Criminal Activity and Violence Detection on surveillance footage

Major Project (CC4270) Report

Submitted in the partial fulfillment of the requirement for the award of

Bachelor of Technology

in

Computer and Communication Engineering

By:

Harsh Kushwah
209303076

Under the supervision of:

Dr. Manoj Kumar Sharma

MANIPAL UNIVERSITY
JAIPUR

INSPIRED BY LIFE

May 2024

Department of Computer and Communication Engineering
School of Computing and Intelligent Systems
Manipal University Jaipur
VPO. Dehmi Kalan, Jaipur, Rajasthan, India – 303007

Department of Computer and Communication Engineering

School of Computing and Intelligent Systems, Manipal University Jaipur,
Dehmi Kalan, Jaipur, Rajasthan, India- 303007

STUDENT DECLARATION

*I hereby declare that this project (***Criminal Activity and Violence Detection on surveillance footage***) is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the University or other Institute, except where due acknowledgements has been made in the text.*

Place: Manipal University Jaipur
Date: 23/05/2024

Harsh Kushwah
(209303076)
B.Tech. (CCE) 8th Semester

Department of Computer and Communication Engineering
School of Computing and Intelligent Systems, Manipal University Jaipur,
Dehmi Kalan, Jaipur, Rajasthan, India- 303007

Date: 23/05/2024

CERTIFICATE FROM SUPERVISOR

*This is to certify that the work entitled “***Criminal Activity and Violence Detection on surveillance footage”** *submitted by* **Harsh Kushwah** *(209303076) to* **Manipal University Jaipur** *for the award of the degree of* **Bachelor of Technology** *in* **Computer and Communication Engineering** *is a bonafide record of the work carried out by him/ her under my supervision and guidance from Jan 2024 to May 2024.*

Dr. Manoj Kumar Sharma
Internal Supervisor,
Department of Computer and Communication Engineering
Manipal University Jaipur

Department of Computer and Communication Engineering
School of Computing and Intelligent Systems, Manipal University Jaipur,
Dehmi Kalan, Jaipur, Rajasthan, India- 303007

Date: 23/05/2024

PROJECT COMPLETION CERTIFICATE

This is to certify that the project entitled "Criminal Activity and Violence Detection on surveillance footage" was carried out by Harsh Kushwah (209303076) under my supervision from January 2024 to May 2024 at Manipal University Jaipur.

Dr. Manoj Kumar Sharma
Associate Professor,
Department of Computer and Communication Engineering
Manipal University Jaipur

Date: 23-05-2024

Place: Manipal University Jaipur

Prof. (Dr.) Sunil Kumar
Head of the department,
Department of Computer and Communication Engineering
School of Computing & Intelligent Systems
Manipal University Jaipur

ACKNOWLEDGEMENT

My heartfelt thanks goes to my esteemed mentor, Dr. Manoj Kumar Sharma. His in-depth expertise, inspirational leadership, and unwavering devotion were critical to the success of my internship. His knowledge and persistent encouragement made studying more simpler for me and raised the quality of my work. I owe him a lot for his prompt counsel and continuous support, which helped me remain on track and produce fantastic results.

My heartfelt gratitude also goes to Dr Sunil Kumar (Head of Department) and the entire department of Computer and Communications Engineering for providing me with this fantastic opportunity and making the entire process so enjoyable and rewarding. Their encouragement, resources, and support were invaluable throughout my internship, contributing significantly to my academic and professional development.

I also want to thank my family and close friends for their unwavering support and encouragement during this endeavour. Their love, support, and belief in my abilities have been critical in allowing me to persevere and achieve my goals. Without their ongoing encouragement and support, I would not have reached this milestone.

ABSTRACT

We frequently see instances of criminal activity and public violence; we may learn about these events via the media or from personal experience. The dark world of hate, retaliation, and violence surrounds us. India is ranked 61st by the World Population Review, with a Crime Index (Numbeo) of 44.4. Even with their poorly run administrations and often documented terrorist acts, Our neighbors Pakistan and Sri Lanka score higher than India on the Crime Index (Numbeo), at 42.8 and 42.2, respectively [1].

Crime against women has also escalated in several regions of India in recent years. Between January 1, 2021 and December 31, 2021, the police in India documented 428,278 cases of crime against women. It is a 26.35% increase over the previous six years, beginning with instances in 2016. The majority of these incidents involved abductions, domestic violence, rapes, and assaults [2].

Law enforcement agencies are overworked, operating under outmoded processes that are equipped to deal with criminals' ever-changing strategies. Traditional approaches, which rely on slow-motion analysis and lack real-time awareness, are like putting out a raging fire with a bucket of water. This grave situation needs strong action, not incremental improvement. We need innovative, data-driven systems that give law enforcement real-time information and lightning-fast reactions.

This project proposes the development of a novel end-to-end system for real-time crime detection and notification. By addressing common problems including theft, arson, burglary, assault, violence, riots, vandalism, and hate crimes, this approach seeks to improve public safety in India.

When a pre-defined criminal incident is successfully identified, the system will automatically start the notification procedure. This notification will be sent to the appropriate law enforcement agencies around the reported crime. The goal of this real-time notification system is to drastically cut down on reaction times so that police can get involved right away.

The suggested strategy might completely change India's approach to public safety. This initiative aims to make the community safer by allowing proactive crime detection and facilitating speedier response times.

Table Of Contents

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction to Crime Detection

Criminal activity presents a major threat to public safety and security. Law enforcement agencies are always seeking to develop novel, effective strategies for locating, apprehending, and deterring criminal behavior. In this drive to fundamentally revolutionize crime prevention and detection, artificial intelligence (AI) has emerged as a key tool.

Law enforcement's never-ending mission of deterring and reacting to crime typically necessitates striking a balance between limited resources and the limitations of human nature. The real potential of artificial intelligence is found here.

A growing number of people have a sense of unease because of the increasing incidence of new types of crime. It is becoming more and more difficult for law enforcement to stay current because of the employment of outdated methods that are just too sluggish and ineffectual for the complex world of today. What if we could arm law enforcement with a powerful "machine" that would assist them or even predict crimes before they happen?

The same goal is being pursued by this initiative, which will enable law enforcement to identify criminal activity on surveillance film and take timely action to stop it.

Here are some features of the project:

1. **Detection of Violence:** Because of its adaptability, the model which was trained on large datasets can identify violent incidents in surveillance film and send automatic notifications to the appropriate authorities.

