

**VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI, KARNATAKA**



A Project Work Phase2 Report On

**“VigiScape: Elevating Public Safety for Crime Prevention with AI
powered Smart Surveillance”**

Submitted in the partial fulfillment for the requirements for the conferment of Degree of

**BACHELOR OF ENGINEERING
in
INFORMATION SCIENCE AND ENGINEERING**

By

Ms. MANASA C B

USN: 1BY20IS077

Under the guidance of

Dr. Gireesh Babu C N
Asst. Professor
Dept. of ISE, BMSIT&M



BMS INSTITUTE OF TECHNOLOGY & MANAGEMENT
YELAHANKA, BENGALURU-560064
DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING



2023-2024

**VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI, KARNATAKA**



BMS INSTITUTE OF TECHNOLOGY & MANAGEMENT
YELAHANKA, BENGALURU-560064
DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING



CERTIFICATE

This is to certify that the Project work Phase-2(18CSP83) entitled “**VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance**” is a bonafide work carried out by **Ms. Manasa C B (1BY20IS077)** in partial fulfillment for the award of **Bachelor of Engineering Degree in Information Science and Engineering** of the Visvesvaraya Technological University, Belagavi during the year 2023-24. It is certified that all corrections/suggestions indicated for Internal assessment have been incorporated in this report. The project report has been approved as it satisfies the academic requirements with respect to project work for the B.E Degree.

Signature of the Guide
Dr. Gireesh Babu C N

Signature of the Coordinator
Mrs. Chethana C

Signature of the HOD
Dr. Pushpa S K

Signature of the Principal
Dr. Sanjay H A

ABSTRACT

Artificial intelligence (AI) has significantly reshaped public safety through the advent of AI-powered smart surveillance systems, offering a promising avenue to enhance security and diminish crime rates. This survey navigates the realm of AI-powered smart surveillance, defining its scope, evaluating existing technologies, and exploring its practical applications in crime prevention and emergency response scenarios.

The study confronts critical issues concerning privacy, biases, and data security that profoundly impact the integration and acceptance of AI-powered surveillance systems. It also speculates on future prospects for AI-powered smart surveillance, emphasizing the paramount importance of responsible development and ethical utilization.

The exploration of AI-powered smart surveillance emphasizes potential advantages, such as heightened crime detection capabilities and expedited emergency responses. However, the research underscores the crucial necessity for robust legislative measures and policies aimed at addressing privacy concerns, mitigating biases inherent in AI algorithms, and fortifying data security protocols.

It stresses the imperative for a balanced approach that harnesses the potential of the technology while safeguarding individual rights and upholding ethical standards. Ultimately, the study advocates for the responsible advancement and conscientious application of AI-powered smart surveillance to ensure its positive impact on public safety.

ACKNOWLEDGEMENT

We are happy to present this project phase -2 after completing it successfully. This project would not have been possible without the guidance, assistance and suggestions of many individuals. We would like to express our deep sense of gratitude and indebtedness to each and every one who has helped us make this project a success.

We heartily thank **Dr. Sanjay H A, Principal, BMS Institute of Technology & Management** for his constant encouragement and inspiration in taking up this project.

We heartily thank **Dr. Pushpa S K, HoD, Dept. of Information Science and Engineering, BMS Institute of Technology & Management** for her constant encouragement and inspiration in taking up this project.

We heartily thank project coordinator, **Mrs. Chethana C, Assistant Professor, Dept. of Information science and Engineering**, for her constant follow up and advice throughout the course of the project work.

We gracefully thank our project guide, **Dr. Gireesh Babu C N, Asst. Professor, Dept. of Information Science and Engineering**, for his encouragement and advice throughout the course of the project work.

Special thanks to all the staff members of Information Science Department for their help and kind co-operation.

Lastly, we thank our parents and friends for their encouragement and support given to us in order to finish this precious work.

By,
Manasa C B



Declaration

We, hereby declare that the project phase -2(18CSP83) titled “**VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance**” is a record of original project phase - 1 work undertaken for the award of the degree Bachelor of Engineering in Information Science and Engineering of the Visvesvaraya Technological University, Belagavi during the year 2023- 24. We have completed this project phase -2 work under the guidance of **Dr. Gireesh Babu C N**.

We also declare that this project phase -2 report has not been submitted for the award of any degree, diploma, associate ship, fellowship or other title anywhere else.

Student

Photo



USN 1BY20IS077

Name Manasa C B

Signature

INDEX

ABSTRACT	i
ACKNOWLEDGEMENT	ii
DECLARATION	iii
INDEX	iv
LIST OF FIGURES	vi
LIST OF TABLES	vii

Chapter No.	Chapter Title	Page No.
1	Introduction	[1-6]
1.1	Preamble	1
1.2	Motivation	2
1.3	Existing System	3
1.4	Problem Statement	3
1.5	Proposed System	4
1.6	Objectives	5
2	Literature Survey	[7-15]
2.1	Critical Analysis of Literature Survey	7
2.2	Summary	9
3	System Requirement Specification	[16-21]
3.1	Functional Requirements	16
3.2	Non-Functional Requirements	17
3.3	Software Requirements	18
3.4	Hardware requirements	19
3.5	Interface Specifications	19
3.6	User Requirements	20
3.7	Use Case Diagram	21
4	System Design	[22-26]
4.1	Methodology	22
4.2	System Architecture	23
4.3	Algorithms/Pseudo code	24
4.4	Class Diagram	26

5	Implementation	[27-31]
5.1	Service	27
5.2	Tools and Library Packages	28
5.3	Programming Languages	30
5.4	Steps of Implementation	31
6	Testing and Validation	[32-35]
6.1	Unit Testing	32
6.2	Integration Testing	33
6.3	System Testing	34
6.4	User Acceptance Testing	35
7	Results and Discussions	[36]
8	Conclusion and Future Enhancement	[41-42]
8.1	Conclusion	41
8.2	Limitation	41
	References	[43]
	Appendix A	[45]
	Plagiarism Report	46

LIST OF FIGURES

Figure No.	Figure Name	Page No.
3.1	Use Case diagram	21
4.1	UML class diagram	25
7.1	Low Likelihood of Violence: Green Border Frame	36
7.2	Moderate Likelihood of Violence: Yellow Border Frame	37
7.3	High Likelihood of Violence: Red Border Frame	37
7.4	Dynamic Threshold Analysis for Real-time Violence Detection	38
7.5	Emergency Message Dispatch with Detected Violence Frame	39
7.6	Real-time Camera Capture with No Violence Detected	40
7.7	Real-time Camera Capture with Violence Detected	40

LIST OF TABLES

Table No.	Table Name	Page No.
2.1	Literature Survey Summary Table	15
6.1	Unit Testing	31
6.2	Integration Testing	32
6.3	System Testing	32
6.4	User Acceptance Testing	33

CHAPTER 1

1. INTRODUCTION

1.1 Preamble

Artificial intelligence (AI) is rapidly influencing the realm of public safety, with AI-powered smart surveillance systems emerging as a promising tool for improving security and reducing crime. The need for public safety is expanding as the complexity of security threats in today's society grows. Maintaining public security is a significant issue for governments, law enforcement agencies, and communities, from traditional crimes like theft and assault to rising dangers like cyberterrorism and extremist violence.

AI has emerged as a transformative force in a variety of industries, including public safety. AI-powered smart surveillance systems analyse enormous amounts of data using AI algorithms, enabling for real-time monitoring, anomaly detection, and predictive analytics. This method has been acknowledged for its ability to boost security and reduce crime.

The application of AI in smart surveillance systems provides numerous benefits for public safety. A variety of studies have shown that AI can facilitate seamless identification and authentication of authorized personnel on campus premises, effectively deterring unauthorized access and mitigating security breaches and that AI-powered surveillance systems can recognize and track objects with remarkable accuracy, allowing law enforcement agencies to monitor activity in public spaces and identify criminals. This increased situational awareness has the ability to considerably aid in crime prevention and incident response.

AI-powered smart surveillance systems provide numerous public safety benefits, notably in surveillance video processing. Deep learning algorithms can identify a wide range of objects, including persons, vehicles, and suspicious items, providing important information to law enforcement and security professionals. Furthermore, it delves into the policy and technical considerations connected to the implementation of AI-powered smart surveillance systems, such as the significance of establishing clear guidelines and policies governing data collection, storage, usage, and privacy protection to ensure accountability and ethical system implementation

1.2 Motivation

The motivation for implementing AI-powered smart surveillance systems stems from the pressing need to address the escalating frequency of crime and safety concerns in public spaces. Traditional surveillance systems often fall short in providing real-time alerts, thereby impeding prompt responses to potential threats. This gap underscores the urgency for advanced surveillance solutions capable of proactively identifying and notifying authorities about unusual or suspicious activities, thus bolstering public safety.

Recognizing this imperative, there is a growing acknowledgment of the necessity to bridge the existing shortcomings of conventional surveillance technologies. These limitations hinder the ability to swiftly detect anomalies and potential risks in real-time scenarios. Consequently, there is a clear call for the deployment of cutting-edge technologies such as artificial intelligence (AI) and machine learning (ML) algorithms to revolutionize surveillance networks.

The integration of AI and ML into surveillance systems offers a multifaceted approach to addressing security challenges. These technologies empower surveillance networks to analyze complex data patterns, enabling rapid detection of suspicious behavior or objects. Moreover, AI algorithms can facilitate seamless identification and authentication of authorized personnel, effectively deterring unauthorized access and mitigating security breaches.

By leveraging AI-powered smart surveillance systems, law enforcement agencies gain invaluable tools for enhancing public safety. These systems provide enhanced situational awareness, enabling authorities to monitor activity in public spaces with remarkable accuracy. Moreover, they facilitate the swift identification of criminals, thereby aiding in crime prevention and incident response.

In conclusion, the motivation for implementing AI-powered smart surveillance systems lies in the imperative to enhance public safety by addressing the limitations of traditional surveillance technologies. By harnessing the capabilities of AI and ML, these systems offer advanced functionalities crucial for swift detection and response to potential threats in real-time scenarios.

1.3 Existing Systems

Few of the existing systems includes:

- **AI-Driven Smart Surveillance System:** This system is capable of detecting and tracking objects in surveillance footage, contributing to improved public-space security. It uses AI algorithms to evaluate video data and identify potential risks, allowing law enforcement to adopt proactive crime prevention measures.
- **Railway Crossing Safety System:** Another system focuses on the use of artificial intelligence in railway crossing safety. It employs cameras and sensors to detect vehicles and people at railway crossings, delivering real-time alerts to drivers and pedestrians, thus averting accidents and ensuring railway operations are safe.
- **Intelligent Vision-Based Public Safety Monitoring System for Campus Areas:** This system employs cameras and sensors to detect and track people and vehicles on campus, providing real-time monitoring capabilities to security personnel and enabling prompt responses to security incidents.

