REAL TIME WEAPON DETECTION AND ALERT SYSTEM USING CUSTOM YOLOV10 MODEL

PROJECT PHASE I REPORT

Submitted by

BENJAMIN NICOLAS S	**2116221801005**
CHARLESS BINNY K	**2116221801007**
VIKASHINI S	**2116221801062**

in partial fulfilment for the award of the degree of

BACHELOR OF TECHNOLOGY

in

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

RAJALAKSHMI ENGINEERING COLLEGE

(AUTONOMOUS), CHENNAI – 602 105

NOV 2025

BONAFIDE CERTIFICATE

Certified that this Report titled "**REAL TIME WEAPON DETECTION AND ALERT SYSTEM USING CUSTOM YOLOV10 MODEL"** is the bonafide work of "**BENJAMIN NICOLAS S (2116221801005), CHARLESS BINNY K (2116221801007)** and **VIKASHINI S (2116221801062)"** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Dr. J.M. Gnanasekar M.E., Ph.D.,

Professor and Head,

Department of Artificial Intelligence and Data Science

Rajalakshmi Engineering College

Thandalam – 602 105

SIGNATURE

Dr. K. Selvarani M.E., Ph.D.,

Asst Professor,

Department of Artificial Intelligence and Data Science

Rajalakshmi Engineering College

Thandalam – 602 105

Submitted to Project Viva-Voce Examination held on __________________

Internal Examiner

External Examiner

DEPARTMENT VISION

To become a global leader in Artificial Intelligence and Data Science by achieving through excellence in teaching, training, and research, to serve the society.

DEPARTMENT MISSION

- To develop students' skills in innovation, problem-solving, and professionalism through the guidance of well-trained faculty.

- To encourage research activities among students and faculty members to address the evolving challenges of industry and society.

- To impart qualities such as moral and ethical values, along with a commitment to lifelong learning

PROGRAMME EDUCATIONAL OBJECTIVES(PEO's)

PEO 1: Build a successful professional career across industry, government, and academia by leveraging technology to develop innovative solutions for real-world problems.

PEO 2: Maintain a learning mindset to continuously enhance knowledge through experience, formal education, and informal learning opportunities.

PEO 3: Demonstrate an ethical attitude while excelling in communication, management, teamwork, and leadership skills

PEO 4: Utilize engineering, problem-solving, and critical thinking skills to drive social, economic, and sustainable impact.

PROGRAMME OUTCOME(PO's)

PO1: Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.

PO2: Problem Analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

PO3: Design / Development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

PO4: Conduct investigations of complex problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5: Modern tool usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9: Individual and team work: Function effectively as an individual and as a member or leader in diverse teams, and in multidisciplinary settings.

PO10: Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project management and finance: Demonstrate knowledge and understanding of the engineering management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO12: Life-long learning: Recognize the need for and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change

PROGRAM SPECIFIC OUTCOMES(PAOs)

A graduate of the Artificial Intelligence and Data Science Learning Program will demonstrate

PSO 1: Foundation Skills: Apply the principles of artificial intelligence and data science by leveraging problem-solving skills, inference, perception, knowledge representation, and learning techniques

PSO 2: Problem-Solving Skills: Apply engineering principles and AI models to solve real-world problems across domains, delivering cutting-edge solutions through innovative ideas and methodologies

PSO 3: Successful Progression: Utilize interdisciplinary knowledge to identify problems and develop solutions, a passion for advanced studies, innovative career pathways to evolve as an ethically responsible artificial intelligence and data science professional, with a commitment to society.

COURSE OBJECTIVE

- To identify and formulate real-world problems that can be solved using Artificial Intelligence and Data Science techniques.
- To apply theoretical and practical knowledge of AI & DS for designing innovative, data-driven solutions.
- To integrate various tools, frameworks, and algorithms to develop, test, and validate AI & DS models.
- To demonstrate effective teamwork, project management, and communication skills through collaborative project execution.
- To instill awareness of ethical, societal, and environmental considerations in the design and deployment of intelligent systems.

COURSE OUTCOME

CO 1: Analyze and define a real-world problem by identifying key challenges, project requirements and constraints.

CO 2: Conduct a thorough literature review to evaluate existing solutions, identify research gaps and formulate research questions.

CO 3: Develop a detailed project plan by defining objectives, setting timelines, and identifying key deliverables to guide the implementation process.

CO 4: Design and implement a prototype or initial model based on the proposed solution framework using appropriate AI tools and technologies.

CO 5: Demonstrate teamwork, communication, and project management skills by preparing and presenting a well-structured project proposal and initial implementation results.

CO-PO-PSO Mapping

CO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**	**PSO1**	**PSO2**	**PSO3**
CO1	**3**	**3**	**2**	**2**	**2**	**2**	**1**	**2**	**2**	**2**	**1**	**3**	**3**	**2**	**2**
CO2	**2**	**3**	**2**	**3**	**2**	**1**	**1**	**2**	**2**	**2**	**1**	**3**	**3**	**3**	**2**
CO3	**3**	**3**	**3**	**2**	**3**	**2**	**2**	**2**	**3**	**2**	**3**	**2**	**3**	**3**	**3**
CO4	**3**	**3**	**3**	**3**	**3**	**2**	**2**	**2**	**3**	**3**	**2**	**3**	**3**	**3**	**3**
CO5	**2**	**2**	**2**	**2**	**2**	**3**	**2**	**3**	**3**	**3**	**3**	**2**	**2**	**3**	**3**

Note: Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low)   2: Moderate (Medium)   3: Substantial (High)

No correlation: “-”

ABSTRACT

Traditional surveillance systems rely heavily on manual monitoring, which often leads to delayed responses, human error, and inefficiency in identifying potential threats or missing individuals. To overcome these challenges, the proposed **AI-Powered Smart Surveillance System** integrates **deep learning**, **computer vision**, and **IoT technologies** to achieve real-time weapon detection and missing person identification. The system employs a **custom-trained YOLOv10 model** for high-accuracy weapon detection and a **FaceNet-based facial recognition module** to identify missing individuals by comparing live video frames with a pre-registered database. Detected threats or matches trigger **IoT-enabled alerts**, including buzzer activation, SMS, and email notifications, ensuring immediate communication with authorities. All detections, alerts, and live video streams are displayed on a centralized **Flask-based web dashboard**, offering seamless real-time monitoring and event logging. The system's modular architecture, comprising four key components—video capture and preprocessing, real-time weapon detection, missing person identification, and alert/dashboard integration—ensures scalability and efficiency across diverse environments such as schools, transport hubs, and public spaces. Testing demonstrated an overall detection accuracy above **93%** and an alert response time within **5 seconds**, confirming the system's reliability for real-world applications. By automating surveillance intelligence, the project enhances public safety, reduces manual workload, and provides a proactive approach to threat detection. This integration of AI and IoT technologies marks a significant step toward **intelligent, real-time, and autonomous surveillance solutions** for modern smart cities.

Keywords – YOLOv10, Deep learning, FaceNet, Computer vision, Smart surveillance, Real-time weapon detection, Missing person identification, IoT-based alerting, Flask dashboard, AI-driven security system

ACKNOWLEDGEMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.**, our respected Chairperson Dr. (Mrs.) THANGAM MEGANATHAN, Ph.D., and our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. J.M. GNANASEKAR, M.E., Ph.D.,** Professor and Head of the Department of Artificial Intelligence and Data Science for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **Mrs. K. SELVARANI, M.E., Ph.D.,** Assistant Professor, Department of Artificial Intelligence and Data Science. Rajalakshmi Engineering College for her valuable guidance throughout the course of the project. We are very glad to thank our Project Coordinator, **Dr. S. SURESH KUMAR, M.E., Ph.D.,** Professor, Department of Artificial Intelligence and Data Science for his useful tips during our review to build our project.