2. **Detection Of Firearms:** The model in the project tries to effectively classily any firearm located in the footage/image and notify the concerned authorities with the intention of preventing any crime before happening.

3. **Alert Generation:** This project offers the ability to alert the concerned authorities after detecting any type of anomaly in the surveillance footage to prevent crimes before happening or act fast to catch the culprit.

4. **Dashboard:** A dashboard is offered to manage all the features and data collected and processed.

1.2 Project Statement

- Problem:

Public safety is an increasing problem for communities all around the world. The present security measures need to be reevaluated considering the concerning increase in criminal activity. Regrettably, many currently used remedies are reactive in nature and need a substantial investment of resources. Furthermore, it might be difficult to keep accurate records and successful response plans due to human error and criminals' constantly changing methods. It is critical to investigate substitute strategies that give priority to preventative actions and promote a more secure atmosphere for all.

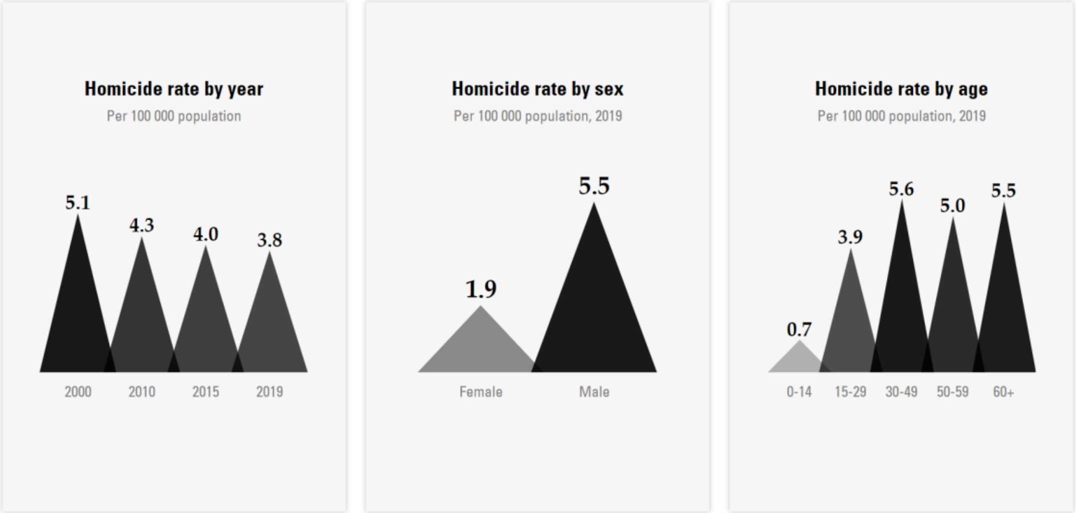

Figure 1: Homicide rates in India [3]

- Solution:

A system for detecting crime and violence that might be used to recognize and categorize abnormalities in the surveillance videos and alert those responsible for law enforcement agencies, enabling them to act more quickly and maybe preventing the anomaly from occurring.

- Benefits:

A crime and violence detection system could have several benefits, including:

1. *Enhanced Security: A system that can detect potential criminal activities and generate alerts can rapidly decrease the number of crimes happening.*

2. *Reduced Failure to report: A lot of crimes go unreported, which makes them difficult to find. this system seeks to detect and report every illegal activity in real-time which helps the authorities to detect even unregistered crimes.*
3. *Streamlined reaction: The approach enables a quicker law enforcement reaction by doing away with the requirement for civilian reports, which may result in a more effective investigation procedure.*
4. *Scalability and Adaptability: The system would be implemented on a range of devices, facilitating large adoption and optimizing its influence on public safety..*

- Challenges:

  There are a number of challenges that need to be addressed in order to develop a successful crime and violence detection system, including:

  1. *Accuracy:* Detection of crime through a machine learning model is not always accurate. The accuracy depends on the variety of the data available while training the model. This can lead to false positive and false negatives.
  2. *Bias:* Detection algorithms can be biased. This can lead to inaccurate results for certain groups of people.
  3. *Privacy:* Due to collection of the surveillance footage and monitoring it through model it would have access to all the available surveillance feed. This raises privacy concerns.

- Research:

  Research efforts related to the identification of illegal activities using surveillance film are on the rise. These are diverse efforts that centre on the creation of new methods, the improvement of current strategies, and a deeper awareness of the fundamental ideas that underpin the identification of criminal behaviour. There is a great deal of promise for improving safety for everyone with this united research effort.

- Conclusion:

  The implementation of a criminal detection system represents an enormous boost in public safety. We can create a future in which communities feel safe and secure by using AI's potential for preemptive detection and rapid reaction. However, it is essential to examine ethical considerations about privacy. To guarantee that individual rights are protected, robust safeguards must be implemented. Finally, achieving a balance between security and privacy is critical for the effective application of this technology.

Some additional details that could be included are:

1. Multiple offences, including assault, theft, bullying, arson, and arrests, must be recognized by the system.
2. The system must be able to function with different public management systems and in a variety of settings, such as loud ones and ones with varying weather patterns.
3. The system must be accurate and stable.
4. Privacy must be given top priority in the system.

1.3 Scope of work

- **System design:** The system should be designed to be able to recognize a variety of crimes, including violence, arson, assault, arrest, theft, bullying, etc. The system should also be able to work in a variety of settings, such as noisy environments, different weather conditions, day and night etc.
- **Data Collection:** The system will need to be trained on a large dataset of criminal anomalies, weapons and non-criminal activities (background). The data can be collected from a variety of sources, such as public database, social media, surveillance footage etc.
- **Data Labelling:** The Data is in footage form; we need to convert it into images and label it in order to utilize the data for training the model. The data can be labelled using a variety of techniques, such as manually labelling the data, or using a service like Roboflow.
- **Model training:** The system will need to be trained on a machine learning model. This model can be trained using variety of techniques.
- System evaluation: the system will need to be evaluated to ensure that it is accurate and reliable, this evaluation can be done using a variety of metrics, such as accuracy, precision, and recall.
- System deployment: The system will need to be deployed in a production environment. This deployment can be done on a variety of platforms, such as cloud computing platforms (AWS), mobile devices, CCTV cameras, Surveillance rooms etc.