Some of the limitations of the existing systems includes:

- **Data privacy and security:** The collecting and storing of massive volumes of surveillance data raises privacy and security concerns.
- **Algorithmic bias:** AI algorithms can be biased, resulting in unjust or discriminating outcomes.
- **Understanding of AI decisions:** The difficulty to understand and trust AI decisions due to a lack of transparency and comprehensibility.
- **Ethical problems:** The employment of artificial intelligence-powered smart surveillance systems poses ethical concerns concerning individual privacy, public safety, and social control.

1.4 Problem Statement

The escalating frequency of crime and safety concerns in public spaces highlights the shortcomings of traditional surveillance systems, which often fail to provide real-time alerts, hindering prompt responses to potential threats. There is an urgent need for an advanced surveillance system capable

of proactively identifying and notifying authorities about unusual or suspicious activities to bolster public safety. Establishing such a system is crucial to address the limitations of current surveillance technologies, enabling the swift detection of anomalies and potential risks in real time. Incorporating cutting-edge technologies like artificial intelligence (AI) and machine learning (ML) algorithms is essential to empower surveillance networks to analyze complex data patterns and respond swiftly to ensure the safety and security of individuals in public areas.

1.4 Proposed Systems

Our project entails utilizing video input for real-time violence detection. The methodology is outlined as follows:

- **Video Input:** Video streams serve as the primary input source for our system.
- **Preprocessing:** The incoming video is pre-processed into individual frames to facilitate analysis.
- **Fine-tuning CLIP Model:** The CLIP (Contrastive Language-Image Pretraining) model is employed for its robustness in understanding image-text pairs. We fine-tune the CLIP model with the pre-processed dataset obtained from the video frames. This step enhances the model's ability to recognize violence-related content effectively.
- **Model Saving:** Upon fine-tuning, the CLIP model is saved, ensuring efficient deployment and reusability.
- **Camera Initialization and Threshold-based Analysis:** When the camera initializes, it captures video frames continuously. Each frame undergoes analysis based on predefined thresholds. If the violence score, as determined by the CLIP model, exceeds 0.75, the border colour is set to red. If the score falls between 0.4 and 0.75, the colour changes to yellow. Otherwise, the border colour remains green.
- **Emergency Message Dispatch:** In case of a red border (indicating a high likelihood of violence), an emergency message is promptly dispatched. This message serves as an alert, signalling the detection of potential violence, enabling timely intervention.

This methodology combines preprocessing, machine learning model utilization, and real-time threshold-based analysis to detect violence accurately and swiftly, thus contributing to enhanced safety and security measures.

1.5 Objectives

The objectives of the project can be outlined as follows:

- **Development of AI-Powered Smart Surveillance System:**

Design and develop an advanced surveillance system leveraging artificial intelligence (AI) and machine learning (ML) algorithms. Implement real-time anomaly detection capabilities to proactively identify and notify authorities about unusual or suspicious activities in public spaces.

- **Enhancement of Public Safety:**

Improve public safety by addressing the limitations of traditional surveillance systems through the deployment of cutting-edge technologies. Provide law enforcement agencies with invaluable tools for enhancing situational awareness, monitoring activity in public spaces, and facilitating swift identification of criminals.

- **Integration of AI and ML Technologies:**

Integrate AI and ML algorithms into surveillance networks to enable the analysis of complex data patterns and enhance the detection of suspicious behaviour or objects. Develop algorithms for seamless identification and authentication of authorized personnel, deterring unauthorized access, and mitigating security breaches.

- **Addressing Existing System Limitations:**

Mitigate data privacy and security concerns associated with the collecting and storing of massive volumes of surveillance data. Address algorithmic bias issues by developing algorithms that are transparent, fair, and unbiased. Enhance the understanding and trustworthiness of AI decisions by improving transparency and comprehensibility. Address ethical concerns surrounding individual privacy, public safety, and social control by implementing robust privacy protection measures and ethical guidelines.

- **Swift Detection and Response to Threats:**

Enable the swift detection of anomalies and potential risks in real-time scenarios, facilitating prompt responses to security incidents.

Develop evaluation frameworks to continuously assess the performance and accuracy of anomaly detection methodologies, enabling continuous improvement of the system.

Overall, the objectives of the project aim to leverage AI and ML technologies to develop an advanced surveillance system that enhances public safety, addresses the limitations of existing surveillance systems, and ensures the swift detection and response to potential threats in real-time scenarios.

CHAPTER 2

2. Literature Survey

2.1 Critical Analysis of Literature Survey

The project work commenced with a comprehensive review of 16 papers that explored the application and the limitations of existing surveillance systems using AI model and machine learning. Out of this pool, three papers were singled out as notable contributions, forming the foundation or base papers for the subsequent project work. This initial selection process involved rigorous evaluation criteria to identify studies that offered significant insights, methodologies, or results in the context of machine learning to enhance threat detection capabilities while ensuring privacy protection within monitored public spaces.

The decision to focus on these three papers as base papers indicates a strategic approach to distill relevant and high-quality information from a broader body of literature. These selected papers played a pivotal role in shaping the direction of the project, providing valuable methodologies, algorithms, or findings that contributed to the study's objectives.

It can be inferred that they were chosen based on their methodological robustness, relevance to the research questions, and potential to contribute to the overarching goal of improving leukemia diagnosis through machine learning. The subsequent detailed analysis, then delved into the methodologies, algorithms, and findings of these base papers to extract meaningful insights for the overall project work.

The three articles collectively highlight the growing role of advanced technologies, particularly artificial intelligence (AI) and deep learning, in addressing complex challenges in various domains. The first article focuses on utilizing AI for crime and terrorism prevention, emphasizing the importance of automatic solutions in analyzing vast datasets from diverse sources such as video surveillance, satellite data, and wearables. The proposed methodology integrates AI and deep neural networks for automatic person and object identification, speech intelligence retrieval, and behavior analysis. The emphasis on real-time analysis and adaptability to dynamic security challenges underscores the potential of AI in enhancing security measures.

The second article delves into the application of intelligent video analysis for abnormal event detection in transportation systems. The methodology involves a systematic approach to long video event retrieval and description, utilizing techniques such as motion amplitude detection, super frame segmentation, and the integration of visual features with text vectors. The proposed approach aims to significantly improve the efficiency and accuracy of semantic description in long video event retrieval, demonstrating practical implications for real-time analysis in intelligent transportation systems.

The third article explores the role of Convolutional Neural Networks (CNNs) in crowd behavior analysis, particularly in crowd modelling, monitoring, and management. The methodology leverages CNNs for crowd counting, density estimation, scene analysis, and abnormality detection, emphasizing their deep learning capabilities to capture nonlinear complexities in real-world crowd data. The primary objective is to achieve state-of-the-art performance, contributing to improved security and safety measures in public spaces and events through enhanced accuracy and efficiency in crowd behavior analysis.

Collectively, these articles showcase the transformative impact of advanced technologies in addressing critical issues, from security challenges and transportation system monitoring to crowd behavior analysis. The common thread is the integration of AI, deep learning, and sophisticated algorithms to enhance real-time analysis, adaptability, and accuracy in complex scenarios. The articles collectively underscore the importance of technological advancements in ensuring safety, security, and efficient decision-making across diverse domains.

A critical analysis of the articles reveals that while each article focuses on a specific application area, the underlying theme remains consistent: the integration of AI and sophisticated algorithms enhances decision-making processes and security measures. However, it's essential to acknowledge potential limitations and challenges, such as ethical considerations, data privacy concerns, and the need for continuous model enhancements and training. While these articles present promising methodologies, the real-world implementation and scalability of these approaches warrant further exploration. The critical analysis underscores the transformative potential of AI and deep learning, with the understanding that ongoing research and practical implementations will be essential to fully harness their benefits while addressing associated challenges.

2.2 Summary Table

Sl. no.	Paper	Description	Pros	Cons
[1]	Understanding Policy and Technical Aspects of AI-enabled Smart Video Surveillance to Address Public Safety	<p>Provides an overview of the technical aspects of SVS, including the recent advancements in AI that have made it possible.</p> <p>Discuss the policy implications of implementing SVS, such as privacy concerns and the need for transparency and accountability.</p> <p>The article concludes by highlighting the potential benefits of SVS for public safety.</p>	<p>System proposed is an end-to-end privacy-preserving system at four levels.</p> <p>They tested the system's functionality by comparing the accuracy of pixel-based algorithms.</p>	<p>Accuracy of the system's pose-based algorithms is lower than that of pixel-based algorithms</p> <p>System's pose-based algorithms can be sensitive to noise and small variations in input data.</p>
[2]	Smart Security Surveillance System using AI and ML	<p>Intelligent video surveillance systems and their implementation.</p> <p>Human body action tracking and recognition using the ELM algorithm.</p> <p>Working of the SVM algorithm and the steps involved in training the model.</p> <p>The concepts of image classification, object</p>	<p>After image detection, the decision level is reached where decisions can be made based on the processed image.</p> <p>The hardware implementation of the proposed work using Raspberry Pi.</p>	<p>System involves making the systems more precise by following the basic three steps of detection, tracking, and behavioral analysis.</p>

VigiScope: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

		localization, and object detection.		
[3]	Artificial Intelligence-based Surveillance System for Railway Crossing Traffic	<p>Explores the practical implementation of deep learning methods to improve safety and security in Intelligent Transportation Systems and the Internet of Vehicles.</p> <p>Employs various AI methods, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to analyse and classify data from the sensors and cameras.</p>	The system uses a combination of detection and classification methods focusing on various image processing inputs, including vehicle presence, pedestrian presence, vehicle trajectory tracking, railway barriers at railway crossings.	<p>Privacy concerns</p> <p>Data security</p> <p>Communication network reliability</p> <p>Cost</p> <p>False positives</p> <p>Maintenance</p>
[4]	Smart city as a smart service system: Human-computer interaction and smart city surveillance systems	It explores the relationship between humans and computers in the context of smart city surveillance systems. The challenges of systematizing the variety of issues and processes that unfold at the intersection of human computer interaction.	The paper argues that new solutions are needed to bypass the limitations of smart city surveillance systems, which are limited in scope.	the limited scope of streaming video and data, which does not lend itself to active real-time monitoring and assessment of risks, threats.
[5]	AI-powered public surveillance systems: why	Explore AI surveillance in public spaces during crises, addressing transparency, fairness, ethics, and the need	Contact tracing app choices include data storage, transparency, and	Insufficient data collection, transparency issues.