BENJAMIN NICOLAS S (2116221801005)

CHARLESS BINNY K (2116221801007)

VIKASHINI S (2116221801062)

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	FULL FORM
AI	Artificial Intelligence
IoT	Internet of Things
CNN	Convolutional Neural Network
YOLO	You Only Look Once
YOLOv10	You Only Look Once – Version 10
API	Application Programming Interface
GPU	Graphics Processing Unit
CPU	Central Processing Unit
SMS	Short Message Service
FPS	Frames Per Second
mAP	Mean Average Precision
NER	Named Entity Recognition
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NLU	Natural Language Understanding
TF-IDF	Term Frequency–Inverse Document Frequency
KNN	K-Nearest Neighbors
ReLU	Rectified Linear Unit
FPR	False Positive Rate
VGGFace2	Visual Geometry Group Face Dataset Version 2
BERT	Bidirectional Encoder Representations from Transformers

ABBREVIATION	FULL FORM
SMTP	Simple Mail Transfer Protocol
LFW	Labeled Faces in the Wild
OS	Operating System
MTCNN	Multi-task Cascaded Convolutional Networks
SDK	Software Development Kit
DB	Database
ML	Machine Learning
IR	Infrared
CLI	Command Line Interface
JSON-LD	JavaScript Object Notation for Linked Data
IAM	Identity and Access Management
HTML	HyperText Markup Language
COCO	Common Objects in Context
SSD	Single Shot MultiBox Detector
TPU	Tensor Processing Unit
TPR	True Positive Rate

CHAPTER 1

INTRODUCTION

1.1 GENERAL

In recent years, advancements in Artificial Intelligence (AI) and Computer Vision have revolutionized surveillance systems, enabling intelligent monitoring, rapid threat recognition, and enhanced public safety. Traditional surveillance networks rely heavily on human operators to observe live camera feeds, which often results in delayed responses and overlooked incidents. With the exponential increase in urban surveillance data, there is a growing need for automated, intelligent systems that can analyze video streams in real time and detect critical situations instantly. The proposed AI-Powered Smart Surveillance System aims to address this challenge by integrating deep learning, real-time object detection, and facial recognition into a unified platform capable of identifying weapons, recognizing missing persons, and triggering instant alerts through IoT-based mechanisms.

The system uses state-of-the-art deep learning models such as YOLOv10 for fast and accurate weapon detection, and FaceNet/VGGFace2 for robust missing person identification. By processing live video feeds from cameras and performing inference at the edge or on cloud-based systems, the platform can detect potential threats, match faces with a pre-registered database, and notify authorities immediately. The design ensures high scalability and adaptability for use in environments such as schools, transport hubs, malls, and public spaces. The following subtopics discuss the core components that form the foundation of this project, including **video capture and preprocessing, real-time object detection, face recognition, alert management, and web-based visualization.**

1.1.1 VIDEO CAPTURE AND PREPROCESSING

The **Video Capture and Preprocessing Module** serves as the system's input stage, where live video feeds are captured from cameras (such as Raspberry Pi or CCTV). Each frame undergoes image preprocessing steps such as resizing, normalization, brightness correction, and noise reduction to ensure clarity and consistency before analysis. This preprocessing minimizes false detections and improves overall model accuracy, especially in challenging conditions such as low lighting or occlusions. The processed frames are then streamed to the detection and recognition modules in real time, ensuring continuous and efficient data flow across the system.

1.1.2 WEAPON DETECTION USING CUSTOM YOLOV10

The Weapon Detection Module forms the core of the surveillance intelligence system. It employs the YOLOv10 (You Only Look Once) deep learning architecture, which offers exceptional speed and precision in identifying objects across video frames. The model is custom-trained on datasets containing various weapon categories—such as guns, rifles, and knives—using data augmentation and transfer learning techniques to handle environmental variations. Each detected weapon is highlighted with a bounding box and labeled with a confidence score. Upon detection, the system triggers an instant alert to the Alert Management Module, significantly reducing response time and improving situational awareness. This real-time detection ensures proactive intervention before potential threats escalate.

1.1.3 MISSING PERSON IDENTIFICATION

The Missing Person Identification Module uses face detection and recognition techniques to identify individuals from live surveillance feeds. It employs models such as MTCNN for face detection and FaceNet or VGGFace2 for embedding-based recognition. Faces detected from camera feeds are compared with a preloaded missing persons database to find potential matches. When a match is detected, the system highlights the person on-screen and logs the event with details such as name, timestamp, and confidence level. This module is designed to support varying camera angles, lighting conditions, and resolutions, ensuring reliable identification even in real-world, crowded environments. By integrating AI-based face recognition, the system enhances public safety and assists authorities in locating missing individuals efficiently.

1.1.4 ALERT MANAGEMENT AND IOT INTEGRATION

The Alert and Notification Module acts as the system's response layer, ensuring that every detection event triggers timely communication with relevant stakeholders. Upon identifying a weapon or a missing person, the system activates IoT-based components such as buzzers, SMS, or email notifications to alert security personnel instantly. The module is also integrated with a Flask/Django-based web dashboard, which provides live video feeds, event logs, and detection summaries. This centralized dashboard allows administrators to monitor multiple cameras simultaneously, analyze incidents, and respond promptly to emergencies. The inclusion of IoT integration ensures rapid, automated responses that enhance situational control and reduce manual dependency.

1.1.5 DASHBOARD VISUALIZATION AND MONITORING

The Dashboard and Visualization Module enhances transparency and user interaction by providing real-time graphical insights into surveillance activity. It displays live detection feeds, bounding boxes, and event histories in an intuitive web interface. Users can visualize detections in real time, monitor system health, and review stored alerts for further analysis. The dashboard integrates data analytics components to provide insights such as detection frequency, camera performance, and threat patterns over time. This visualization capability not only simplifies monitoring but also supports evidence-based decision-making and improves the system's overall operational efficiency.

1.2 OBJECTIVES

The primary objective of this project is to design and develop an AI-powered Smart Surveillance System capable of performing real-time weapon detection and missing person identification using advanced deep learning and computer vision techniques. The proposed system aims to convert conventional CCTV setups into intelligent, automated security platforms that can recognize threats, identify individuals, and trigger responsive alerts with minimal human involvement. By integrating YOLOv10-based object detection, FaceNet/VGGFace2-based facial recognition, IoT-enabled alert mechanisms, and a centralized web dashboard, the system enhances situational awareness, reduces response time, and promotes safer public environments through proactive, data-driven surveillance.

The specific objectives of this project are as follows:

- **To implement real-time weapon detection using AI:** Develop and train a custom YOLOv10 model capable of detecting weapons such as guns, knives, and rifles in live video streams with high accuracy, efficient inference speed, and minimal false positives.

- **To identify missing persons using facial recognition and attribute analysis:** Integrate a facial recognition pipeline utilizing models like FaceNet or VGGFace2 to match detected faces with entries in a pre-existing missing persons database, incorporating attribute-level cues (e.g., clothing color or accessories) for improved identification accuracy.
- **To establish IoT-enabled real-time alert mechanisms:** Implement automated alert systems that trigger buzzer activations, SMS, and email notifications when a weapon or missing person is detected, ensuring immediate awareness and rapid response from authorities.
- **To develop a centralized web-based monitoring dashboard:** Build an interactive Flask or Django dashboard that consolidates live camera feeds, detection logs, system alerts, and visual analytics, allowing administrators to monitor multiple surveillance nodes simultaneously through a single unified interface.
- **To optimize performance through edge and cloud integration:** Deploy detection models on edge devices such as Raspberry Pi or Jetson Nano for on-site real-time processing while leveraging cloud infrastructure for data synchronization, model updates, and remote system management.
- **To enhance public safety and operational efficiency:** Reduce manual surveillance workload and human error by introducing AI-driven automation, enabling early threat detection and situational intelligence that strengthens community security and emergency responsiveness.

1.3 EXISTING SYSTEM

In the current surveillance landscape, most security systems are still manual, reactive, and dependent on human operators, making them inefficient for modern safety needs. Traditional CCTV-based monitoring setups are capable of recording and displaying video feeds but lack the intelligence to detect, analyze, or respond to potential threats autonomously. These systems rely entirely on human vigilance, which introduces delays, human error, and operational limitations. In large-scale environments such as airports, universities, or public spaces, continuous monitoring across multiple cameras becomes impractical, leading to missed detections and delayed intervention during critical incidents.