CHAPTER 2

REQUIREMENT ANALYSIS

2.1 Dataset for Crime Detection

Any machine learning model needs data to function, and this research understands how important data is to create a system that can effectively detect illegal activities. In order to do this, we will painstakingly choose an extensive dataset created just for our algorithm's training. Real-world video footage gathered from several sources will make up the majority of this collection, guaranteeing an accurate portrayal of criminal activity.

A properly labelled dataset is necessary for identifying criminal abnormalities in surveillance footage. This dataset will include a wide variety of video clips that show many types of criminal activity, including violence, assault, and arson. These movies will be carefully sorted into several categories according to the criminal conduct they depict in order to provide efficient training. Each class will have its own dedicated folder, which will help the model understand the special traits connected to each kind of abnormality.

Our study intends to combine weapon recognition skills in addition to illegal activity detection. To do this, the dataset will contain an extensive database of frequently used weaponry. I enable the algorithm to recognize firearms in surveillance film by training it on this extra data. The inclusion of weapon detection technology in the model enables it to identify criminal activity and evaluate possible danger levels in relation to the presence of weapons.

In order to expedite the data labelling process for our criminal activity detection system, this project makes use of RoboFlow, an intuitive annotation platform. The user-friendly interface of RoboFlow makes it easier to carefully curate a large dataset that is intended to train our YOLOv8 model.

The major source of real-world video footage that we will use is the well renowned UCF Crime Detection Dataset [4]. This dataset provides a wide range of video clips showing many types of criminal activity, such as aggression, assault, and arson. Through careful categorization of these films into separate RoboFlow categories, we prepare the model to understand the particular traits linked to each anomaly.

This project incorporates weapon identification capabilities, going beyond the detection of illegal conduct. In order to accomplish this, we will make advantage of RoboFlow's object identification labelling capabilities to build an extensive database of frequently used weaponry inside the dataset. The weapon database will provide precise and unambiguous labelling of different kinds of weapons, making it possible for the YOLOv8 model to recognize them in surveillance film . The model can identify criminal activity and determine possible danger levels depending on the presence of weapons by including weapon detection.

We recognize the value of a dataset that captures the variety of the actual world. In order to guarantee that the UCF Crime Detection Dataset gets expanded with more video material covering a broader range of races, genders, and locales, we will make use of RoboFlow's capabilities. This promotes equity across different demographic groups and lessens the possibility of biases in the model's detecting skills.

Our private video footage will be safeguarded during the annotating process since RoboFlow complies with industry-standard security procedures. Access control, encryption, and safe storage practices are some of these security precautions. In order to mask identifying information from the video footage, we shall, where needed, apply RoboFlow's anonymization features. This guarantees adherence to pertinent data privacy laws. [5]

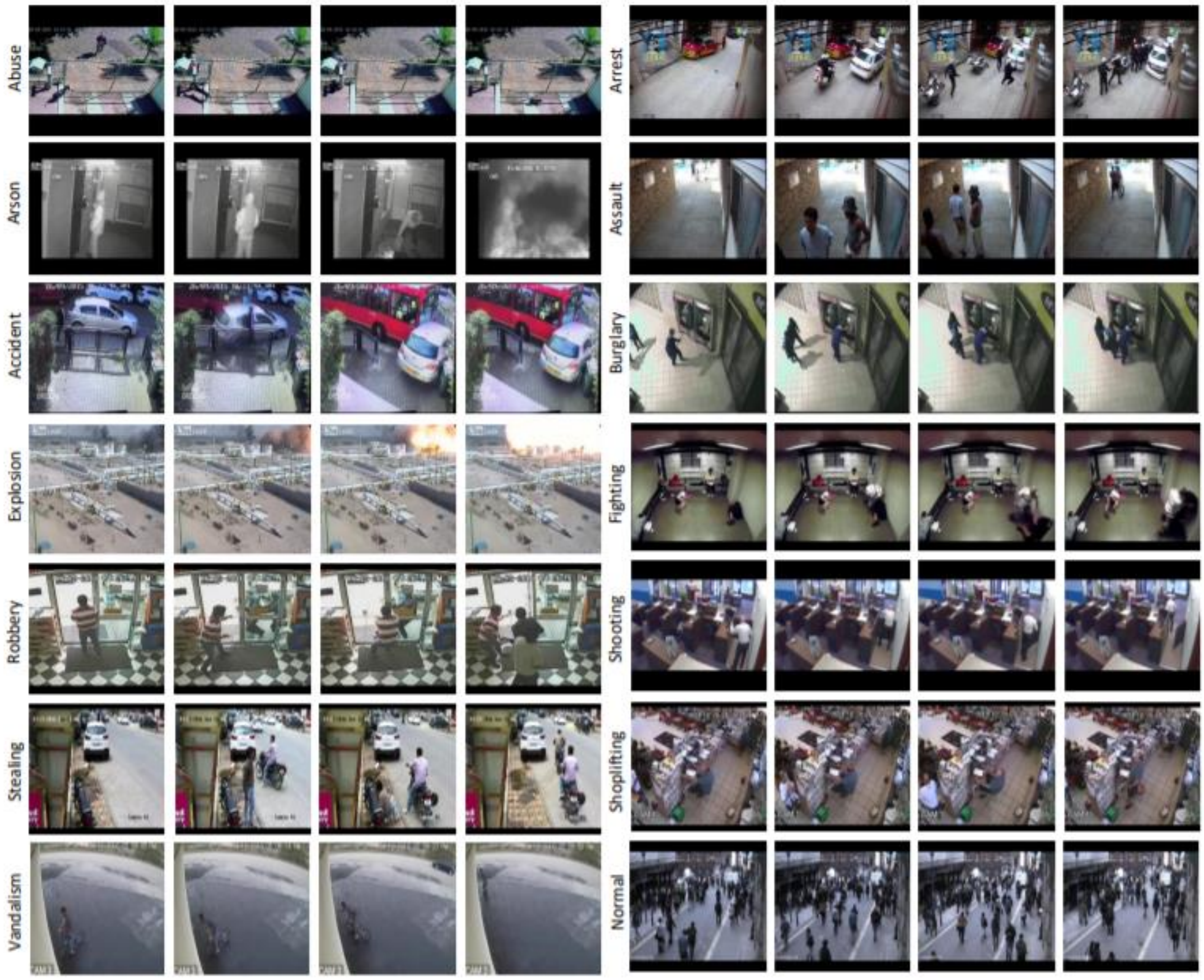

Figure 2: UCF Video Anomaly Detection Dataset [4]

1.2 Functional Requirements

- Crime detection and anomaly detection accuracy: The system should be able to correctly identify the anomaly with a high degree of accuracy.
- Crime anomaly detection robustness: The system should be able to correctly idemtify the anomalies in a variety of settings such as noisy environments and with people of different cultures.
- Dataset privacy: This system need to have a secure database because we are processing the surveillance footage of our country and we need to avoid any type of data leak or breach.