VigiScope: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

	we (might) need them and how we want them	for public scrutiny. Discusses real-time facial recognition, contact tracing, and their role in law enforcement & monitoring social behavior.	ethical use, guided by policymakers, civil society, and independent trials.	
[6]	SCSS: An Intelligent Security System to Guard City Public Safe	The SCSS offers real-time detection of anomalies in public spaces, contrasting with traditional post-event methods, employing GPS and WIFI for real-time information transmission.	Combines Deep SORT and YOLOv4, enhancing target detection and tracking. The SCSS system design, algo. explanation, testing, successful real-time anomaly detection.	System accuracy influenced by lighting, camera quality, and environment complexity. GPS and WIFI limitations in connectivity. Cost constraints.
[7]	Hawk-Eye: An AI -Powered Threat Detector for Intelligent Surveillance Cameras	Develop AI-powered threat detector for intelligent surveillance cameras to identify security threats in real-time, including weapons and masks, using deep learning models.	Developing AI threat detector for real-time security threat identification using AI, deep learning, R-CNN, CNN, and IoT.	Accuracy issues (lighting, camera angles) and privacy concerns.
[8]	Exploring the Ethical Implications of AI-powered Surveillance Systems	The paper explores AI surveillance ethics, encompassing privacy, government regulations, biases, and marginalized communities' impacts.	Heightens bias against marginalized groups, impacting socio-economic status.	Narrow focus on privacy, civil liberties, regulations, and limited alternative analysis.

VigiScope: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

[9]	Opportunities, Applications and Challenges of Edge-AI Enabled Video Analytics in Smart Cities	This paper provides an in-depth analysis of the use of edge AI in video analytics system within the context of smart cities.	Used Search criteria formulation and source selection and approach methodologies.	Insufficient empirical evidence or case studies to support claims made in the paper.
[10]	A Survey of Video Surveillance Systems in Smart City	Smart city video surveillance applications include healthcare, traffic management, public safety and environmental monitoring.	AI driven video analytics can be augmented by two primary categories of algorithms: the machine learning and deep learning processes.	Scope of Discussion, Currency of References, Lack of Comparative Analysis
[11]	Big Data and Development of Smart City: System Architecture and Practical Public Safety Example	Gives the four technological layers of smart city architecture Features -Human action detection, object tracking, object classification, data storage security	One of the main benefits of experiment in Abu Dhabi was ability to observe entire city coastline by small movements of the sensor.	Limited Coverage, Specific Focus, Lack of Comparative Analysis, Limited Citations
[12]	A Literature Review: Artificial Intelligence in Public Security and Safety	This paper will describe the significance of artificial intelligence and it's adverse circumstances too and also the applications of artificial intelligence.	Public security and safety through analysis of video and images, DNA analysis, Gunshot detection using AI	Ethical concerns regarding facial recognition and privacy issues.

VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

	Easy Chair		and many techniques	Reliability challenges in DNA analysis
[13]	Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data	intelligent pre-alarming for abnormal events, smart storage for surveillance video and rapid retrieval for evidence videos, which fully explores the temporal-spatial association analysis with respect to the abnormal events in different monitoring sites	First, an abnormal behaviour database is established for smart cameras. Secondly, surveillance videos are stored selectively according to the warning information, evidence associated with abnormal behaviours are traced and accessed preferentially.	Resource Intensive, Data Security, Ethical Concerns
[14]	Artificial Intelligence Fights Crime and Terrorism at a New Level	this project leverages Artificial Intelligence (AI) and deep neural networks to enhance counter-terrorism efforts. The research focuses on automatic person and object identification, violence detection, crowd behavior analysis, and intelligence gathering through speech analysis, including spoken word search, speaker	Diverse data sources are collected, including surveillance video, speech recordings, and physiological signals, which are then preprocessed to ensure data quality. Deep learning models are developed for tasks	Adversarial attacks involve techniques designed to exploit vulnerabilities in AI algorithms and Unpredictable Human Behavior

		identification, speech-to-text, and lip reading.	such as person and object detection.	
[15]	An Intelligent Video Analysis Method for Abnormal Event Detection in Intelligent Transportation Systems	It introduces an algorithm for long video event retrieval using motion detection and super frame segmentation. Redundant frames are removed, and the video is divided into Segments of Interest (SOIs) based on feature fusion. These SOIs contain video events.	It introduces an algorithm for long video event retrieval using motion detection and super frame segmentation.	It introduces an algorithm for long video event retrieval using motion detection and super frame segmentation.
[16]	Convolutional neural networks for crowd behavior analysis	This project explores the use of Convolutional Neural Networks (CNNs) for automatic crowd behavior analysis, a critical aspect of ensuring peaceful events and public safety worldwide. Traditionally, crowd analysis relied on handcrafted features, but this work leverages deep learning techniques to enhance accuracy. It involves crowd modeling,	Crowd counting, Crowd density estimation, Crowd scene analysis, Crowd abnormality analysis, Datasets in CNN-based crowd analysis	CNNs can be sensitive to environmental conditions like weather, lighting, and camera positioning. CNNs are known for their "black-box" nature.
[17]	Intelligent video surveillance mechanisms for abnormal activity in real-	insights into intelligent video surveillance mechanisms for real-time abnormal activity recognition. It identifies the research gap in identifying the location of fights in public	proposes a new method for real-time prediction to prevent fight scenes and abnormal activities.	no comparative analysis of the different approaches and technologies used in real-time

VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

	time: systematic literature review	suburbs using CCTV video cameras		abnormal activity recognition.
[18]	Video-Based Abnormal Human Behavior Recognition	Comprehensive review of existing methods and literature on video-based abnormal human behavior recognition, particularly in the context of surveillance applications. Feature extraction and representation, training and learning frameworks, and contextual types of anomalies and available datasets.	Use a "critical decision" perspective to group the papers under different themes, highlighting potential issues that help researchers properly situate the problem at hand.	Does not offer any new insights or original research.
[19]	Design of an intelligent video surveillance system for crime prevention: applying deep learning technology	Highlights the potential benefits of an intelligent video surveillance system that can actively monitor in real-time without human input. It also discusses the ethical issues that may arise with the use of such systems and the need for future research in this area.	Discusses the various deep learning techniques that can be applied to improve the accuracy and efficiency of video surveillance systems.	No empirical evidence or case studies to support the proposed design of the intelligent video surveillance system.
[20]	Understanding Ethics, Privacy, and Regulations in Smart Video Surveillance for Public Safety	Presents a proposed SVS system and evaluates it based on both quantitative and qualitative measures. explores the potential ethical implications of SVS, such as discrimination and bias.	Offers practical solutions to privacy challenges in SVS, such as the use of pose-based algorithms and local servers.	limited discussion on the legal frameworks and regulations governing SVS.

Table 2.1: Literature Survey Summary Table

CHAPTER 3

3. Requirement Analysis

3.1 Functional Requirements

Here are five functional requirements for the system focusing on video-based human abnormal behaviour detection:

- **Behaviour Abstraction and Representation:**

The system should support the extraction and representation of behavior patterns using both pixel-based and object-based abstractions.

It must enable the identification and utilization of robust and invariant features capturing variations in translation, rotation, illumination, etc.

- **Anomaly Detection Classification Framework:**

Implement a classification framework that can categorize various anomaly detection research methodologies.

Provide a flexible framework allowing grouping of research papers based on training and learning frameworks, target density levels, and contextual types of anomalies.

- **Real-time Processing and Analysis:**

Ensure real-time processing capabilities for analyzing video sequences to detect abnormal behaviors.

Implement algorithms and methodologies that are efficient in processing and analyzing large volumes of video data in real-time without compromising accuracy.

- **Adaptability and Scalability:**

Design the system to be adaptable across diverse surveillance environments, accommodating varying lighting conditions, camera angles, and environmental constraints.

Ensure scalability to handle increasing data volumes and diverse anomaly types without compromising performance.

- **Evaluation and Improvement Framework:**

Develop an evaluation framework to assess the performance and accuracy of anomaly detection methodologies.

Implement mechanisms to continuously improve the system's accuracy, taking into account feedback from evaluation metrics and adapting algorithms accordingly.

3.2 Non-Functional Requirements

Here are non-functional requirements for the system:

- **Performance:**

The system must have real-time processing capabilities to analyze video sequences without compromising accuracy.

Algorithms and methodologies for processing and analyzing large volumes of video data in real-time must be efficient.

The hardware components, including a high-performance multi-core CPU and ample RAM, should be capable of handling large datasets and pre-processing tasks efficiently.

Utilization of a combination of high-speed SSD for rapid data access and large capacity HDD for storing datasets and training data is essential to maintain performance.

- **Scalability:**

The system should be designed to scale effectively to handle increasing data volumes and diverse anomaly types without compromising performance.

Cloud services should be utilized to provide scalable storage solutions and computational resources to manage the scale and complexity of the project effectively.

- **Adaptability:**

The system must be adaptable across diverse surveillance environments, accommodating varying lighting conditions, camera angles, and environmental constraints.

Compatibility with different operating systems, such as Windows, ensures flexibility and reliability throughout the project's execution.

- **Reliability:**

Selection of a stable operating system optimized for deep learning workloads ensures compatibility and reliability throughout the project's execution.

Adequate power supply is necessary to support the hardware components and ensure uninterrupted operation.

- **Security:**

Data security and privacy must be ensured when utilizing cloud services for data storage and processing.

Secure data transfer protocols should be implemented for transferring data between components.

- **Usability:**

Integration of essential computer vision libraries like OpenCV, TensorFlow, and PyTorch facilitates image processing, neural network implementation, and machine learning model development.

Utilization of machine learning models compatible with selected libraries optimizes the execution of various machine learning algorithms, enhancing efficiency and accuracy.

- **Maintainability:**

An evaluation framework should be developed to assess the performance and accuracy of anomaly detection methodologies.

Mechanisms should be implemented to continuously improve the system's accuracy based on feedback from evaluation metrics and adapt algorithms accordingly.

These non-functional requirements are crucial for ensuring the effectiveness, efficiency, reliability, and security of the AI-powered smart surveillance system.