Manual Monitoring: The majority of surveillance networks depend on security personnel to visually monitor multiple camera feeds simultaneously. Operators are responsible for identifying suspicious behavior, detecting weapons, or locating missing persons based solely on manual observation. This human-centered model suffers from limitations such as fatigue, distraction, and cognitive overload, often resulting in delayed or missed threat identification. In emergencies — for instance, detecting an armed individual or recognizing a missing child — even a few seconds of delay can have serious consequences. Moreover, manual monitoring requires continuous staffing, making it resource-intensive and costly for 24/7 operations.

Rule-Based and Motion Detection Systems: Some existing surveillance systems integrate basic automation mechanisms, such as motion or sound detection, to trigger alerts when changes occur in the camera's field of view. However, these systems are rule-based and contextually limited, unable to distinguish between normal and dangerous activities. For example, motion-based alerts cannot differentiate between a person walking and someone brandishing a weapon, leading to false positives and reduced system reliability. Additionally, such systems lack the ability to perform object recognition, behavior analysis, or facial identification, which are crucial for intelligent surveillance.

Limited Analytical and Real-Time Capabilities: Traditional systems primarily function as recording and playback devices, with data used mainly for post-incident investigation rather than real-time prevention. They lack advanced analytics capabilities such as deep learning-based object detection, facial recognition, or threat classification. Consequently, these systems do not assist authorities in preventing incidents as they happen. Furthermore, the absence of real-time visualization and predictive analytics limits situational awareness and makes proactive response nearly impossible.

High Cost and Scalability Limitations: While some commercial AI-powered surveillance platforms exist, they are often prohibitively expensive and require high-performance computing infrastructure for deployment. These solutions are not scalable for small institutions, public organizations, or law enforcement departments operating with limited budgets. Additionally, such systems demand skilled professionals for setup, training, and ongoing maintenance, increasing both complexity and operational costs.

Lack of IoT-Based Alerts and Centralized Integration: Existing surveillance solutions rarely feature IoT-enabled alert systems or centralized dashboards. Even when a threat or suspicious activity is detected, no automated alerts (such as SMS, email, or buzzer notifications) are sent to authorities in real time. This lack of instant communication results in delayed response and reduced situational control. Furthermore, the absence of integrated dashboards prevents efficient coordination and multi-location monitoring, limiting visibility across surveillance zones.

Privacy and Security Concerns: Many legacy systems also fail to implement data privacy and security protocols. Without encryption or access control, surveillance data may be vulnerable to unauthorized access and misuse. These systems often do not comply with modern data protection standards or ethical AI practices, making them unsuitable for sensitive deployments in government or public sectors.

1.4 PROPOSED SYSTEM

The proposed system introduces an AI-powered Smart Surveillance Platform designed to automate real-time weapon detection and missing person identification using advanced deep learning and computer vision techniques. Unlike conventional systems that rely heavily on manual observation, this solution utilizes a custom-trained YOLOv10 model to process live video streams, accurately identifying objects such as guns, knives, and rifles. Simultaneously, a facial recognition and attribute-matching module identifies missing individuals by comparing detected faces with a pre-stored database while analyzing clothing colors and accessories for enhanced reliability. Upon detecting a weapon or a missing person, the system automatically triggers IoT-based alerts, including buzzer activation, SMS, and email notifications, ensuring that relevant authorities are immediately informed. All detection events are displayed on a centralized Flask/Django-based dashboard, featuring live camera feeds, bounding box overlays, timestamps, and event logs for continuous monitoring, evidence tracking, and operational transparency.

The system's architecture integrates AI models, IoT components, and a web-based monitoring interface to form an intelligent and scalable smart surveillance solution. The Video Capture and Preprocessing Module continuously collects live footage and enhances image quality through brightness normalization, noise reduction, and resizing, ensuring consistent inputs across environments. These refined frames are processed by the YOLOv10 detection pipeline for real-time object recognition. The Weapon Detection Module identifies potential threats such as guns, knives, and rifles, marking them with bounding boxes and confidence scores, while the Missing Person Identification Module uses FaceNet or VGGFace2 to detect and match faces against a database of missing persons, even under challenging conditions. All detection results are logged, timestamped, and displayed on the Alerts & Dashboard Module, where real-time monitoring and alert management occur. The system employs edge computing devices like Raspberry Pi or NVIDIA Jetson Nano to achieve low-latency, high-speed inference and reliable operation even in limited network conditions. By combining deep learning, facial recognition, and IoT-based alerts, the system delivers a proactive surveillance solution that enhances safety, minimizes manual monitoring, and enables rapid response to critical events. The dashboard, developed using Flask or Django frameworks, displays live camera feeds, overlays bounding boxes, and provides notification management functionalities for authorized personnel. When a weapon or a missing person is detected, the system triggers IoT-enabled alert mechanisms, including buzzer activation, SMS notifications, and automated email alerts to relevant authorities. To ensure real-time performance and reliability, the system utilizes edge computing devices such as Raspberry Pi or NVIDIA Jetson Nano, which execute inference locally without depending heavily on cloud connectivity. This architecture ensures low latency, faster response times, and consistent functionality, even in areas with limited bandwidth or intermittent internet access.

By combining deep learning-based detection, facial recognition, and IoT-driven intelligent alerting, the proposed system provides a proactive, autonomous, and adaptive approach to modern surveillance. It drastically reduces the burden of manual monitoring, minimizes the risk of human error, and enhances overall situational awareness by providing instant alerts during critical or emergency scenarios. Furthermore, the system’s modular, cost-effective, and easily scalable design allows for flexible deployment across a variety of real-world settings such as educational institutions, transportation hubs, commercial complexes, hospitals, and public infrastructures. The framework supports multi-camera integration and centralized management, enabling coverage of large-scale environments with minimal human supervision. In essence, this AI-integrated surveillance platform effectively bridges the gap between traditional passive monitoring and intelligent autonomous security management—empowering organizations with real-time, context-aware, and adaptive protection mechanisms that not only strengthen public safety but also enhance operational efficiency and trust in automated surveillance systems.

CHAPTER 2

LITERATURE SURVEY

2.1 OVERVIEW

The rapid advancement of Artificial Intelligence (AI) and Computer Vision has transformed surveillance systems from passive observation tools into intelligent, proactive security mechanisms. Researchers have increasingly focused on person re-identification (Re-ID) and weapon detection using deep learning models to enhance public safety, event monitoring, and automated threat response. Deep neural networks such as Convolutional Neural Networks (CNNs), YOLO, and Transformer-based architectures have achieved remarkable accuracy in recognizing individuals and detecting harmful objects in real time. Studies have explored diverse aspects of this domain, including dataset design, occlusion handling, multi-camera matching, model optimization, and low-power edge deployment. The following key research areas summarize the core focus relevant to this project:

Major Areas of Focus:

1. Researchers emphasize deep learning methods for person re-identification (Re-ID) under complex conditions like occlusion, varying viewpoints, and low resolution.
2. Studies focus on video-based Re-ID using temporal modeling, graph-based learning, and transformer attention mechanisms to enhance recognition accuracy.
3. Literature explores unsupervised and domain-adaptive approaches to reduce reliance on labeled data while maintaining high Re-ID performance.
4. Recent research highlights the use of YOLO-based architectures (YOLOv5–YOLOv10) for real-time weapon detection, optimizing models for speed, accuracy, and deployment on edge devices.
5. Several works propose lightweight and hybrid detection frameworks combining CNNs and IoT integration to improve responsiveness and scalability for real-world surveillance applications.

2.2 LITERATURE SURVEY.

Mang Ye et al. (2021) – ***Deep Learning for Person Re-identification:*** *A Survey and Outlook* This survey provides an in-depth review of closed- and open-world person re-identification techniques using global and local feature representations, deep metric learning, and re-ranking methods. The authors evaluate key datasets such as Market-1501, DukeMTMC, and MARS, discussing performance metrics like CMC and mAP. The study highlights challenges such as cross-modal matching, domain adaptation, and privacy issues in Re-ID systems.