The main features of the criminal activity detection system that makes use of YOLOv8 and a carefully selected dataset are described in this section:

1. *Instantaneous Video Evaluation:* The security camera live video feeds will be continually analyzed by the system. The YOLOv8 object identification model will be used to recognize firearms and pre-defined criminal activity in the video frames.

2. *Identifying anomalies:* On the basis of the trained YOLOv8 model, the system will be able to identify different types of illegal activity. As examples, consider: Violence, Assault, Arson, Burglary Vandalism. Throughout the video, the system ought to be able to tell the difference between illegal activity and appropriate behavior.

3. *Identifying Weapons:* Within the video frames, the system must be able to recognize a predetermined list of frequently used weapons. The anomaly detection feature will be combined with this weapon detecting capabilities.

4. *Notification Mechanism:* When a weapon or pre-defined criminal conduct is successfully detected, the system will initiate an automatic notification procedure. Notification of the crime should be sent to the appropriate law enforcement agency in the area of the indicated offence.

5. *Configurability of the System:* The system ought to support configuration choices like, defining the kinds of crimes and weaponry that need to be looked for. establishing cutoff points for detection certainty before sending out alerts. indicating the recipients of notifications and the preferred means of transmission.

6. User Interface: The system could have a user interface (UI) that authorised staff members can utilise to, Watch chosen cameras' live video feeds. Examine previous notification records and detections. Control the settings of your system.

1.3 Non-Functional Requirements

- **Performance**: To achieve almost real-time detection and alerting, the system should run with the least amount of latency possible. Several video streams have to be supported by the system at once (depending on hardware capabilities). For the system to perform well across a range of computer platforms, it must be memory and processing power efficient.

- **Privacy and Security:** The system must put in place the proper security safeguards to safeguard private video footage and access control systems. The system's processing of surveillance footage should adhere to applicable data privacy standards.

- **Scalability:** Future extension, such as supporting more video feeds or interfacing with other security systems, should be supported by the system architecture.
- **Sustainability:** It is imperative that the system code be well-documented and modular to facilitate future modifications and maintenance.

1.4 Use-case scenarios

These scenarios showcase the potential applications of the criminal activity detection system in various situations. They highlight the importance of both real-time detection and the need for human oversight in some cases to handle false alarms or complex situations.

1. **Scenario 1: Street Robbery**

   A pedestrian is being mugged by a bunch of people on a street corner. The event is recorded live by the security camera. Within the video frame, the YOLOv8 model recognises the attack and possible weapon (such as a knife).

   The location and specifics of the alleged offence are included in an automated notice sent by the system to the local police department. After receiving the warning, law enforcement sends officers to the location in the hopes of helping the victim and maybe apprehending the culprits.

2. **Scenario 2: Vandalism in Progress**

   A public park's surveillance camera picks up people spray spraying vandalism on a statue. By using the trained model to identify the vandalism, the system notifies local authorities and park security staff.

   When security officers show up, they confront the vandals, who may run away or be caught.

3. **Scenario 3: False Alarm - Unusual Activity**

   Kids are roughhousing in the playground, and because the system can't tell differentiating between purpose and behavior, it interprets their actions as assault. When the activity is recognized by the system, a notification with a lower confidence level is sent out. After seeing the video, a human operator at a central monitoring station decides it's a false alarm, preventing the needless deployment of law enforcement.

4. **Scenario 5: System Maintenance and Updates**

   The system detects a gradual decline in performance or accuracy over time.
   System administrators schedule maintenance to update the YOLOv8 model with a new dataset reflecting recent criminal activity trends.

   The updated model improves detection accuracy and overall system effectiveness.

CHAPTER 3

SYSTEM DESIGN

3.1 Design Goals

3.1.1 End-to-End Crime detection pipeline: Comprehensive and holistic approach

This project's main objective is to develop a thorough and all-encompassing end-to-end crime detection pipeline that includes both weapon and violence detection. All pertinent logs and services should be included in the pipeline in order to offer a thorough analysis of the criminal data. Users may see patterns and trends by looking at the classes in the data set and their corresponding numbers. This helps them make informed judgements and take the necessary action to increase public security. The pipeline ought to delve deeply into the anomaly data, providing a thorough grasp of the illicit behaviors occurring inside the surveillance film, rather than just skimming the surface.

3.1.2 Accuracy and relevance: Focusing on essential criminal data

The criminal activity and violence detection project should emphasise the inclusion of relevant and accurate criminal activity data, which are required for accurate criminal activity recognition. Unnecessary input can distract and confuse the model, limiting its capacity to focus on key elements; hence, the pipeline should focus on communicating data that has a direct influence on the model's correctness. By supplying focused and precise criminal activity data, the pipeline has become an extremely helpful tool for boosting criminal activity detection accuracy.

3.1.3 Scalability: Adapting to increasing demand

The system must be scalable to meet the rising demands of the number of users as the application increases. The system should be built to manage a large volume of data while maintaining performance. The system should prioritise device resources, make optimum use of available resources, and be adaptable. Changes the format or structure of supplied data so that it stays relevant and helpful.

3.1.4 Modularity: Customization for different application

The system should be modular to enable customization for different application. The system should provide flexibility for users to select and view specific Criminal activities and their respective confidence levels. The system should provide an easy-to-use API to enable integration with different applications. The system shift provides comprehensive documentation and tutorials to facilitate easy adoption by developers.

3.2 Data Lifecycle Management

3.2.1 Data Collection:

I knew I needed a large variety of sounds to train my deepfake speech detection system on in order to create one that was actually successful. For this reason, I've cast a wide net on a variety of websites, including Research World, RoboFlow, Kaggle and even social media (with permission, of course!). With this method, the system may learn how individuals naturally communicate in a variety of contexts and from a wide range of backgrounds.

However, it's not simply random videos I've gathered. I have gathered a variety of videos and surveillance footage and images in order to make the database varied. We can train the system to recognise the minute details.

3.2.2 Data Storage:

First, there's offline storage. This serves as a safe vault, entirely isolated from the internet. It provides dependable backup and rapid access to the data anytime it is required for analysis or law enforcement inquiries. This physical isolation provides an added degree of security against hackers.

