp. 2070–2091.
5. Antoniou A. and Angelov P. (2016 A General-Purpose Intelligent Surveillance System for Mobile Devices Using Deep Learning. Proc. of the Int. Joint Conf. on Neural Networks, pp. 2879–2886.
6. Savran A 2007 Multifeedback-Layer Neural Network IEEE Transactions on Neural Networks, 18 (2), pp. 373–384.
7. Bilinski P. and Bremond F. (2016 Human Violence Recognition and Detection in Surveillance Videos. Proc. of the 13th IEEE Int. Conf. on Advanced Video and Signal-Based Surveillance. 30-36.
8. Fu E Y, Leong H V, Nga G, and Chan S (2016) Automatic Fight Detection in Surveillance Videos Proc. of the 4th Int. Conf. on Advances in Mobile Computing and Multimedia, pp. 225-234
9. Manoharan S. 2019, Image Detection Classification and Recognition for Leak Detection in Automobiles. Journal of Innovative Image Processing, 01 (02), pp. 61–70.
10. H.W. Kang, H.B. Kang, Prediction of crime occurrence from multi-modal data using deep learning, PLoS One 12 (4) (2017), e0176244.