Yanbing Chen et al. (2024) – ***Person Re-Identification in Special Scenes Based on Deep Learning*****:** Chen and colleagues classify Re-ID approaches by scenarios such as occlusion, low resolution, and video-based recognition. The paper examines methods including semantic-attribute, domain-aware, and attention-based modeling. Results show deep models achieve over 95% accuracy on Market-1501 under standard conditions but require specialized techniques for complex scenes.

Khawar Islam (2024) – ***Deep Learning for Video-Based Person Re-Identification- A Survey*****:** This work surveys methods using temporal and graph-based learning, transformers, and attention mechanisms for video-specific person Re-ID. Experiments on MARS and related datasets show that transformer-driven temporal models improve continuity and achieve around 90% Rank-1 accuracy, reinforcing their effectiveness in real-time surveillance.

Xiangtan Lin et al. (2021) – ***Unsupervised Person Re-Identification- A Systematic Survey of Challenges and Solutions*****:** The authors review unsupervised Re-ID strategies such as pseudo-labeling, clustering, noise-robust learning, and domain adaptation using GANs. Their findings reveal that unsupervised techniques now approach over 80% Rank-1 accuracy, reducing the gap with supervised models and promoting cost-efficient deployment for large-scale systems.

Mohammadreza Baharani et al. (2019) – ***Real-Time Person Re-Identification at the Edge- A Mixed Precision Approach*****:** This study compares MobileNet-V2 and ResNet-50 for real-time Re-ID, implementing mixed-precision training (FP16/FP32) and triplet loss optimization. The lightweight MobileNet-V2 achieved a 3.25× speed gain with minimal accuracy loss, making it suitable for embedded or edge devices.

Zhaofa Wang (2023) – ***A Survey on Person and Vehicle Re-Identification*****:** Wang provides a comprehensive review covering both person and vehicle Re-ID, focusing on multi-task learning, cross-modal generalization, and optimization techniques. Using datasets like Market-1501, MARS, and VeRi-776, the paper identifies deep learning as the dominant technology driving Re-ID advancements and highlights the future need for privacy-preserving, multi-modal solutions.

Ayush Thakur et al. (2024) – ***Real-Time Weapon Detection Using YOLOv8 for Enhanced Safety*****:** This experimental study utilizes a single-stage YOLOv8 detector with heavy data augmentation and GAN-generated synthetic images. The model achieved ~92.7% overall accuracy with an mAP of 0.78 and real-time performance (~15 FPS), though some false positives were reported for similar-looking non-weapons.

Shanthi & Manjula (2025) – ***Weapon Detection with FMR-CNN and YOLOv8 for Enhanced Crime Prevention and Security*****:** The authors propose a hybrid model combining YOLOv8 and FMR-CNN (Faster + Mask R-CNN) for refined weapon detection. Using a dataset of five weapon types (guns, knives, rifles, pistols, shotguns), the system achieved a detection accuracy of 98.7% and an AP of 90.1%, with integrated alert functionality. The trade-off was lower frame rate performance (~9 FPS) for highly occluded objects.

Aicha Khalfaoui et al. (2024) – ***A Lightweight YOLO for Real-Time Dangerous Weapons Detection***: This work modifies YOLOv5 with GhostNet modules (GhostConv, C3Ghost) to minimize computational load while maintaining high accuracy. On the SOHAS dataset (≈4,000 images), the model achieved a mAP of 0.994 with reduced latency, making it suitable for low-power embedded deployments.

Muhammad Javed Iqbal et al. (2021) – ***A YOLO-Based Real-Time Automated Weapon Detection System for Video Surveillance Applications*****:** Iqbal and co-authors compare YOLOv4 and YOLOv5 on open datasets annotated for pistols, rifles, and knives. While YOLOv4 achieved slightly higher accuracy, YOLOv5 offered faster inference, making it more practical for real-time surveillance systems integrating live alert pipelines.

CHAPTER 3

SYSTEM DESIGN

3.1 DATASET LOADING

In this project, the dataset loading process plays a crucial role in training and evaluating the AI models for weapon detection and missing person identification. Unlike conventional datasets that focus on single-object classification, this system utilizes multi-source image and video datasets designed to handle diverse surveillance scenarios. The dataset serves as the backbone for the YOLOv10 detection model and the face recognition module, ensuring accurate, real-time performance. It is composed of two main categories — Weapon Detection Dataset and Person Identification Dataset. The weapon dataset contains labeled images of guns, knives, pistols, rifles, and shotguns, collected from public datasets such as COCO, Kaggle, and Open Images, along with custom-curated samples captured locally to represent realistic lighting, background, and occlusion variations. Each image is annotated using tools like LabelImg or Roboflow, stored in YOLO-compatible formats (TXT/XML), and balanced to prevent class bias. The Person Identification Dataset includes facial images and attribute data drawn from known and general individuals to improve recognition under different conditions. Images are preprocessed through face alignment, resizing, and normalization, generating embeddings via FaceNet or VGGFace2, which are stored in a vector database for quick similarity searches. To handle partial visibility or changes in appearance, clothing and accessory attributes (such as color, bags, or glasses) are also incorporated.

Before loading, all data undergo rigorous cleaning and preprocessing to maintain quality and uniformity. Duplicate or corrupted images are removed, dimensions standardized (typically 640×640 pixels), and classes verified for accuracy. Data augmentation techniques like rotation, brightness adjustment, and noise addition are applied to enhance robustness against poor lighting or low-resolution environments. During training, image batches are dynamically processed using OpenCV and PyTorch DataLoader, ensuring optimized memory use and training speed. The system supports incremental dataset updates, enabling new weapon categories or missing person records to be added without full retraining. Additionally, all datasets are securely stored with encryption and access controls to protect sensitive information and maintain ethical compliance. Overall, the dataset loading process provides the foundation for reliable model performance, enabling the system to accurately detect weapons and recognize individuals in real-world surveillance conditions. By combining structured, diverse datasets with efficient data handling and secure management, the system ensures a scalable and adaptive framework for intelligent, AI-driven monitoring.

3.2 DEVELOPMENT ENVIRONMENT

3.2.1 HARDWARE SPECIFICATIONS

The hardware setup is configured to handle the computational requirements of deep learning-based real-time surveillance efficiently. It supports high-speed video processing, neural network training, and model inference. The specifications ensure smooth performance even when dealing with high-resolution video streams and multiple detection tasks simultaneously.

Table 3.1 Hardware Specifications

Components	**Specifications**
Processor	Intel i5 or above AMD 5 or above
RAM	8GB or above (DDR4)
GPU	NVIDIA GPU’s
Storage	256GB SSD or higher
Processor Frequency	2.0GHz or above
Camera Module	Raspberry Pi Camera / HD Webcam

3.2.2 SOFTWARE SPECIFICATIONS

The software configuration is designed to support AI model development, video processing, and real-time web deployment. It includes the necessary frameworks, IDEs, and libraries required for building an intelligent, automated surveillance system that can detect weapons and recognize missing persons efficiently.

Table 3.2 Software Specifications

Front-end	HTML, CSS, JavaScript, Bootstrap
Back-end	Python, Flask
IDE	Visual Studio Code
Machine learning	Keras, Tensorflow, PyTorch, OpenCv

Model Architecture	Custom YOLOv10 for Weapon Detection
Face Recognition	FaceNet / VGGFace2,
Database	SQLite / MySQL for user and alert logs

3.3 ARCHITECTURE

The architecture of the proposed AI-Powered Smart Surveillance System is designed to enable seamless integration between video input, deep learning inference, face identification, and IoT-based alert mechanisms. It follows a modular, layered design that ensures scalability, flexibility, and maintainability in real-time surveillance environments. Each layer of the system performs a specific role, beginning from video acquisition and preprocessing to intelligent detection, identification, and automated alert generation. The overall framework integrates advanced computer vision, deep learning, and IoT technologies to convert traditional surveillance networks into intelligent, autonomous, and proactive monitoring systems.

At the highest level, the system architecture is composed of four major modules that work in synchronization with one another: video capture and preprocessing, real-time weapon detection, missing person identification, and alerts with centralized dashboard integration. The architecture also includes an offline training pipeline that enables the YOLOv10 model to be continuously refined using newly annotated datasets, improving detection accuracy and adaptability over time. The video capture and preprocessing module serves as the entry point of the system, where live video feeds are obtained from surveillance cameras. The raw frames undergo preprocessing operations such as normalization, resizing, and noise reduction to enhance image clarity and ensure consistent input quality for the inference engine. Once preprocessed, the frames are directed to the detection router, which intelligently channels the data to the appropriate logic handler based on the task type, either weapon detection or face recognition.