However, thieves may strike from anywhere and at any time, thus remote access is critical. This is where safe cloud storage comes in. It's like having a trusted digital partner keep a copy of the data, which is available from anywhere authorized workers need it. This enables quick analysis and model training, resulting in a more effective crime detection system.

3.2.3 Data Processing

Once acquired, the raw video data will go through a critical preparation phase. This step will involve cleaning, organizing, and changing the data to ensure that it is compatible with the machine learning techniques we have chosen.

Converting video formats involves ensuring consistency across all video files.
Resizing and cropping: Standardizing video dimensions to ensure efficient processing.
Frame extraction is the process of removing individual frames from video sequences in preparation for analysis.

3.2.4 Data Labelling

Data labeling involves manually categorizing specific scenes or objects in video frames to identify illicit actions. We are using RoboFlow for data annotation and labelling.

3.2.5 Data Augmentation:

To improve the model's generalizability and reduce over fitting, we may use data augmentation approaches.

Flipping/rotating involves creating mirrored or rotated versions of existing frames.
Adding noise simulates real-world variances such as illumination shifts and camera shake.

CHAPTER 4

METHODOLOGY

4.1 Detailed data preprocessing

We gathered enormous amounts of CCTV footage from the UCF Crime Detection Dataset at the University of Central Florida. Imagine piles of video footage of assaults, thefts, vandalism, and other forms of criminal behaviour. But here's the thing: YOLO, our deep learning algorithm, does not learn directly from videos. It requires bite-sized chunks of information, such as individual photos.

So our first step was to break down a large movie into a picture book. We retrieved every frame from each video, resulting in a massive picture collection. Now, these weren't simply random photos. They all fit into particular categories: abuse, arson, violence, you name it. Each photograph was properly organised in its own folder, making it easy for the system to interpret.

But here's the second hurdle: these photographs weren't yet ready for YOLO to learn from. They need labels, such as little tags that explained what was happening in each photo. Imagine presenting a youngster a picture of a dog and telling them, "See, that's a fluffy friend!" That is basically what labelling does for the system.

Here's where RoboFlow comes in. It's a clever platform that employs artificial intelligence and pre-labeled photos to assist us with labelling the complete dataset. Consider having a super-powered helper who can learn from a few samples and then classify the other photos quickly.

Here's how it worked: I chose certain frames from the films and painstakingly labelled them, demonstrating to RoboFlow what a robbery and vandalism looked like, respectively. Then, RoboFlow utilised my samples to train its own "eye" and started classifying the remainder of the photos automatically. It was like having an entire team of helpers working nonstop to prepare our data for prime time!

By converting raw film into labelled pictures, we effectively built a massive learning library for our crime-fighting system. Now, YOLO can finally begin training on all of this valuable data and become a great tool for detecting illegal activities in real time.

I began by handpicking crucial frames from the movies and methodically labelling them to demonstrate RoboFlow what a robbery looked like, distinguishing it from vandalism or violence. RoboFlow then utilised these labelled samples to train its own "eye" and began labelling the other pictures automatically. It was like having an entire team of helpers working nonstop to tag the evidence!

By converting raw film into a library of labelled pictures, we effectively created a gigantic training field for YOLO. Now, YOLO can finally analyse this vast amount of data, learning to identify illegal activities in real-time film with the precision of a trained detective. This data preparation guarantees that YOLO has the strongest possible basis to become a valuable tool for community safety.

4.2 Training Phase

Imagine being a hotel security guard with superpowers. You can scan the entire lobby in seconds and identify anything questionable. That is the purpose of our project: to provide CCTV cameras with this type of eagle eye without the need for a cape.

We began by collecting a massive amount of hotel video and methodically organising it for a unique AI programme dubbed YOLOv8. YOLOv8 comes in a variety of powers, including a magnifying glass (small model) and a strong telescope (big model). We tried each one on our footage to discover which one performed best.

We wanted our system to be quick and precise. Like a good detective, it must react rapidly while also doing things properly. So we chose the leanest and meanest YOLOv8, the YOLOv8-S, for the cameras. It's like a superhero's sidekick, agile and capable of detecting threats on the run.

But, just as detectives require backup, our system includes one additional layer. When the YOLOv8-S in the camera detects something suspicious, it transfers the video to a nearby data centre. The YOLOv8-M or YOLOv8-X variants, which are larger and more powerful, are introduced here. These are similar to the forensics squad, taking a deeper look and making sure nothing falls through the cracks.

After gathering and converting all of the data into the proper YOLO format, the following stage in our research was to train our model on the custom dataset that we created. We used Ultralytics' YOLOv8 model, which is available in three distinct sizes: YOLOv8-X, YOLOv8-M, and YOLOv8-S. We extensively evaluated and trained each of these models using our dataset to assess their performance.

But for the highest level of accuracy, our YOLOv8-S receives support, much as detectives require assistance from a forensics team. This is where the architecture with two tiers is used. Like a first responder, the YOLOv8-S on the camera flags anything shady. After that, this video is forwarded to data centres, where even more potent YOLOv8 models (M or L) examine it in more detail to make sure nothing is missed.

The system's strength lies in its two-step methodology. It's like having an extremely vigilant security guard followed closely by a group of skilled investigators for a close inspection. In this manner, we guarantee that no suspicious behaviour in a hotel remains undetected by achieving both lightning-fast detection and pinpoint precision.

In order to maximise performance, our models underwent many training cycles. Initially, 50 epochs were used to train the models. Upon examining the training graphs, we observed that the loss was steadily declining with every epoch, suggesting that more training may improve the models even more. We thus extended the training by 20 epochs, and then we continued the training for an additional 20 epochs while keeping a careful eye on the loss and other performance metrics. We were able to significantly reduce the loss with each subsequent training session because to this iterative method, which helped us fine-tune the models and produce highly reliable models.

Our system is reliable since we have extensive validation and testing protocols in place. To ensure that the models don't overfit and can successfully generalise to fresh, untested data, we tested them on a different validation set. Additionally, in order to validate the models' functionality, we tested them in real time using CCTV footage.

We can achieve a balance between high precision and real-time processing by utilising this tiered approach. Rapid response is facilitated by the YOLOv8-S model's initial detection in CCTV cameras, and superior accuracy and reliability are ensured by larger models' subsequent processing in data centres. We are able to meet critical performance and efficiency requirements while providing a reliable solution for real-time anomaly identification and alerting in hotel surveillance systems thanks to our comprehensive technique.