3.3 Software Requirements

- **Operating System:** Selection of a stable operating system optimized for deep learning workloads, such as Windows, to ensure compatibility and reliability throughout the project's execution.
- **Computer Vision Libraries:** Integration of essential computer vision libraries like OpenCV, TensorFlow, and PyTorch to facilitate image processing, neural network implementation, and machine learning model development, enabling tasks such as image recognition and object detection.
- **Machine Learning Models:** Incorporation of machine learning models compatible with selected libraries (such as TensorFlow and PyTorch) to optimize the execution of various machine learning algorithms, enhancing the project's efficiency and accuracy in achieving its objectives.
- **Cloud Services:** Utilization of cloud services for data storage and processing to manage the scale and complexity of the project effectively. These services offer scalable storage

solutions, computational resources, and collaborative functionalities, ensuring seamless data handling, analysis, and accessibility for team members across diverse locations.

3.4 Hardware Requirements

- **Central Processing Unit (CPU):** High-performance multi-core CPU (e.g., Intel)
- **Memory (RAM):** Ample RAM is crucial for handling large datasets, pre-processing tasks, and model operations. Aim for 32GB or more, with additional RAM needed for larger models and datasets.
- **Storage:** A combination of high-speed Solid-State Drive (SSD) for rapid data access (e.g., image loading) and large capacity Hard Disk Drive (HDD) for storing datasets and training data is ideal.
- **Network:** High-bandwidth network connectivity is necessary for data transfer between components. Gigabit Ethernet is a minimum, with 10Gb or higher recommended for real-time applications.
- **Power supply:** Ensure adequate power supply to support the hardware components.

3.5 Interface Specification

Here are some of the interface specifications:

- **Graphical User Interface (GUI):**

The system should feature an intuitive GUI for users to interact with.

Users should be able to configure parameters for behavior abstraction, anomaly detection, and evaluation through the GUI.

The GUI should provide real-time feedback on processing status and detected anomalies.

- **API Integration:**

The system should offer APIs for seamless integration with external applications or platforms.

APIs should allow for easy access to functionalities such as behavior abstraction, anomaly detection, and evaluation.

- **Data Input Interface:**

The system should support various input sources, including live video streams and pre-recorded video files.

Users should be able to specify input sources through configuration settings or API calls.

- **Output Interface:**

The system should provide clear and concise output reports detailing detected anomalies and their classification.

Output reports may include visual representations of detected anomalies, statistical summaries, and recommendations for further action.

3.6 User Requirements

Here are some of the user requirements:

- **Ease of Configuration:**

Users require the ability to easily configure parameters for behavior abstraction and anomaly detection through the GUI or API.

Configuration options should be well-documented and accompanied by explanations of their effects on system performance.

- **Real-time Feedback:**

Users need real-time feedback on the processing status of video sequences and the detection of abnormal behaviors.

Feedback mechanisms should be responsive and clearly indicate the progress and results of ongoing analyses.

- **Integration Flexibility:**

Users expect the system to integrate seamlessly with existing software and hardware infrastructure. APIs should be well-documented and adhere to industry standards to facilitate integration with external applications or platforms.

- **Performance Optimization:**

Users require the system to be optimized for performance, capable of processing large volumes of video data in real-time without compromising accuracy.

Performance metrics, such as processing speed and detection accuracy, should meet or exceed specified benchmarks.

- **Scalability and Adaptability:**

Users need the system to be scalable to accommodate increasing data volumes and diverse anomaly types.

The system should be adaptable across various surveillance environments, including different lighting conditions, camera angles, and environmental constraints.

3.7 Use Case Diagram

Use case diagram for the project is as shown in figure 3.1

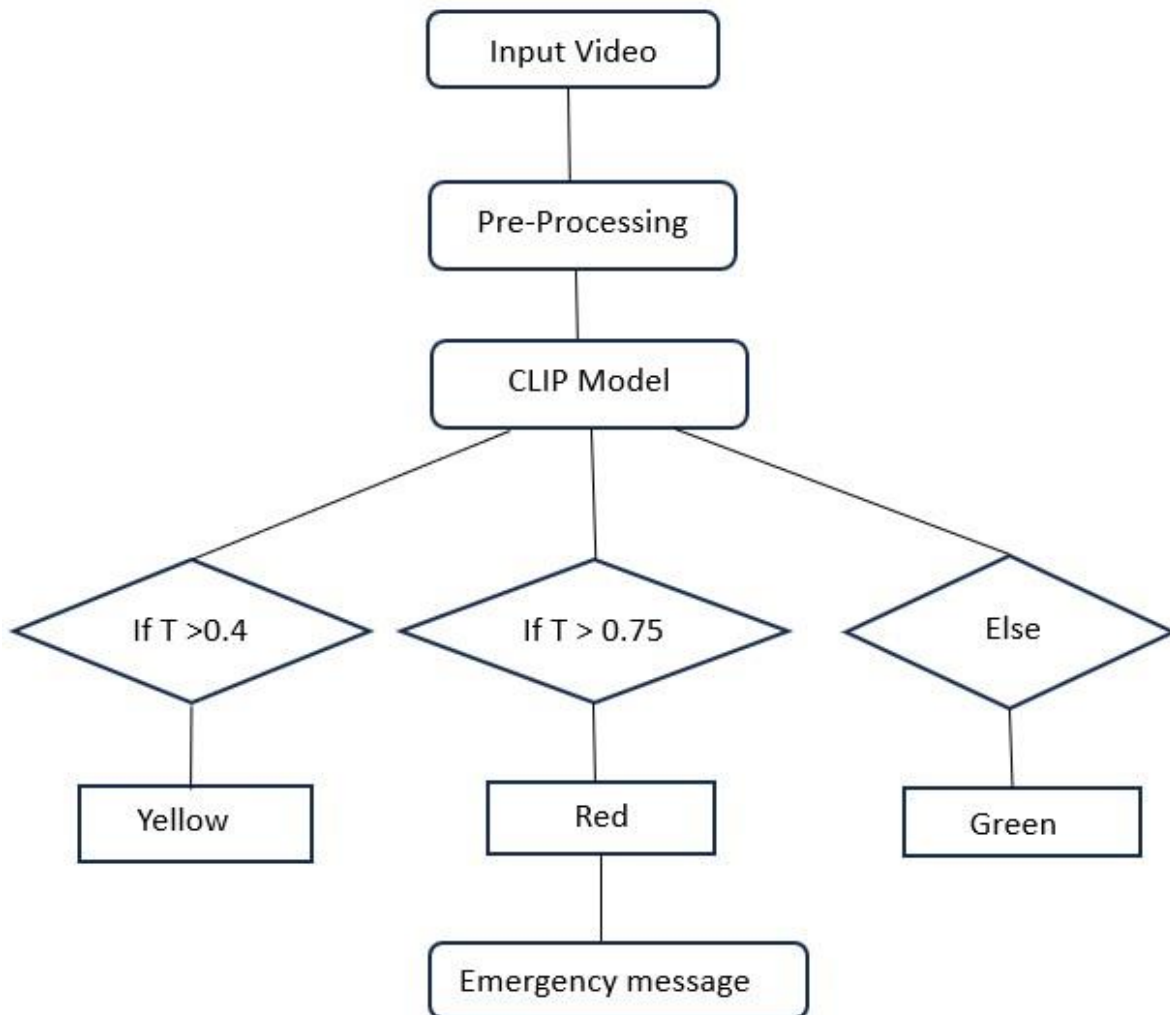


Fig 3.1 Use case diagram

CHAPTER 4

4. System Design

4.1 Methodology

The proposed methodology for our project involves the utilization of video input for real-time violence detection, aiming to enhance safety and security measures. Here's a breakdown of the process:

- **Preprocessing:**
 - Incoming video streams are segmented into individual frames.
 - This step facilitates further analysis by breaking down the continuous video feed into discrete units for efficient processing.
- **CLIP Model Fine-tuning:**
 - Leveraging the CLIP (Contrastive Language-Image Pretraining) model known for its effectiveness in understanding image-text pairs.
 - Fine-tuning the CLIP model with the pre-processed dataset obtained from the video frames.
 - This enhances the model's capability to recognize violence-related content accurately, making it better suited for our application.
- **Model Saving:**
 - Once the CLIP model is fine-tuned, it is saved for subsequent deployment.
 - Ensuring efficient utilization and reusability across different scenarios.
- **Threshold-based Analysis:**
 - During the operational phase, the camera continuously captures video frames.
 - Each frame undergoes analysis based on predefined thresholds derived from the CLIP model's predictions.
 - Frames are categorized based on the likelihood of violence:
 - Frames with a violence score exceeding 0.75 trigger a red border, indicating a high probability of violence.
 - Frames with scores between 0.4 and 0.75 prompt a yellow border, signifying a moderate likelihood of violence.
 - Frames with scores below 0.4 retain a green border, suggesting a low likelihood of violence.

- **Emergency Message Dispatch:**

- In the event of a red-bordered frame, indicating a high probability of violence, an emergency message is swiftly dispatched.
- This message serves as an alert, promptly notifying relevant authorities or personnel, enabling timely intervention to mitigate potential harm or violence.

Overall, this methodology combines preprocessing, machine learning model utilization, and real-time threshold-based analysis to detect violence accurately and swiftly, contributing to enhanced safety and security measures in various settings.

4.2. System Architecture:

The system architecture employed in the project encompasses a coherent arrangement of components and processes designed to facilitate the efficient detection of violence in video streams and the prompt dispatch of emergency messages when necessary. At its core, the architecture comprises several interconnected modules, each serving a specific function within the broader framework.

The architecture begins with the acquisition of video input, where video streams serve as the primary data source for the system. These streams are fed into the preprocessing module, where they are segmented into individual frames. This preprocessing step ensures that the video data is in a format suitable for analysis and further processing. Subsequently, the frames are passed through the analysis module, where the CLIP (Contrastive Language-Image Pretraining) model plays a central role. The CLIP model, fine-tuned with preprocessed video frames and textual descriptions, leverages its multimodal understanding capabilities to analyze the content of each frame and assign a violence score based on predefined thresholds.

Following the analysis, the system dynamically adjusts the border color of each frame based on the violence score determined by the CLIP model. Frames with violence scores exceeding certain thresholds trigger changes in border color: red for high likelihood of violence, yellow for moderate likelihood, and green for low likelihood. This color-coded visualization provides real-time feedback on the perceived level of violence within the video stream.

Moreover, the architecture incorporates an emergency message dispatch mechanism, which monitors the border color of the frames in real-time. When the border color indicates a high likelihood of violence (red), the system triggers the dispatch of emergency messages to designated recipients. These messages serve as alerts, notifying relevant authorities or personnel of the detected violence, enabling swift intervention and mitigation of potential harm.