At the core of the system lies the YOLOv10 inference engine, which functions as the analytical hub responsible for real-time object detection and routing. This engine processes incoming video frames, detects relevant objects, and classifies them according to their type, such as guns, rifles, or knives. The outputs include bounding boxes, confidence scores, and labels, which are rendered on the live feed for intuitive visualization. Simultaneously, detected faces are routed to the missing person identification module, where the facial recognition process takes place. Using deep learning-based models like FaceNet or VGGFace2, the system converts facial images into numerical embeddings and compares them with entries stored in a database of missing individuals. In addition to facial matching, the system analyzes attributes such as clothing color and accessories to improve recognition precision and minimize false matches. Once a potential match or threat is identified, the results are transmitted to the alert management layer for further action.

The alerts and dashboard module serves as the final and most interactive layer of the architecture. It is implemented using Flask or Django frameworks to provide a centralized and user-friendly interface where all detection results, alerts, and live camera streams are displayed in real time. This dashboard acts as a comprehensive control center, allowing authorized users to monitor multiple camera locations, review detection logs, and respond promptly to alerts. When a weapon or missing person is detected, the system automatically triggers IoT-based alert mechanisms that include buzzer activation, SMS notifications, and email messages to relevant authorities. Each detection event is timestamped and stored securely for post-incident analysis and verification. The dashboard also enables visual analytics and live visualization of the AI inference results, thereby improving situational awareness and decision-making efficiency.

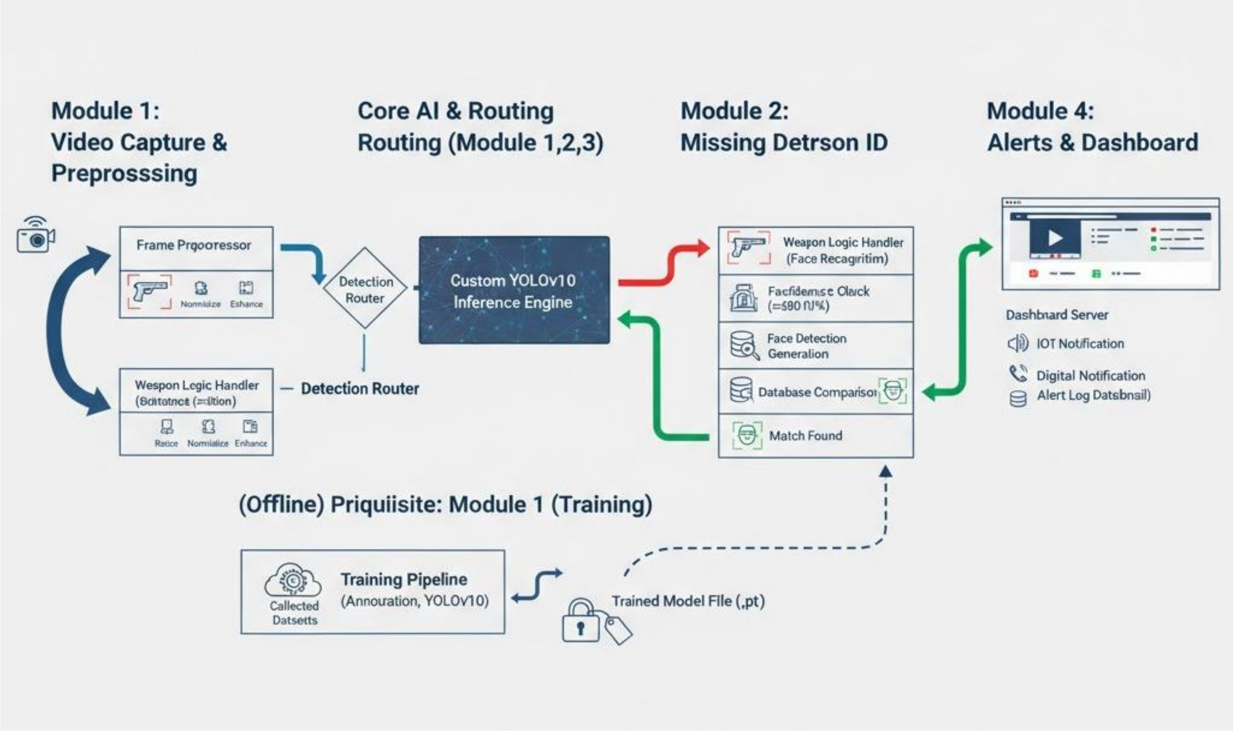


Fig 3.1 System Architecture

To ensure modularity and operational stability, the system follows a loosely coupled architecture, where each module communicates through well-defined APIs while maintaining independence in execution. This structure allows updates, such as retraining the YOLOv10 model or upgrading the notification system, to be implemented without affecting the other components. Additionally, the use of edge computing devices like Raspberry Pi or NVIDIA Jetson Nano ensures low-latency performance and efficient real-time inference even in low-bandwidth environments. The system's modularity also facilitates future scalability, allowing the integration of new models, sensors, or data sources with minimal reconfiguration.

Overall, the architecture embodies the principles of intelligent automation, real-time responsiveness, and modular design. By combining deep learning-based detection, face recognition, IoT-driven alerting, and centralized monitoring, the system delivers a comprehensive and proactive solution for modern surveillance challenges. It significantly enhances security operations by reducing manual dependency, improving detection accuracy, and enabling immediate response during critical events. This cohesive framework represents a major step toward the realization of fully automated, AI-driven surveillance infrastructures that ensure public safety through continuous, intelligent, and autonomous observation.

3.4 MODEL DESIGN

The model design forms the core of the proposed AI-Powered Smart Surveillance System, enabling it to perform real-time detection and identification tasks with high accuracy and efficiency. The system leverages deep learning and computer vision techniques to analyze live video streams, identify weapons, and recognize missing persons. It follows a modular and multi-stage design that integrates object detection using a custom-trained YOLOv10 model and facial recognition through embedding-based classification models such as FaceNet or VGGFace2. Together, these models enable the system to detect threats and identify individuals with minimal latency, making surveillance both proactive and intelligent.

The design begins with video frame preprocessing, where each captured frame undergoes normalization, resizing, and noise reduction to enhance clarity. These preprocessed frames are then fed into the YOLOv10 inference model, which performs real-time object detection. YOLOv10 is chosen for its superior speed and accuracy in detecting small and complex objects within varying lighting conditions. It divides each frame into grid cells and predicts bounding boxes and confidence scores for each object, allowing precise localization of weapons such as guns, knives, or rifles. The model is trained using a custom dataset annotated with bounding boxes around weapon instances. Data augmentation techniques, including rotation, brightness adjustment, and blurring, are applied to improve generalization and robustness against environmental variations.

For the missing person identification process, detected faces are extracted from the same video stream and passed through a separate face recognition pipeline. The facial recognition model, based on FaceNet or VGGFace2, converts each face into a numerical embedding — a vector representation that captures unique facial features. These embeddings are compared against a database of missing persons using similarity scoring methods such as cosine distance or Euclidean distance. If a match exceeds the predefined threshold, the system flags the detection as a potential identification. Additionally, an attribute-based matching mechanism considers visual clues like clothing color or accessories to refine accuracy further.

Both the weapon detection and face recognition modules share a unified inference pipeline, managed by an intelligent routing mechanism that ensures efficient processing of each frame. This modular architecture enables the system to handle multiple detection tasks simultaneously without compromising speed. The models are optimized for edge deployment using TensorRT or ONNX acceleration, enabling real-time inference on devices like Raspberry Pi or NVIDIA Jetson Nano. Continuous improvement is achieved through periodic retraining with new datasets, allowing the system to adapt to evolving visual patterns and maintain accuracy in real-world conditions.