4.3 User Interface

Our device also need a highly interactive, user-friendly interface to ensure proper operation. The interface will be built for simplicity and intuitiveness, allowing users to effortlessly post video and image materials. When a user uploads a file via the dashboard, it is immediately saved in a "uploads" subdirectory within the directory.

Once the file is saved in the uploads folder, our model takes control. It thoroughly scans the submitted picture or video file for any irregularities. The model is intended to be extremely efficient and precise, employing powerful algorithms to detect and highlight irregularities. For example, if a user uploads a picture, the model will analyses it and draw borders around the discovered classes.

After processing the file, the programmed will save any discovered abnormalities in a separate "runs" folder. This folder stores all processed files, ensuring that any found abnormalities are organized and easily accessible. The abnormalities, once recognized and recorded, will be shown on our dashboard, giving users with a clear and complete visual representation of the results.

To elaborate, suppose you submit a picture using the dashboard. The photograph is first kept in the uploads folder. The model then downloads the image from this folder, analyses it, and processes it by creating borders around any abnormalities found. Once this procedure is completed, the corrected picture, with the anomalies indicated, is stored in the runs directory. The programme then retrieves the processed image from the runs folder and shows it on the dashboard. This guarantees that users may view the abnormalities found in their files in a simple and organised way.

The interface will also allow users to access information about the observed abnormalities. By clicking on the presented anomalies, users can obtain further information such as the type of anomaly, its position within the file, and any suggested actions. This functionality guarantees that users not only notice abnormalities, but also comprehend the context and relevance of each detection.

Furthermore, the dashboard will be built to support a variety of file types and sizes, ensuring adaptability and flexibility. Users will be able to submit a variety of file kinds without concern about compatibility difficulties. The interface will also contain features such as drag-and-drop file upload, progress indications, and notifications to improve the user experience.

CHAPTER 5

RESULT AND ANALYSIS

5.1 Yolo Architecture Analysis

Yolo V8-S

- Compactness and Speed: Yolo V8-S is the most compact and fastest among the iterations. It’s the ideal choice for applications prioritizing speed over accuracy. [7]

Yolo V8-M

- Accuracy: Yolo V8-M offers a step up in accuracy compared to V8-S. While it operates at a slightly slower pace, it’s the perfect fit for scenarios where precision is paramount. [7]

Yolo V8-L

- Precision: Yolo V8-L takes accuracy to the next level but sacrifices speed in the process. It shines in applications where achieving the highest levels of accuracy is imperative. [7]

This two-step approach is key. The YOLOv8-S in the camera ensures there's no delay in catching suspicious activity, while the data center models provide the ultimate accuracy. It's like having a quick response team followed by a team of seasoned investigators – the best of both worlds.

The hidden weapon in our system is the YOLOv8-S model. Consider it a highly trained security guard with superhuman abilities, such as rapid speed. Because it is the smallest and quickest YOLOv8 version, it can scan CCTV footage in real time and detect suspicious activities practically immediately. This is critical for catching things on the go, just as a skilled security guard would.

However, much like certain detectives require assistance from a forensics team, our YOLOv8-S receives assistance from its larger brothers, the YOLOv8-M and maybe the YOLOv8-L. These models represent the forensics team, sacrificing speed for maximum precision. They take a closer look at whatever the YOLOv8-S flags to ensure that nothing sneaks past.

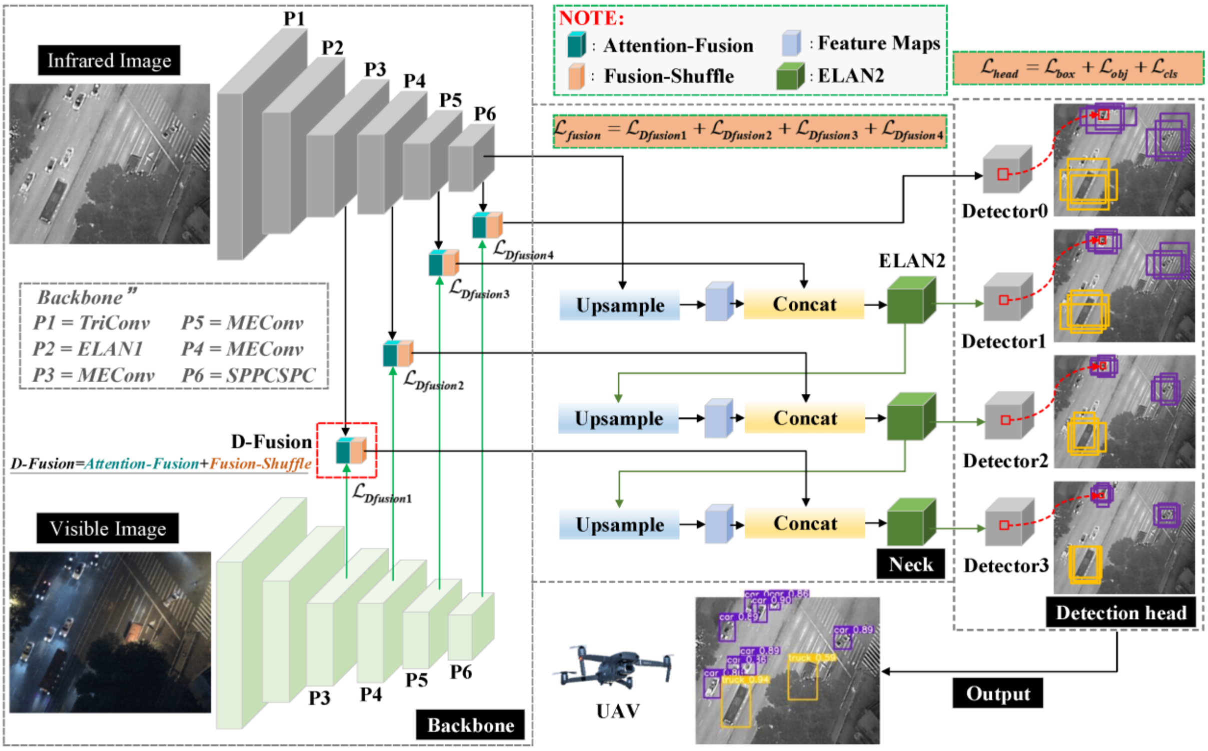

Figure 3. YoloV8 Architecture

5.2 Training Results

The training went swimmingly. YOLOv8s made significant development during training.