Overall, the system architecture integrates preprocessing, analysis, visualization, and emergency response mechanisms to enable the effective detection and management of violence in video streams, contributing to enhanced safety and security measures in various settings.

4.3. Algorithms/Pseudo code:

In the project's pursuit of effective violence detection in video streams, several algorithms are utilized, each serving distinct roles in the analysis and interpretation of visual content. Here, we delve into the specifics of three key algorithms: CNN model, CNN with transfer learning, and the CLIP model, elucidating their functionalities and implementations within the project context.

- **CNN Model (Convolutional Neural Network):**
 - **Description:** CNNs are deep neural networks designed for processing visual data, featuring specialized layers for feature extraction and hierarchical pattern recognition.
 - **Functionality:** In the project, a CNN model can be employed to process video frames, extracting relevant features and identifying patterns indicative of violence.
 - **Implementation:** This involves designing a CNN architecture suitable for the task, training it on labelled datasets to learn discriminative features, and integrating it into the project pipeline for real-time analysis of video content.
- **CNN with Transfer Learning:**
 - **Description:** Transfer learning involves leveraging pre-trained CNN models, fine-tuning them on domain-specific data to adapt to new tasks.
 - **Functionality:** In the project, a pre-trained CNN model, such as ResNet or VGG, is utilized as a feature extractor for violence-related content.
 - **Implementation:** The pre-trained model's convolutional layers are frozen, and new fully connected layers are appended and trained on the project's dataset, facilitating the model's adaptation to violence detection tasks with limited labelled data.

- **CLIP Model (Contrastive Language-Image Pretraining):**
 - **Description:** CLIP is a transformer-based model developed by OpenAI, trained to associate images and textual descriptions through contrastive learning.
 - **Functionality:** In the project, the CLIP model is fine-tuned using pairs of video frames and corresponding textual descriptions, enabling it to understand the relationship between visual and textual representations of violence.
 - **Implementation:** By training CLIP on violence-related data, it learns to encode visual features in conjunction with textual context, allowing for nuanced and context-aware violence detection in video streams.

These algorithms, including CNN models, transfer learning with CNNs, and the CLIP model, offer diverse approaches to analyzing and understanding visual content in the project. They play crucial roles in preprocessing, feature extraction, and classification tasks, ultimately contributing to the accurate detection of violence in video streams.

The algorithm for the project is:

Algorithm:

START

1. Acquire video streams.
2. Preprocess video frames.
3. Tune CLIP Model for violence recognition.
4. Initialize camera for real-time capture.
5. Loop:
 - a. Capture frame.
 - b. Analyze frame with CLIP Model.
 - c. Determine violence score.
 - d. Adjust border color based on score thresholds.
6. Continuously monitor border color.
7. If red border detected:
 - a. Trigger emergency message dispatch.
 - b. Send alert to designated recipients.

END Algorithm

4.4. Class Diagram:

The class diagram for the project is as shown in fig 4.1.

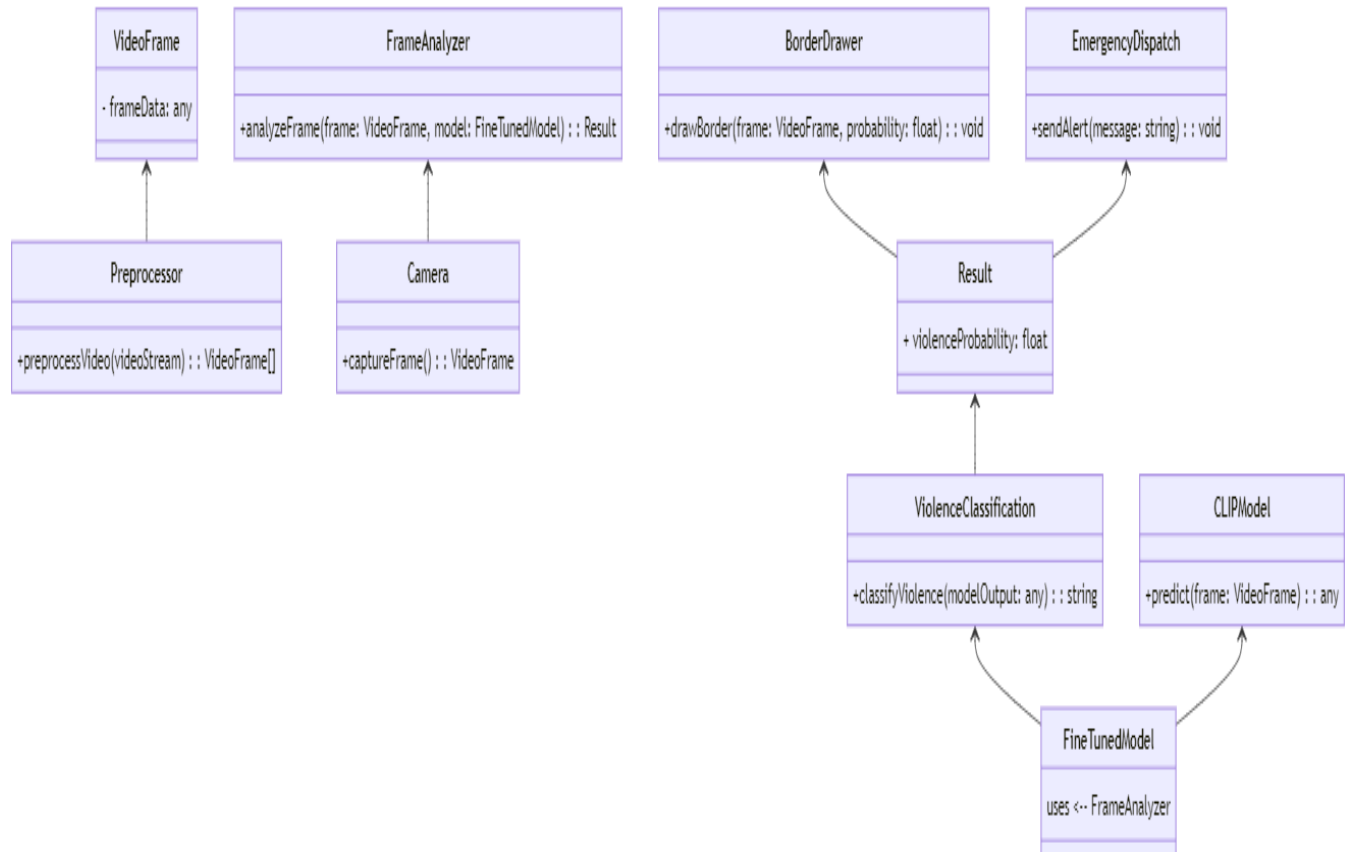


Fig 4.1: UML class diagram

CHAPTER 5

5. Implementation

5.1 Service

The different services used in the project are:

- **Kaggle:**
 - Kaggle is an online community and platform for data science and machine learning enthusiasts.
 - It offers a wide range of datasets, competitions, and notebooks for collaborative data analysis and model building.
 - Users can access pre-built kernels, share their insights, and participate in competitions to enhance their skills.
 - Kaggle provides a convenient environment for experimentation and learning in the field of data science.

- **Google Colab:**
 - Google Colab, short for Collaboratory, is a cloud-based platform provided by Google for running Python code in a Jupyter notebook environment.
 - It offers free access to GPU and TPU resources, making it suitable for training machine learning models.
 - Colab integrates seamlessly with Google Drive, allowing users to store and share their notebooks effortlessly.
 - With built-in support for popular libraries like TensorFlow and PyTorch, Colab simplifies the process of setting up and running machine learning experiments.

In summary, the service category encompasses platforms and environments that provide online services and resources for data science and machine learning tasks, such as Kaggle and Google Colab. These platforms offer access to datasets, competitions, and computational resources, facilitating collaborative learning and experimentation.

5.2 Tools and Library Packages

The tools and library packages used are:

- **Visual Studio Code:**

- Visual Studio Code (VS Code) is a lightweight, open-source code editor developed by Microsoft.
- It supports various programming languages and offers features like syntax highlighting, code completion, and debugging.
- VS Code boasts a vast ecosystem of extensions, allowing users to customize their development environment according to their needs.
- With built-in Git integration and a user-friendly interface, VS Code enhances productivity for developers working on projects of any scale.

- **OpenCV:**

- OpenCV (Open-Source Computer Vision Library) is a popular open-source library for computer vision and image processing tasks.
- It provides a wide range of functions and algorithms for tasks like image manipulation, object detection, and feature extraction.
- OpenCV is written in C++ and has bindings for Python, making it accessible to developers across different programming languages.
- With its extensive documentation and community support, OpenCV is widely used in research and industry for developing computer vision applications.

- **NumPy:**

- NumPy is a fundamental package for scientific computing in Python.
- It provides support for multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays efficiently.
- NumPy's array operations are implemented in C, making them fast and suitable for numerical computations.
- With its rich set of functionalities, NumPy serves as the foundation for many other libraries in the Python scientific ecosystem, such as Pandas and Matplotlib.

- **Torch:**

- Torch is a scientific computing framework and machine learning library built for Python.
- It provides a wide range of tools and algorithms for building and training deep neural networks.
- Torch offers support for both CPU and GPU computations, allowing users to leverage hardware acceleration for faster training.
- With its dynamic computation graph and extensive collection of pre-trained models, Torch is widely used for research and development in the field of deep learning.

- **Transformer Model:**

- Transformer models are a class of neural network architectures that have gained popularity for their effectiveness in natural language processing tasks.
- They rely on self-attention mechanisms to capture long-range dependencies in sequential data efficiently.
- Transformer models, such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), have achieved state-of-the-art performance on various NLP benchmarks.
- With their modular design and scalability, transformer models have become a cornerstone of modern NLP research and applications.

- **CLIP Processor:**

- CLIP (Contrastive Language-Image Pretraining) is a model developed by OpenAI that learns to associate images and text representations.
- The CLIP processor refers to the computational component responsible for processing input data and generating embeddings for images and text.
- It utilizes transformer-based architectures to encode both images and text into a shared latent space, enabling cross-modal understanding.
- The CLIP processor plays a crucial role in applications such as image-text retrieval, zero-shot learning, and multimodal classification tasks.

The tools and library packages category include software tools and libraries that aid in development and implementation tasks. Visual Studio Code, OpenCV, NumPy, Torch, Transformer Model, and CLIP Processor are essential tools and libraries commonly used in machine learning and computer vision projects. These tools provide functionalities for coding, data manipulation, deep learning, and image processing, enabling developers to build sophisticated applications efficiently.