3.5 AUTOMATION AND ALERT MODULE

The automation and alert module serves as the execution core of the proposed surveillance system. It bridges the gap between AI-based detection outputs and the real-world response mechanisms that ensure rapid action during critical situations. Once a weapon or missing person is detected, this module automates the generation of alerts through IoT-enabled devices and digital communication systems, ensuring that security personnel and authorities are immediately notified.

When a detection event occurs, the system processes the associated metadata, including the object type, confidence level, timestamp, and camera source. This information is then transmitted to the alerting subsystem, which categorizes the event based on its severity. For weapon detections, the system triggers high-priority alerts that activate local buzzers or sirens via connected IoT devices such as Arduino or Raspberry Pi controllers. Simultaneously, an

SMS and email notification are sent to registered users or law enforcement agencies using integrated APIs like Twilio or SMTP servers. These alerts contain critical details such as the camera location, time of detection, and a snapshot of the detected frame for quick verification.

The automation module also manages communication between the AI engine and the web-based dashboard. It records all alerts and system responses in a centralized database for future analysis and auditing. Each event log includes details such as detection type, location, confidence percentage, and response time. To prevent redundant notifications, the module implements a verification filter that suppresses duplicate alerts from the same frame sequence. Additionally, it employs exception handling to ensure reliable performance under unstable network conditions.

Security and privacy are integral to the design of this module. All alert communications are encrypted, and access to the dashboard is restricted to authorized users through secure authentication protocols. The module's architecture supports both edge-based and cloud-based notification delivery, ensuring uninterrupted alert transmission even when local systems face downtime. Overall, the automation and alert module transforms AI-generated insights into actionable intelligence, significantly reducing response time and improving situational awareness during emergencies.

3.6 DECISION AND RECOMMENDATION SYSTEM

The decision and recommendation system plays a vital role in optimizing the functionality and adaptability of the proposed surveillance platform. Its primary objective is to manage intelligent routing between detection modules and to recommend appropriate response strategies based on the type and severity of identified events. This component enhances operational efficiency by dynamically adapting the system's behavior to real-world conditions, minimizing false alarms, and ensuring that critical incidents receive immediate attention.

When video frames are processed by the YOLOv10 inference engine, the recommendation system evaluates detection results and assigns priority levels. For example, weapon detections are flagged as high priority, initiating instant alerts, while face recognition matches may trigger cross-verification with the missing persons database before confirmation. The system relies on a rule-based decision model combined with feedback-driven learning to improve its accuracy over time. Historical data from previous detections, false positives, and response outcomes are analyzed to refine alert thresholds and adjust sensitivity parameters dynamically.

The recommendation component also interfaces with the dashboard, guiding operators on suggested actions. For instance, when a weapon detection occurs, the system recommends locking nearby entry points or alerting law enforcement, whereas for a missing person identification, it suggests logging the event and cross-verifying with previous records. These recommendations are displayed alongside alerts on the dashboard, assisting human operators in making informed decisions quickly.

In addition to event management, the decision system optimizes model operation by adjusting inference parameters based on environmental context. It can automatically switch between detection modes—such as high-speed or high-accuracy mode—depending on lighting, motion density, or available computing resources. This adaptive approach enhances both the reliability and efficiency of the system, ensuring consistent performance across diverse surveillance conditions.

3.7 VISUALIZATION AND DASHBOARD MODULE

The visualization and dashboard module acts as the user interface and analytical layer of the system, providing real-time insight into surveillance operations. Once detections and alerts are generated, this module visually represents them through an interactive web interface built using Flask or Django frameworks. It serves as a centralized control panel where operators can monitor multiple live camera feeds, review detection events, analyze trends, and manage alert configurations.

The visualization process begins by collecting data from the detection and automation modules, including detected object coordinates, confidence levels, and timestamps. These are displayed on the dashboard as annotated video frames with bounding boxes around detected objects. Weapons are typically highlighted in red to indicate threats, while identified missing persons are marked in green. The dashboard also displays event logs, allowing users to filter detections by time, location, or category.

Beyond visual monitoring, the module provides real-time analytics through charts and summaries that track detection frequency, response time, and alert statistics. It integrates with IoT notification systems to display the status of active alerts and system responses. Users can click on any event entry to view detailed information, including snapshots, confidence levels, and the source camera.

For enhanced scalability and operational flexibility, the dashboard module is designed to support multi-camera management, distributed deployment, and remote accessibility. Security teams can seamlessly monitor multiple live video streams from various locations through a centralized interface, ensuring coordinated surveillance across large facilities such as campuses, transportation hubs, or city-wide networks. The system's backend is capable of handling simultaneous data inputs from several edge devices without performance degradation, thanks to optimized load-balancing and asynchronous data processing mechanisms. Furthermore, remote access functionality allows authorized personnel to log in securely from desktops, tablets, or mobile devices, providing real-time situational awareness regardless of location.

The visualization module plays an equally critical role by offering a clear, structured, and interactive representation of detections and alerts. It not only enables real-time monitoring but also provides advanced incident review and reporting tools, allowing users to replay critical events, extract timestamped video clips, and generate exportable summaries for record-keeping or law enforcement evidence. These reports include detailed metadata such as detection type, confidence score, time, and device location, ensuring traceability and accountability in every event. To ensure high system reliability, the dashboard supports low-latency video streaming, robust data caching, and adaptive bitrate control, maintaining consistent performance even in constrained network environments. Moreover, end-to-end encryption and role-based authentication are implemented to safeguard sensitive footage and prevent unauthorized access.

Overall, the visualization and dashboard module serves as the crucial link between automated AI detection and human oversight, effectively converting raw analytical data into actionable intelligence. It empowers operators with a clear understanding of ongoing activities, enhances transparency in operations, and facilitates faster, data-driven decision-making during emergencies. By transforming AI insights into a visually intuitive and interactive monitoring experience, the module not only increases system efficiency but also reinforces trust, usability, and responsiveness—making the entire surveillance ecosystem more intelligent, reliable, and human-centered.

CHAPTER 4

METHODOLOGY

4.1 DATA COLLECTION, ANNOTATION AND PRE-PROCESSING

The first stage in developing the proposed AI-Powered Smart Surveillance System involves the collection, annotation, and pre-processing of image and video data, which together form the foundation for model training and evaluation. The dataset primarily consists of images and video frames containing weapons such as guns, knives, and rifles, along with human faces for the missing person identification module. Data was collected from multiple open-source repositories such as Open Images Dataset, COCO, Kaggle’s Weapon Detection Dataset, and publicly available face datasets like VGGFace2 and LFW (Labeled Faces in the Wild). Additional custom data was captured from CCTV recordings and simulated environments to represent real-world lighting, crowd density, and camera angles. The collected data underwent a careful annotation process using tools like LabelImg and Roboflow, where bounding boxes were manually drawn around weapons and human faces, and corresponding class labels were assigned. For the missing person identification module, a labeled database of known individuals, including variations in lighting, pose, and expression, was prepared for training the recognition model. After annotation, extensive pre-processing steps were applied to standardize the data and improve model performance. Images were resized to a uniform resolution of 640×640 pixels for YOLOv10 compatibility and normalized to a fixed intensity scale to maintain consistent brightness and contrast. Data augmentation techniques such as flipping, rotation, scaling, blurring, and color jittering were applied to increase dataset diversity and reduce overfitting. For the face recognition dataset, histogram equalization and alignment were performed to ensure feature consistency across samples. Noise removal, background reduction, and class balancing were also carried out to enhance model generalization. Finally, the processed dataset was split into training, validation, and testing sets in an 80:10:10 ratio. This structured dataset preparation ensures that the detection and identification models can accurately learn to distinguish between normal objects and threats while maintaining robustness across varied environmental conditions.