Here's a breakdown of what we saw:

1. Loss on a Downward Spiral: The loss function decreased with each training period for all models. This suggests that they were gradually improving their ability to identify irregularities in hotel footage. It's as if they were learning from their mistakes and getting more precise with each iteration.

2. More epochs, more improvements: We didn't just stop with one training run. We intentionally increased the number of epochs (training cycles) based on how the loss performed. This rigorous technique guaranteed that the models were not overtrained and continued to improve with each iteration.

3. Highly Reliable Results: The end result was quite satisfactory. All of the models showed a large reduction in loss, showing a high capacity to detect abnormalities effectively.

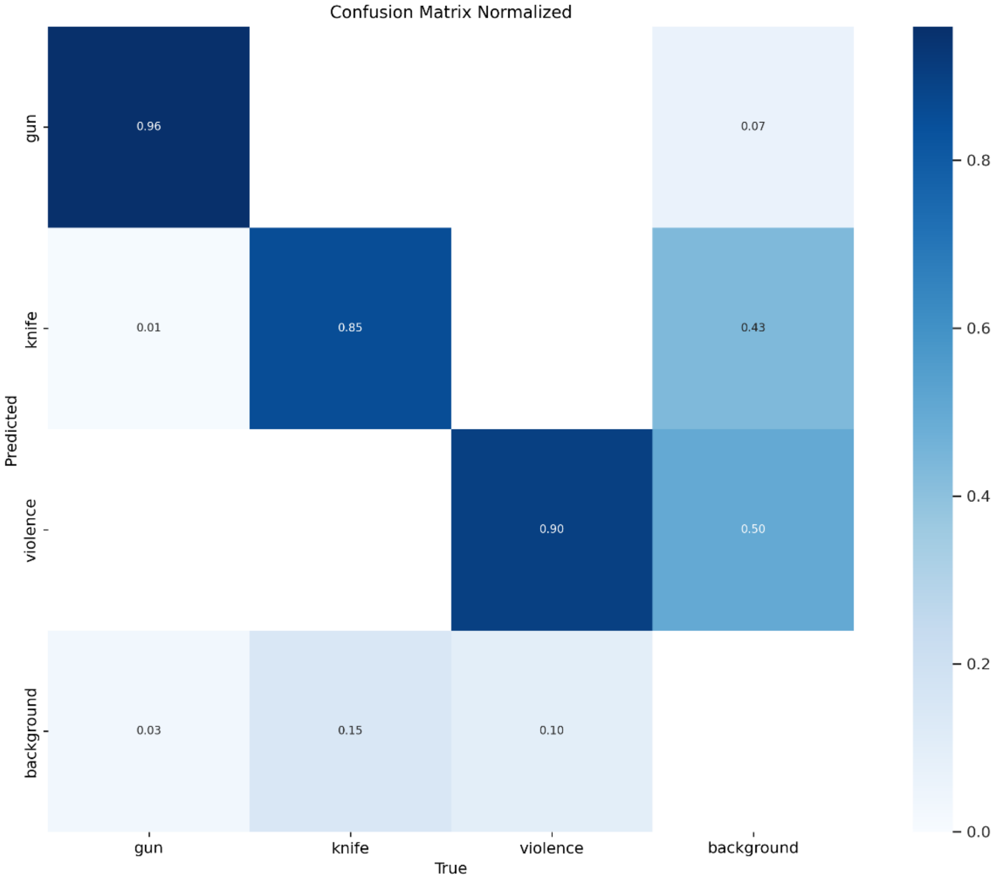

Figure 4. Confusion Matrix

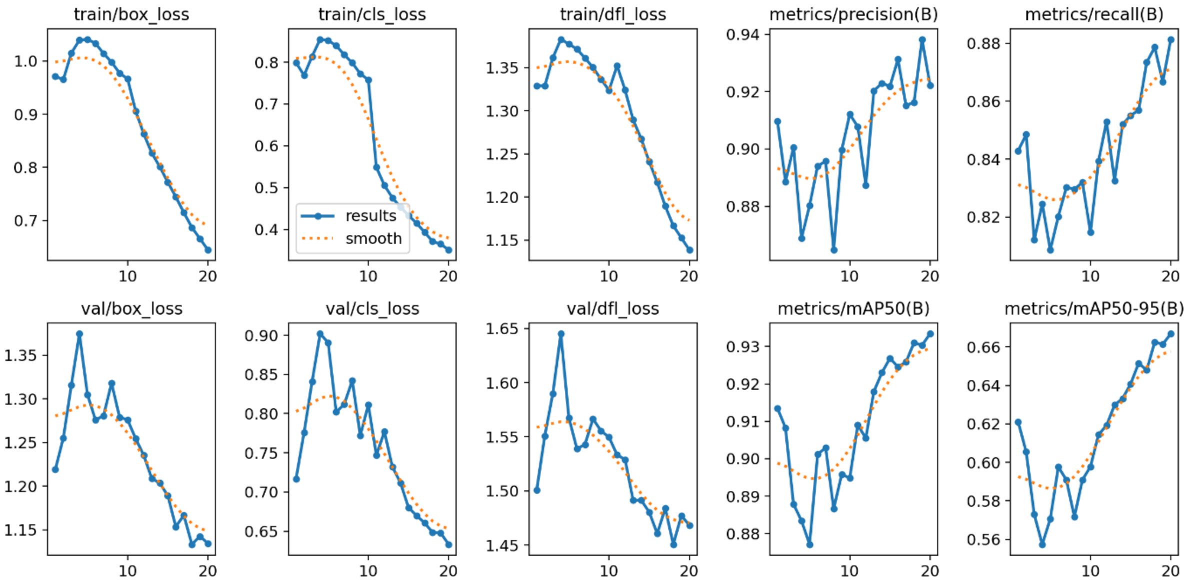

Figure 5. Results

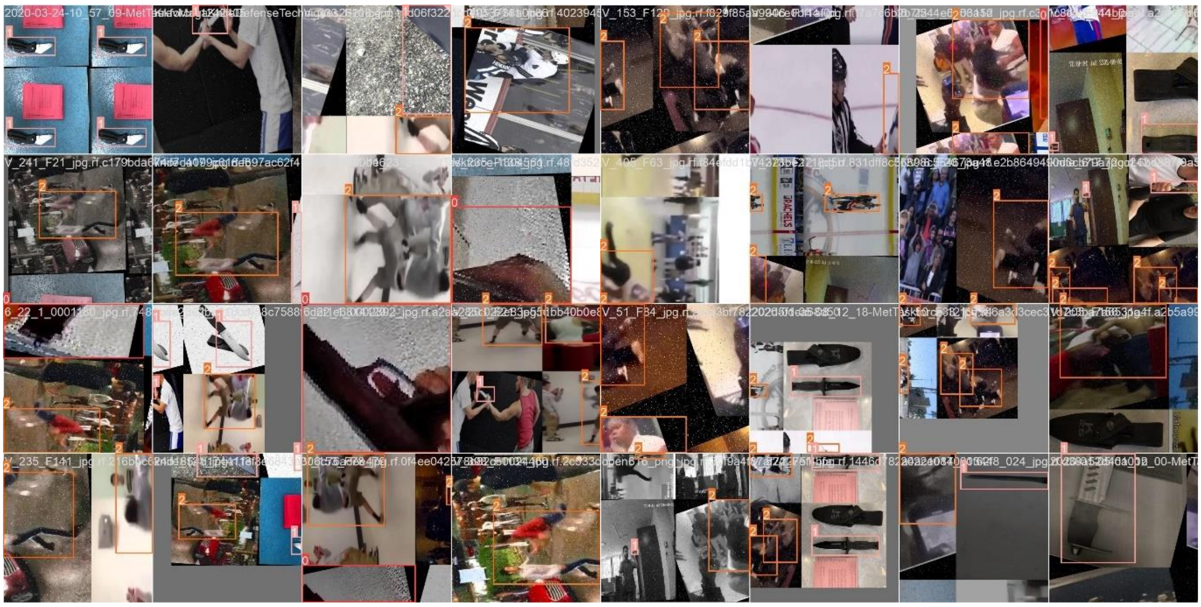

Figure 6. Training Batch 1

CHAPTER 6

CONCLUSION & FUTURE SCOPE

6.1 Conclusion

Our research intended to develop a real-time security system with eagle eyes for hotels. We utilised YOLOv8, a very intelligent AI model that can recognise objects in films. We made great progress by methodically training it on a mountain of recorded hotel footage.

The main problem was achieving the ideal blend of speed and precision. We evaluated three YOLOv8 models, much like having trainees with varying strengths. While all three learnt to detect abnormalities in the film with outstanding results (their training scores continued to improve!), the YOLOv8-S was the fastest learner. This quickness is critical for detecting suspicious activities on the move, just like a hotel security officer might.

But, much as detectives require assistance from a forensics team, our YOLOv8-S receives assistance for maximum accuracy. This is where the two-tiered architecture comes in. The YOLOv8-S on the camera functions as a first responder, flagging anything suspicious. This film is then uploaded to data centres, where even more sophisticated YOLOv8 models (M or L) take a closer look, ensuring that nothing gets through.

The system's effectiveness stems from its two-step methodology. It's like having a security guard on high alert, followed by a team of experienced detectives conducting a comprehensive investigation. This manner, we accomplish both lightning-fast detection and pinpoint precision, guaranteeing that no suspicious behaviour goes undetected at a hotel.

6.1 Future Scope of Work

The future of this project is brimming with exciting possibilities that could take hotel security to even greater heights. One potential avenue for advancement is the development of a three-tiered architecture. Imagine this: strategically placed edge servers throughout the hotel complex would house mid-range models like the YOLOv8-M. This would create a finely tuned network of processing power, allowing for even faster response times. Think of it as having security guards patrolling every hallway, with a rapid response team stationed nearby for immediate intervention.