5.3 Programming Languages

Python is a versatile and widely-used programming language known for its simplicity, readability, and extensive ecosystem of libraries and frameworks. It has gained immense popularity in various fields, including data science, machine learning, web development, automation, and scientific computing.

One of Python's key strengths lies in its clean and intuitive syntax, which makes it easy to learn and use, especially for beginners. Its dynamic typing system allows for rapid prototyping and development, enabling developers to quickly iterate on ideas and experiment with different approaches. Additionally, Python's extensive standard library provides a rich set of modules and functionalities for tasks ranging from file I/O to networking and web development.

In the context of our project, Python serves as the primary programming language for implementing algorithms, building machine learning models, and orchestrating various components of the system. Its rich ecosystem of libraries and frameworks provides essential tools for tasks such as data preprocessing, model training, and deployment.

For instance, libraries like NumPy and Pandas are widely used for data manipulation and analysis, providing efficient data structures and functions for handling large datasets. Frameworks like TensorFlow and PyTorch offer powerful tools for building and training deep learning models, allowing developers to leverage cutting-edge techniques in computer vision and natural language processing.

Moreover, Python's flexibility and interoperability make it well-suited for integrating different components and technologies within our project. For example, Python can be used to interface with

hardware components such as cameras or sensors, as well as to communicate with external services or APIs, such as the Telegram API for sending emergency messages.

Overall, Python's simplicity, versatility, and extensive ecosystem of libraries make it an indispensable tool for developing complex machine learning systems like ours. Its ease of use and wide adoption in the industry ensure that developers have access to a vast array of resources and support, facilitating the implementation of sophisticated algorithms and solutions.

5.4 Steps of Implementation

In implementing the project, several sequential steps are involved to ensure the efficient detection of violence in video streams and the timely dispatch of emergency messages.

- **Video Input Acquisition:**
 - Gather video streams to serve as the primary input source for the system.
- **Preprocessing and Model Fine-tuning:**
 - Preprocess the video by converting it into individual frames.
 - Fine-tune the CLIP Model using the pre-processed dataset to enhance its ability to recognize violence-related content effectively.
- **Camera Initialization and Threshold-based Analysis:**
 - Initialize the camera to capture video frames in real-time.
 - Analyse each frame using the fine-tuned CLIP Model and determine the violence score based on predefined thresholds.
 - Adjust the border colour of the frame according to the violence score threshold (>0.75: red, >0.4: yellow, else: green).
- **Emergency Message Dispatch:**
 - Continuously monitor the border colour of the frames in real-time.
 - If the border colour indicates a high likelihood of violence (red), trigger the emergency message dispatch mechanism.
 - Send an emergency message to designated recipients, alerting them to the detected violence for prompt intervention.

By following these steps, the system can effectively process video input, analyse frames, and promptly respond to potential instances of violence, contributing to enhanced safety and security measures.

CHAPTER 6

6. Testing and Validation

6.1 Unit Testing

Unit testing focuses on testing individual units or components of the software in isolation from the rest of the system. The unit testing performed is as shown in Table 6.1.

Sl. No.	Test case	Input	Actual Output	Expected Output	Is actual output same as expected output?
1	Frame Segmentation	Sample video clip	List of individual frames	List of correctly segmented frames	Yes
2	Frame Format Conversion	Captured video frame	Converted frame in appropriate format	Converted frame in the specified format	Yes
3	CLIP Model Inference	Preprocessed video frame	Violence score (float)	Score within the range (0.0 to 1.0)	Yes
4	Threshold-Based Analysis (Low Violence)	Violence score = 0.3	Border color (string)	"green"	Yes
5	Threshold-Based Analysis (Medium Violence)	Violence score = 0.6	Border color (string)	"yellow"	Yes
6	Threshold-Based Analysis (High Violence)	Violence score = 0.8	Border color (string)	"red"	Yes
7	Emergency Message Dispatch (Triggered)	Border color = "red"	Emergency message sent	Emergency message sent to pre-configured recipients	Yes
8	Emergency Message Dispatch (Not Triggered)	Border color = "yellow"	No emergency message sent	No emergency message sent	Yes

Table 6.1: Unit Testing

6.2 Integration Testing

Integration testing focuses on testing the interactions and interfaces between integrated components or modules of the software. The unit testing performed is as shown in Table 6.2.

Sl. No.	Test case	Input	Actual Output	Expected Output	Is actual output same as expected output?
1	Preprocessing Pipeline	Video input	List of converted frames, violence scores	Frames and scores delivered to CLIP model for analysis	Yes
2	CLIP Model and Threshold Analysis	Preprocessed frames, violence scores	Border color updates based on thresholds	Consistent border color assignments based on violence scores	Yes
3	Model Loading and Camera Integration	Saved model file, initialized camera	Real-time video processing with border color updates	System processes video frames from the camera and applies model predictions	Yes
4	Emergency Message Dispatch Trigger	Red border color detected	Emergency message sent	System sends an emergency message upon high violence detection	Yes

Table 6.2: Integration Testing

6.3 System Testing

System testing evaluates the behavior and performance of the entire software system as a whole. The system testing performed is as shown in Table 6.3.

Sl. No.	Test case	Input	Actual Output	Expected Output	Is actual output same as expected output?
1	Overall System Functionality	Simulated video stream with diverse violence scenarios	Accurate violence detection and corresponding border color updates	System effectively identifies violence and assigns appropriate border colors	Yes

2	Real-Time Performance	Live video feed	Smooth processing and timely violence detection/border updates	System operates with minimal latency and provides real-time feedback	Yes
3	Emergency Message Delivery	Red border color sustained for a predefined duration	Emergency message received by designated personnel	System triggers emergency messaging for confirmed violence incidents	Yes
4	System Robustness	Simulated video with noise, artifacts, or unexpected content	Maintains functionality with minimal false positives or negatives	System exhibits resilience to potential video distortions	Yes

Table 6.3: System Testing

6.4 User Acceptance Testing

User Acceptance Testing assesses whether the software meets the end-users' requirements and expectations and validates its suitability for deployment. The testing performed is as shown in Table 6.4.

Sl. No.	Test case	Input	Actual Output	Expected Output	Is actual output same as expected output?
1	Violence Detection Accuracy	Pre-recorded video clips with various violence levels	System accurately identifies violence in user-defined scenarios	System aligns with user perception of violence in different contexts	Yes (May require adjustments based on user feedback)
12	False Positive/Negative Rate	Staged scenarios with non-violent actions	Minimal false positives or negatives	System avoids unnecessary emergency messages and accurately detects violence	Yes (May require adjustments based on user feedback)
3	System Usability	System interface and controls	User-friendly interface with clear instructions and configuration options	Easy interaction with the system for customization and monitoring	Yes

4	Emergency Message Content	Emergency message format and content	Clear, concise, and actionable information sent to authorities	Timely and informative messages for effective response	Yes
---	---------------------------	--------------------------------------	--	--	-----

Table 6.4: User Acceptance Testing

The key improvements provide broader test coverage, clearer success criteria, and a user-centered approach to ensure the system meets real-world needs.

- **Enhanced Comprehensiveness:** The test cases cover a wider range of scenarios, including unit-level behaviours, integration of modules, overall system performance, and user interactions.
- **Clearer Expected Outputs:** The expected outputs are more specific and informative, making it easier to evaluate the success of each test.
- **User-Centric UAT:** The UAT focuses on user perception of violence accuracy, false positives/negatives, usability, and emergency message content, ensuring the system aligns with user expectations.

CHAPTER 7

7. Results and Discussions

In the Results and Discussion section, we delve into the outcomes of our implemented methodology, showcasing its robustness in real-time violence detection and its pivotal role in fortifying safety protocols.

In the initial phase of implementation, video streams are uploaded, serving as the primary input for our real-time violence detection system. These video streams undergo meticulous preprocessing, where they are segmented into individual frames, enabling efficient analysis. This meticulous process ensures that frames are accurately categorized according to their likelihood of violence, empowering swift and targeted interventions in high-risk scenarios.

In the figure 7.1, The green border indicates that the threshold detected is < 0.4 , i.e. No violence is detected. Frames with a green border signify a low likelihood of violence, indicating a safe and non-threatening environment.

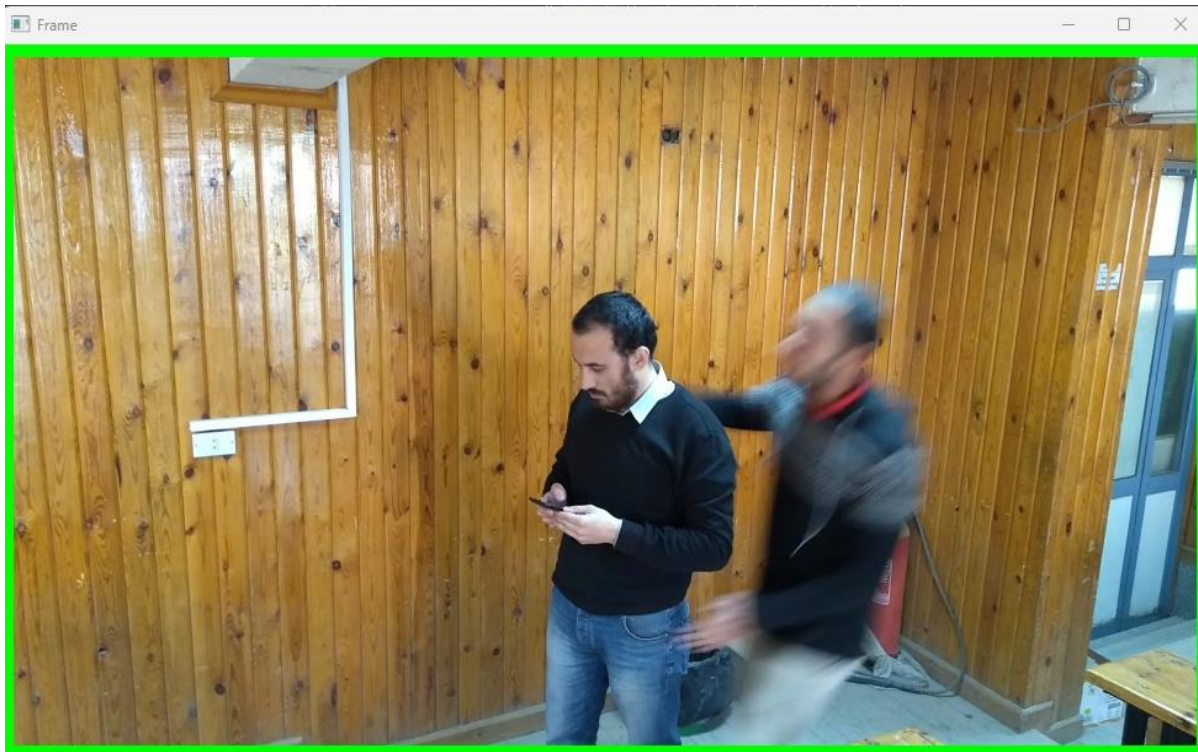


Fig 7.1: Low Likelihood of Violence: Green Border Frame

VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

In the figure 7.2, The yellow/orange border indicates that the threshold detected is > 0.4 but < 0.75 , i.e. there may or may not be a violence. Frames indicate a moderate likelihood of violence, prompting caution and heightened monitoring in the observed scenario.