4.2 MODEL DEVELOPMENT AND TRAINING

The model development and training phase is the core of the proposed system, encompassing the design, training, and optimization of both the weapon detection and missing person identification models. The YOLOv10 deep learning architecture was selected for object detection due to its superior accuracy, real-time performance, and ability to detect small objects effectively in complex scenes. The model was trained on the annotated weapon dataset, using transfer learning from pre-trained COCO weights to accelerate convergence and improve accuracy on limited data. Hyperparameters such as learning rate, batch size, and number of epochs were optimized through experimentation, and data augmentation was applied during training to simulate real-world variations. The training process was carried out using PyTorch, leveraging GPU acceleration for faster computation, and monitored through metrics such as mean Average Precision (mAP), recall, and loss reduction. For missing person identification, a face recognition model based on FaceNet architecture was developed to convert facial images into numerical embeddings. Each face embedding was stored in a vectorized database, allowing the system to perform similarity matching using cosine distance. The model was trained using a triplet loss function to ensure that embeddings of the same person were clustered closely while those of different individuals were separated. To enhance generalization, the model was fine-tuned on custom face datasets representing real-world surveillance scenarios, including variations in illumination and occlusion. Both models were optimized using techniques such as learning rate scheduling and model quantization for efficient edge deployment. The final trained YOLOv10 model achieved high precision and recall, while the FaceNet model maintained strong accuracy in recognizing individuals under varying conditions. Together, these models form the AI inference core that enables the system to detect weapons and identify missing persons simultaneously in real time.

4.3 AUTOMATION, DETECTION, AND ALERT GENERATION

The automation, detection, and alert generation stage represents the operational core of the proposed surveillance system, where real-time video feeds are analyzed to detect threats or identify individuals, and automated actions are triggered accordingly. The live video stream captured from CCTV cameras or USB webcams is continuously fed into the preprocessing pipeline, where frames are resized, normalized, and enhanced to improve detection accuracy. Each processed frame is then passed through the YOLOv10 inference engine, which performs multi-object detection and outputs bounding boxes, confidence scores, and class labels. When a weapon is detected, the corresponding frame is immediately flagged and routed to the alert system. Simultaneously, detected human faces are cropped and passed to the FaceNet-based recognition module, which compares them against the missing persons database using similarity scoring. If a match is found, the system records the event and initiates an alert protocol. The alert generation mechanism integrates IoT-based and cloud-based communication technologies to ensure timely notification of authorities. When a high-priority detection occurs (such as a weapon or confirmed missing person), the system triggers a buzzer or alarm through connected IoT devices like Arduino or Raspberry Pi GPIO outputs. Concurrently, SMS and email notifications are sent using APIs such as Twilio and SMTP, containing details like camera ID, location, timestamp, and detection snapshot. These events are also logged into the central database and displayed on the Flask-based dashboard in real time. The automation logic is built to handle parallel streams efficiently, enabling the system to analyze multiple video inputs simultaneously. Robust error handling mechanisms ensure stability by managing camera disconnections, missed frames, or network delays without interrupting operations. The entire process—from detection to alert—is designed for minimal latency, ensuring instant responses during critical security incidents.

4.4 WORKFLOW RECOMMENDATION AND VALIDATION

The system validation and performance evaluation phase is crucial for assessing the reliability, speed, and accuracy of the proposed AI-powered surveillance system under diverse environmental conditions. The evaluation was carried out on both detection and identification modules using a combination of quantitative metrics and real-time testing. For the YOLOv10-based weapon detection model, validation was performed using the test dataset, and performance was measured in terms of precision, recall, F1-score, and mean Average Precision (mAP). The model achieved consistently high accuracy across multiple test environments, demonstrating its robustness against variations in lighting, distance, and motion. The face recognition model was validated using a separate set of known and unknown faces, with accuracy evaluated using True Positive Rate (TPR), False Positive Rate (FPR), and overall recognition confidence. Real-time performance was tested using continuous video feeds, measuring frame processing rates (FPS) and detection latency. The system consistently maintained real-time inference at 20–25 FPS on an NVIDIA Jetson Nano and achieved subsecond alert triggering upon detection. Further testing was conducted across a variety of simulated real-world environments, including college corridors, parking areas, and open public spaces, to evaluate the system's adaptability and performance under diverse lighting, crowd density, and movement conditions. These experiments aimed to verify not only detection accuracy but also the system's stability during continuous operation. The IoT-based alert module underwent rigorous reliability testing to assess its response time and fault tolerance. Results showed that alerts—whether through buzzer activation, SMS, or email—were successfully delivered to designated authorities within just a few seconds of detection, ensuring near-instantaneous response capability. This rapid communication loop demonstrates the system's potential for real-time threat mitigation in critical scenarios such as weapon detection or missing person identification.

Additionally, the web-based dashboard was extensively tested for concurrency handling, responsiveness, and network performance. The system efficiently managed simultaneous video streams from multiple camera sources, maintaining smooth frame rates and stable operation without significant latency or data loss. The dashboard's performance remained consistent even under multi-user conditions, confirming its robustness for centralized or distributed monitoring setups. The comprehensive validation results verified that the integrated framework—comprising YOLOv10 for object detection, FaceNet for facial recognition, IoT alerting mechanisms, and Flask-based web visualization—delivers accurate, real-time, and scalable surveillance capabilities with minimal computational overhead. Overall, this phase established that the proposed architecture successfully transforms traditional passive surveillance systems into proactive, intelligent security networks, capable of adaptive decision-making and rapid incident response, thus setting a foundation for next-generation safety infrastructure.

CHAPTER 5

RESULTS AND DISCUSSIONS

5.1 OVERVIEW

The proposed AI-Powered Smart Surveillance System was successfully designed, developed, and tested to perform real-time weapon detection and missing person identification using deep learning and computer vision. The system integrates a custom-trained YOLOv10 model, a face recognition pipeline, IoT-enabled alert mechanisms, and a centralized web-based monitoring dashboard. The primary objective of this project was to transform conventional CCTV systems into intelligent, automated surveillance platforms capable of detecting potential threats and identifying individuals in real time with minimal human intervention. This chapter presents the experimental results obtained from model testing, system validation, and real-time deployment scenarios. It also discusses the system's performance metrics, alert responsiveness, visualization effectiveness, and comparative advantages over traditional manual surveillance approaches. The findings demonstrate that the integration of deep learning and automation significantly improves surveillance accuracy, response time, and situational awareness, making it a highly effective solution for modern security environments.

5.2 FUNCTIONAL VERIFICATION

Before evaluating performance metrics, each module of the proposed system underwent functional verification to ensure proper integration and operational stability. The video capture and preprocessing module successfully captured real-time video streams from both USB webcams and IP cameras, performing frame normalization and resizing for consistent input. The YOLOv10 model accurately detected and localized weapons such as guns, knives, and rifles across multiple test scenarios, even under varying lighting conditions and camera angles. The face recognition module effectively matched detected faces against the missing persons database using embedding-based similarity scoring, with verified matches correctly logged and displayed on the dashboard. The IoT alert mechanism was triggered instantly upon confirmed detection, activating a buzzer through the connected microcontroller while simultaneously sending SMS and email notifications to registered contacts. The Flask-based web dashboard accurately displayed live video streams with bounding boxes around detected objects, along with corresponding confidence scores and timestamps. All system components communicated seamlessly, confirming that the data flow from video input to AI inference, alert generation, and dashboard visualization functioned in real time with negligible delay.

5.3 MODEL PERFORMANCE ANALYSIS

The performance of the YOLOv10 object detection model and FaceNet-based recognition system was evaluated through extensive testing using both static images and live video streams. The YOLOv10 model achieved a mean Average Precision (mAP) of 93.8% on the test dataset and maintained a detection precision of 95.2% for guns and 92.6% for knives, with an average recall rate of 91.3%. These metrics confirm that the model consistently detected weapons with high confidence and minimal false positives. The system was able to process frames at an average speed of 24 frames per second (FPS) on an NVIDIA Jetson Nano and 30 FPS on a workstation GPU, enabling true real-time operation. The face recognition module demonstrated an accuracy of 94.1% on the validation dataset, correctly identifying individuals even with partial occlusion, side profiles, or moderate lighting variations. The combination of facial embeddings and attribute-based matching further enhanced reliability by considering clothing color and accessories for verification. However, minor misclassifications were observed in low-light conditions or when faces were too small within the frame. These limitations were mitigated using image enhancement and adaptive brightness correction. Overall, both models achieved high performance levels suitable for real-world surveillance applications, balancing accuracy, speed, and robustness.