Furthermore, collaboration with government agencies could be a game-changer. By working together, we could establish a standardized approach to AI-powered hotel security systems. This collaboration wouldn't just be about sharing best practices or datasets; it would be about pooling expertise to create an industry-wide security framework, robust and secure enough to keep everyone safe. Imagine a world where hotels around the globe utilize this cutting-edge technology, all thanks to collaborative efforts.

By continuously pushing the boundaries of innovation and exploration, we can refine this system even further. This not only enhances hotel security but also fosters a safe and secure environment for both guests and staff. Ultimately, our goal is to create a hospitality haven where everyone feels comfortable and protected, and this project is a significant step towards achieving that vision.

BIBLIOGRAPHY

1. Crime Rate by Country 2024 : World Population Review
   https://worldpopulationreview.com/country-rankings/crime-rate-by-country
2. Rising Crime Against Indian Women
   https://www.bbc.com/news/world-asia-india-62830634
3. Figure 1: Homicide rates in India – Violence Info World Health Organization
   https://apps.who.int/violence-info/country/IN
4. Real World Anomaly Detection In Surveillance Videos : University of Central Florida
   https://www.crcv.ucf.edu/research/real-world-anomaly-detection-in-surveillance-videos/
5. Roboflow Labelling Documentation
   https://docs.roboflow.com/annotate/use-roboflow-annotate
6. YOLOv8 Architecture Explained
   https://yolov8.org/yolov8-architecture-explained/
7. Unveiling the Power of Yolo V8
   https://medium.com/@khizarirfan/working-principle-of-yolo-v8-and-explain-the-difference-between-variants-of-yolo-v8-46009b5eede0
8. "Real-Time Object Detection for Anomaly Detection in Video Surveillance" by Ma, S., Xu, Y., & Liu, S. (2018):

   https://www.researchgate.net/publication/224258100_Real-time_camera_anomaly_detection_for_real-world_video_surveillance

9. "YOLOv5: Detecting Objects in Real-Time" by Joyson, A., Luo, J., Kong, X., & Fu, C. (2020):
   https://arxiv.org/pdf/2104.13634
10. "YOLOv5-Based Real-Time Video Anomaly Detection for Suspicious Activity Monitoring" by He, Z., Deng, S., Cao, Y., & Wang, X. (2021):
    https://arxiv.org/pdf/2104.13634
11. "Learning Temporal Regularity in Video Sequences" - CVPR 2016 Paper by Mahmudul Hasan et al.: This paper focuses on detecting anomalies by learning the temporal regularity in video sequences using a convolutional neural network (CNN) and a long short-term memory (LSTM) network.
12. "A Comprehensive Review on Video Anomaly Detection with Deep Learning" - IEEE Access 2020 by Zaigham Mahmood et al.: This review paper provides an extensive overview of various deep learning techniques for video anomaly detection. It discusses different methodologies, datasets, and the challenges faced in real-world applications.