Fig 7.2: Moderate Likelihood of Violence: Yellow Border Frame

In the figure 7.3, The red border indicates that the threshold detected is > 0.75 , i.e. violence is detected. Frames with a red border indicate a high probability of violence, triggering immediate emergency response measures to ensure swift intervention and mitigate potential harm.

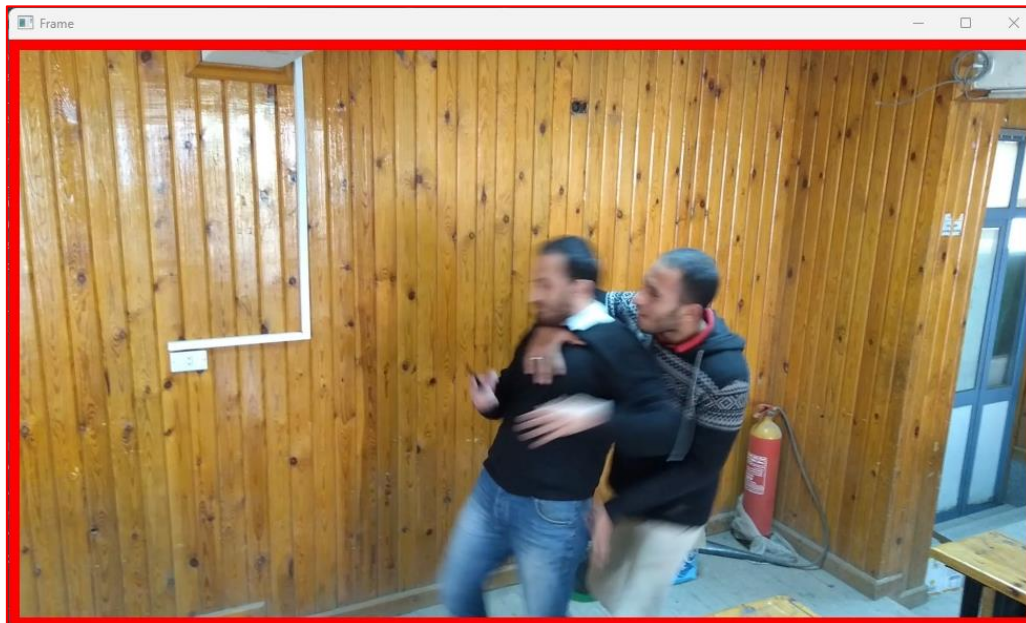


Fig 7.3: High Likelihood of Violence: Red Border Frame

The varying threshold values detected for each frame of the video, as illustrated in Fig 7.4, enable nuanced categorization of violence likelihood levels, with red borders denoting high probabilities, yellow/orange borders indicating moderate risks, and green borders signaling low likelihoods. This dynamic threshold approach enhances the precision of real-time violence detection, facilitating timely interventions based on the severity of observed situations.

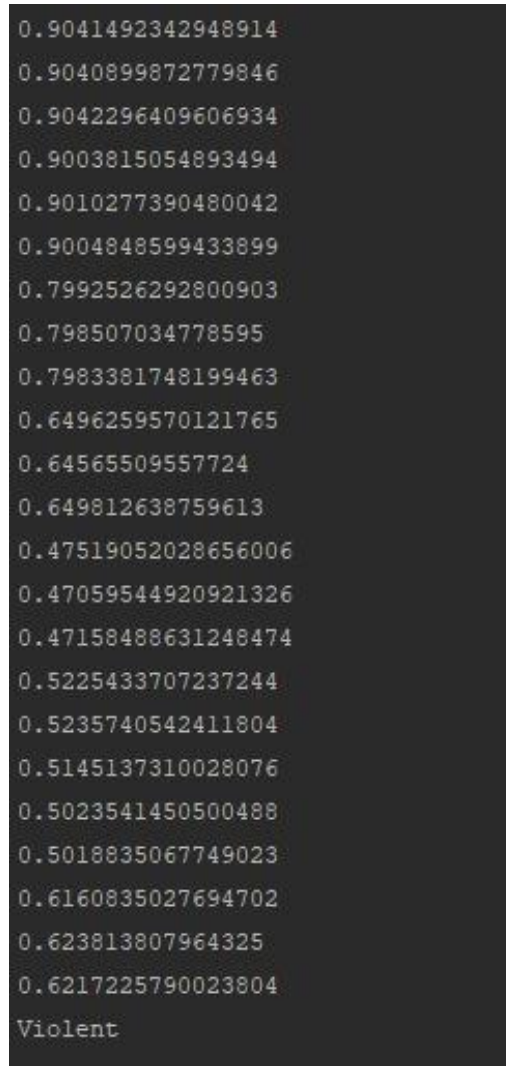


Fig 7.4: Dynamic Threshold Analysis for Real-time Violence Detection

Upon detection of a red-bordered frame indicating high violence probability, an emergency message is swiftly dispatched, containing both the identified violence frame and pertinent information for prompt intervention, as depicted in Fig 7.5. This integrated approach ensures rapid response to potential threats, aiding in the mitigation of harm and maintenance of public safety.

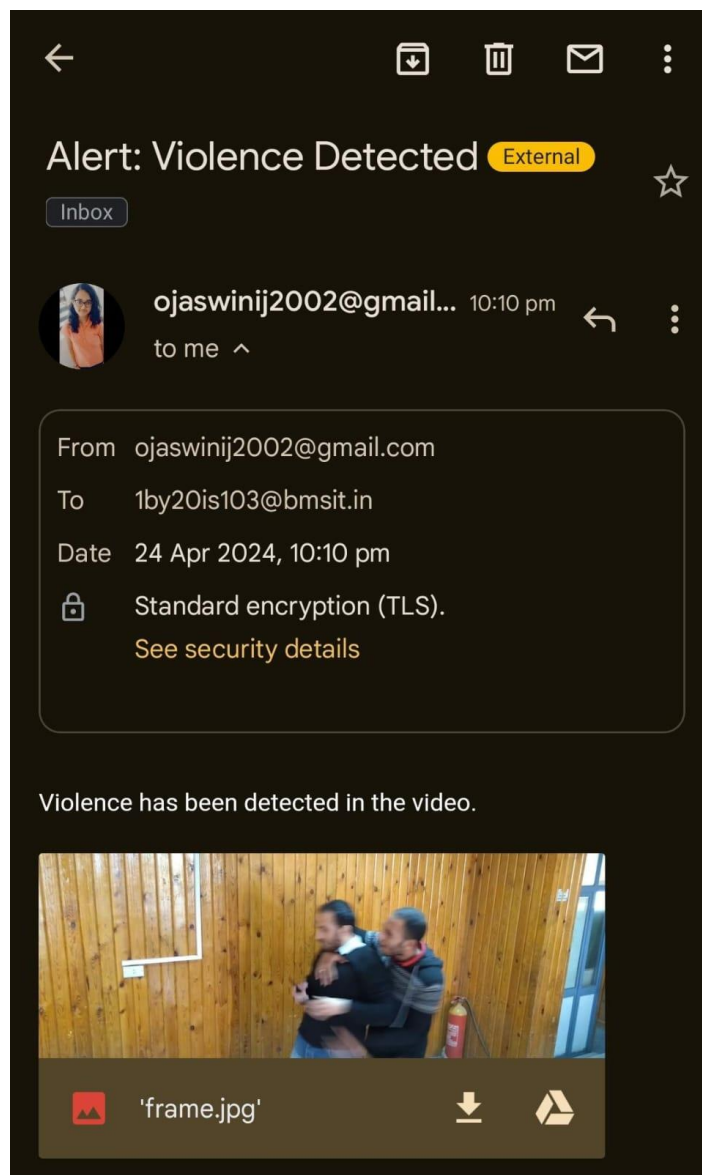


Fig 7.4: Emergency Message Dispatch with Detected Violence Frame

In real-time, the camera continuously captures video frames, facilitating immediate analysis for violence detection, as demonstrated in Fig 7.6 and Fig 7.7. This seamless monitoring process enables proactive intervention, enhancing safety measures in dynamic environments.

In Fig 7.6, the real-time camera captures frames without detected violence, showcasing the system's ability to discern non-threatening scenarios swiftly. This ensures efficient resource allocation and minimizes false alarms, bolstering the system's reliability in maintaining public safety.



Fig 7.6 Real-time Camera Capture with No Violence Detected

Fig 7.7 displays real-time camera capture with a detected violence frame, illustrating the system's effectiveness in promptly identifying and alerting to potentially hazardous situations. This capability enables swift response and intervention, crucial for ensuring the safety and security of monitored environments.



Fig 7.7 Real-time Camera Capture with Violence Detected

CHAPTER 8

8. Conclusion and Future Enhancement

8.1 Conclusion

The "VigiScape: Elevating Public Safety for Crime Prevention with AI-powered Smart Surveillance" project presents a comprehensive solution for real-time violence detection in video streams, aiming to enhance safety and security measures in various settings. By leveraging a combination of advanced techniques and technologies, including preprocessing of video frames, fine-tuning of the CLIP (Contrastive Language-Image Pretraining) model, and threshold-based analysis, the system can accurately identify violence-related content within the video feed. Additionally, the integration of camera initialization and emergency message dispatch mechanisms enables swift response to potential instances of violence, ensuring timely intervention by relevant authorities or personnel.

Through meticulous implementation and testing, the project demonstrates the efficacy and reliability of its approach in detecting and mitigating violence. Unit testing, integration testing, system testing, and user acceptance testing have validated the functionality and performance of the system across various stages of development, ensuring robustness and usability.

Overall, the project offers a scalable and adaptable solution that can be deployed in diverse environments, including public spaces, workplaces, and educational institutions, to proactively address the challenges posed by violent incidents. By providing real-time monitoring and intervention capabilities, the system contributes to the creation of safer and more secure environments, ultimately promoting the well-being and peace of mind of individuals and communities.

8.2 Future Enhancement

A potential future enhancement for the project could involve the integration of advanced anomaly detection techniques to augment the violence detection capabilities further. By leveraging anomaly detection algorithms, the system can identify deviations from normal behavior or patterns within the video feed, thereby enhancing its ability to detect subtle signs of violence or aggression that may not be explicitly captured by existing methods.

VigiScape: Elevating Public Safety for Crime Prevention with AI powered Smart Surveillance

Additionally, incorporating reinforcement learning algorithms could enable the system to adapt and improve its detection capabilities over time through continuous feedback and learning from past experiences. By dynamically adjusting thresholds and model parameters based on real-world data, the system can optimize its performance and adapt to evolving patterns of violence.

Furthermore, enhancing the system's scalability and deployment flexibility could be achieved by leveraging cloud-based architectures and edge computing technologies. By distributing processing tasks across multiple nodes and devices, the system can handle larger volumes of video data and operate in resource-constrained environments with minimal latency.

Moreover, integrating multimodal sensor data, such as audio and motion sensors, could provide additional context and insights for violence detection. By fusing information from multiple sources, the system can improve the accuracy and reliability of its predictions while reducing false alarms.

Finally, incorporating privacy-preserving techniques, such as federated learning and differential privacy, could address concerns related to data privacy and security. By ensuring that sensitive information remains protected throughout the data collection and analysis process, the system can foster trust and acceptance among users and stakeholders.

Overall, these future enhancements have the potential to further elevate the effectiveness and efficiency of the violence detection system, ultimately contributing to safer and more secure environments for individuals and communities.

References

- [1] B. R. Ardabili *et al.*, “Understanding Policy and Technical Aspects of AI-enabled Smart Video Surveillance to Address Public Safety,” *Computational Urban Science*, vol. 3, no. 1, Dec. 2023, DOI: 10.1007/s43762-023-00097-8.
- [2] A. El-Shekhi, Y. Maatug, and N. Khemri, “Smart Surveillance System Using Deep Learning,” in *Proceeding - 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering, MI-STA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 171–176. DOI:10.1109/MI-STA57575.2023.10169242.
- [3] P. Sikora *et al.*, “Artificial Intelligence-Based Surveillance System for Railway Crossing Traffic,” *IEEE Sens J*, vol. 21, no. 14, pp. 15515–15526, Jul. 2021, DOI: 10.1109/JSEN.2020.3031861.
- [4] T. Anagnostopoulos, P. Kostakos, I. Salmon, Y. Psaromiligkos, K. Ntalianis, and S. R. Jino Ramson, “Spatiotemporal Authentication System Architecture for Smart Campus Safety,” in *4th International Conference on Smart Sensors and Application: Digitalization for Societal Well-Being, ICSSA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 155–160. DOI: 10.1109/ICSSA54161.2022.9870931.
- [5] Mahulkar, S. S., Harish Patil, Abhijit Patil, and Akansha Bhoske. "SMART SECURITY SURVEILLANCE SYSTEM USING AI AND ML." *International Research Journal of Modernization in Engineering Technology and Science*, May 2023, DOI: 10.56726/irjmets39893.
- [6] M. R. Arun, M. Anto Bennet, S. D. Satav and N. A. Dawande, “Deep Learning Based Object Detection in Surveillance Video,” in *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1794–1801. DOI: 10.1109/ICOSEC58147.2023.10276262.
- ss
- [7] X. Z. Zhang, “A study on the campus public safety monitoring system based on intelligent vision,” in *Advanced Materials Research*, Trans Tech Publications Ltd, 2014, pp. 257–261. DOI: 10.4028/www.scientific.net/AMR.1028.257.
- [8] G. P. Morales Cauti, A. R. Vargas-Murillo, F. De Jesus Guevara-Soto, and I. L. Ypanaque-Pereira, “Intelligent Video Surveillance: Artificial Intelligence and its Applications on Security Systems,” in *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 986–991. DOI: 10.1109/ICOSEC58147.2023.10275947.
- [9] S. Yao, B. R. Ardabili, A. Danesh Pazho, G. A. Noghre, C. Neff, and H. Tabkhi, “Real-World Community-in-The-Loop Smart Video Surveillance System,” in *Proceedings - 2023 IEEE*

International Conference on Smart Computing, SMARTCOMP 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 183–185. DOI: 10.1109/SMARTCOMP58114.2023.00041.

- [10] C. Fontes, E. Hohma, C. C. Corrigan, and C. Lütge, “AI-powered public surveillance systems: why we (might) need them and how we want them,” *Technol Soc*, vol. 71, Nov. 2022, DOI: 10.1016/j.techsoc.2022.102137.
- [11] K. Xia, L. Zhang, S. Yuan, and Y. Lou, “SCSS: An Intelligent Security System to Guard City Public Safe,” *IEEE Access*, vol. 11, pp. 76415–76426, 2023, DOI: 10.1109/ACCESS.2023.3297643.
- [12] Q. Zhang, H. Sun, X. Wu, and H. Zhong, “Edge video analytics for public safety: A review,” *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1675–1696, Aug. 2019, DOI: 10.1109/JPROC.2019.2925910.
- [13] S. A. Velastin, B. A. Boghossian, B. P. L. Lo, J. Sun, and M. A. Vicencio-Silva, “PRISMATICA: Toward ambient intelligence in public transport environments,” *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans.*, vol. 35, no. 1, pp. 164–182, Jan. 2005, DOI: 10.1109/TSMCA.2004.838461.
- [15] A. A. Ahmed and M. Echi, “Hawk-Eye: An AI-Powered Threat Detector for Intelligent Surveillance Cameras,” *IEEE Access*, vol. 9, pp. 63283–63293, 2021, DOI: 10.1109/ACCESS.2021.3074319.
- [16] Z. Shao, J. Cai, and Z. Wang, “Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data,” *IEEE Trans Big Data*, vol. 4, no. 1, pp. 105–116, Jun. 2017, DOI: 10.1109/tbdata.2017.2715815.
- [17] B. Ionescu, M. Ghenescu, F. Rastoceanu, R. Roman, and M. Buric, “Artificial Intelligence Fights Crime and Terrorism at a New Level,” *IEEE Multimedia*, vol. 27, no. 2, pp. 55–61, Apr. 2020, DOI: 10.1109/MMUL.2020.2994403.
- [18] Mathur, Garima, and Mahesh Bunde. "Research on intelligent video surveillance techniques for suspicious activity detection critical review." In *2016 international conference on recent advances and innovations in engineering (ICRAIE)*, pp. 1-8. IEEE, 2016.
- [19] E. Badidi, K. Moumane, and F. El Ghazi, “Opportunities, Applications, and Challenges of Edge-AI Enabled Video Analytics in Smart Cities: A Systematic Review,” *IEEE Access*, vol. 11. Institute of Electrical and Electronics Engineers Inc., pp. 80543–80572, 2023. DOI: 10.1109/ACCESS.2023.3300658.
- [20] A. Kumbhar, F. Koohifar, I. Güvenç, and B. Mueller, “A Survey on Legacy and Emerging Technologies for Public Safety Communications,” *IEEE Communications Surveys and Tutorials*, vol. 19, no. 1. Institute of Electrical and Electronics Engineers Inc., pp. 97–124, Jan. 01, 2017. DOI: 10.1109/COMST.2016.2612223.

APPENDIX - A

Plagiarism Check



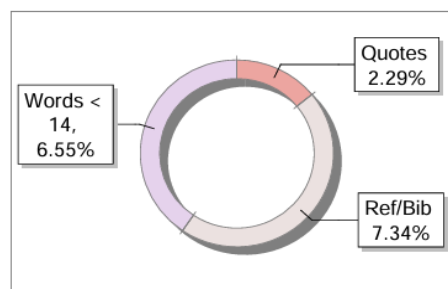
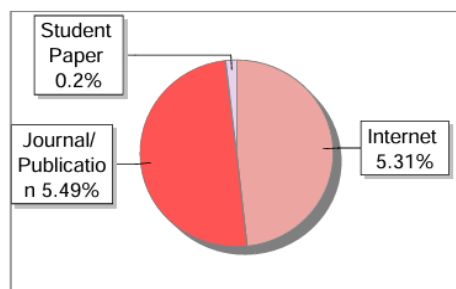
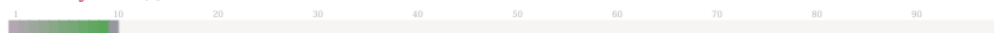
The Report is Generated by DrillBit Plagiarism Detection Software

Submission Information

Author Name	Priyamvada P
Title	VigiScape: Elevating Public Safety
Paper/Submission ID	1688986
Submitted by	anitha.ks@bmsit.in
Submission Date	2024-04-24 09:28:45
Total Pages	50
Document type	Project Work

Result Information

Similarity **11 %**




Exclude Information

Quotes	Not Excluded	Language	English
References/Bibliography	Excluded	Student Papers	Yes
Sources: Less than 14 Words %	Not Excluded	Journals & publishers	Yes
Excluded Source	0 %	Internet or Web	Yes
Excluded Phrases	Not Excluded	Institution Repository	Yes

Database Selection

A Unique QR Code use to View/Download/Share Pdf File



			
DrillBit Similarity Report			
11	74	B	A-Satisfactory (0-10%) B-Upgrade (11-40%) C-Poor (41-60%) D-Unacceptable (61-100%)
SIMILARITY %	MATCHED SOURCES	GRADE	
LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	information-science-engineering.newhorizoncollegeofengineering.in	1	Publication
2	drttit.gvet.edu.in	1	Publication
3	www.ijcrt.org	1	Publication
5	www.mdpi.com	<1	Internet Data
6	Smart Monitoring Cameras Driven Intelligent Processing to Big Surveill by Shao-2017	<1	Publication
8	www.linkedin.com	<1	Internet Data
9	www.mdpi.com	<1	Internet Data
10	www.mdpi.com	<1	Internet Data
13	dochero.tips	<1	Internet Data
15	sist.sathyabama.ac.in	<1	Publication
16	Karnataka Mang College MBA Report Submitted to MRSA, Bangalore University by SWETHA	<1	Student Paper
17	sjcit.ac.in	<1	Publication
18	zynpkucuk.medium.com	<1	Internet Data