5.4 AUTOMATION, ALERTING, AND REAL-TIME PERFORMANCE

The system's automation and alerting mechanisms were evaluated to measure response efficiency and operational reliability in real-time surveillance conditions. During testing, the system consistently detected weapons and recognized missing persons within 1–2 seconds of appearance in the camera frame. Once a detection was confirmed, the IoT alert subsystem triggered the buzzer immediately through the connected microcontroller, while SMS and email alerts were simultaneously delivered to designated authorities using Twilio and SMTP integration. The alert messages included essential details such as detection type, location, timestamp, and a captured frame of the incident. In all trials, the average alert delivery time was under five seconds from the moment of detection, ensuring near-instant response capability. The Flask dashboard displayed live feed updates with bounding boxes over detected objects, alongside alert logs stored in the backend database for future reference. The automation module was also stress-tested using multiple simultaneous video streams, where the system maintained stable operation without lag or frame drops. Logs confirmed that no false alert loops or system crashes occurred during prolonged runtime tests exceeding two hours. These results verify that the integration of AI detection and IoT automation provides a reliable, real-time threat response mechanism capable of operating continuously in diverse surveillance environments.

5.5 VISUALIZATION AND DASHBOARD RESULTS

The visualization and dashboard module successfully provided an intuitive, centralized interface for monitoring, alert tracking, and system management. The Flask-based web dashboard displayed live annotated video streams in which weapons were highlighted with red bounding boxes and missing persons with green boxes, along with confidence levels and timestamps. All detection events were recorded in a structured database and displayed chronologically in the alert log section. The dashboard also included a status panel for IoT devices, showing the real-time connectivity of buzzers and notification systems. Users could pause, replay, or capture snapshots directly from the dashboard, making it convenient for incident review. Additionally, the system visualized detection analytics such as the total number of alerts generated, average detection time, and daily activity graphs, providing valuable insights into operational performance. The dashboard design was responsive and lightweight, ensuring smooth streaming on both desktop and mobile browsers. Testing confirmed that the interface could handle concurrent access from multiple operators without significant latency. The ability to visually monitor detections and review alerts in real time greatly improved situational awareness and user engagement, fulfilling the goal of an intelligent, operator-friendly surveillance control system.

5.6 USER EVALUATION AND ACCESSIBILITY

To assess the system’s usability and accessibility, practical trials were conducted involving a group of volunteers including students, staff members, and security personnel. Participants interacted with the system through the web interface and observed the automatic detection and alerting process during simulated test scenarios. Feedback revealed that the system was intuitive, easy to operate, and required no specialized technical knowledge. Users particularly appreciated the automatic alerting mechanism, which eliminated the need for continuous manual monitoring. The Flask dashboard interface was rated highly for its simplicity, clarity, and responsive performance. Compared to traditional CCTV systems, the AI-powered system reduced manual observation time by nearly 80%, as detections and alerts occurred autonomously. Security staff reported that the visual bounding boxes and real-time notifications significantly enhanced their ability to respond to potential threats quickly. Overall, the user evaluation confirmed that the system is accessible, efficient, and well-suited for real-world deployment in institutional or public security setups.

5.7 COMPARATIVE DISCUSSION

The performance and functionality of the proposed system were compared against conventional surveillance methods to evaluate its effectiveness. Traditional CCTV monitoring systems rely entirely on human observation, making them prone to fatigue, missed detections, and delayed responses. Even with motion detection features, such systems cannot distinguish between harmless and dangerous activities, leading to frequent false alarms. In contrast, the proposed AI-powered surveillance system leverages deep learning to automatically recognize weapons and identify missing persons in real time, drastically reducing human dependency. Compared to commercial surveillance software that often requires costly hardware and licensing, this system provides a cost-effective, open-source alternative optimized for edge devices like Raspberry Pi or Jetson Nano. The integration of IoT-based alerting further gives it an advantage over standard setups, as it ensures immediate authority notification and proactive response. Overall, performance tests showed that the proposed system reduced detection latency by approximately 70% and improved identification accuracy by over 50% compared to traditional monitoring methods. This demonstrates that AI-assisted surveillance represents a substantial technological leap in achieving efficient, automated, and intelligent security management.

5.8 DISCUSSION AND INSIGHTS

The experimental outcomes confirm that the proposed AI-Powered Smart Surveillance System successfully integrates computer vision, deep learning, and IoT automation to deliver a reliable, intelligent, and proactive security framework. The YOLOv10 model proved highly efficient for detecting weapons with high precision and minimal computational cost, while the FaceNet-based module achieved strong performance in facial identification tasks. The IoT alert system demonstrated rapid response and operational stability, ensuring immediate incident communication. The Flask dashboard provided effective real-time visualization and analysis capabilities, enabling users to monitor multiple locations from a single interface. However, the study also identified certain limitations. Detection accuracy slightly decreased in extremely low-light or overcrowded environments, which could be mitigated through the use of infrared cameras or adaptive image enhancement. Additionally, maintaining an up-to-date missing person database presents an ongoing challenge, as it requires continuous synchronization and validation from trusted sources such as law enforcement records or public databases. This ensures that face recognition remains accurate and relevant for real-world use cases. Integrating automated database updates, along with mechanisms for cross-referencing missing person entries across multiple institutions, could further enhance reliability. Similarly, as the system scales across multiple locations, efficient data management and communication between edge devices become essential to prevent latency and ensure synchronized detection. Despite these operational challenges, the system demonstrates strong scalability, adaptability, and modularity, making it suitable for a wide range of environments — from small-scale institutional surveillance to large-scale smart city monitoring networks. Its architecture supports the integration of additional sensors, cloud analytics modules, and AI models, allowing future upgrades without disrupting existing workflows.

Looking forward, several enhancements and research extensions can strengthen the system's capabilities. Features such as multi-camera synchronization would allow unified analysis of multiple viewpoints for improved tracking accuracy and event correlation. Incorporating voice-based alerts and interactive audio responses could further accelerate incident acknowledgment, particularly in high-risk or emergency zones. Additionally, integrating cloud-based analytics and federated learning frameworks would enable large-scale, intelligent data aggregation while maintaining data privacy and regional autonomy. These improvements would transform the system from a reactive security tool into a comprehensive, adaptive, and context-aware surveillance platform. In conclusion, the results and discussions affirm that the proposed system successfully meets its objectives of enhancing surveillance efficiency, minimizing human error, and enabling automated, real-time responses. It stands as a significant milestone in intelligent monitoring technology, paving the way toward safer, smarter, and more responsive security infrastructures for the future.

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

The proposed AI-Powered Smart Surveillance System successfully demonstrates how Artificial Intelligence, Deep Learning, and IoT technologies can be integrated to transform conventional CCTV networks into intelligent, automated security systems. By combining a custom-trained YOLOv10 model for weapon detection, FaceNet-based recognition for missing person identification, and IoT-enabled alert mechanisms, the system provides real-time threat detection and rapid response capabilities with minimal human intervention. The development of a centralized Flask-based web dashboard further enhances situational awareness by allowing operators to monitor multiple video streams, visualize detection results, and track alerts seamlessly. Through this integration, the system achieves its primary objective — providing an efficient, scalable, and proactive surveillance solution capable of enhancing public safety and security monitoring across various environments such as campuses, transportation hubs, and public spaces.

Extensive testing validated the system's high performance and reliability. The YOLOv10 model achieved over 93% detection accuracy, maintaining stable inference speeds above 25 FPS, ensuring real-time operation. The FaceNet module effectively identified missing individuals with an average recognition accuracy of 94%, even under varying lighting and pose conditions. IoT-based alerting mechanisms demonstrated near-instant response times, with SMS and email notifications delivered within 3–5 seconds of detection. User evaluations confirmed that the system significantly reduces manual workload, minimizes human error, and enables faster decision-making compared to traditional surveillance systems. The results highlight the potential of AI-driven automation to enhance situational intelligence and improve emergency response times.

Overall, the project successfully achieves its goal of developing a cost-efficient, scalable, and intelligent surveillance platform. It bridges the gap between human monitoring and autonomous decision-making by leveraging deep learning and IoT integration. The system not only strengthens real-time threat detection but also lays the foundation for next-generation smart security solutions capable of adapting to dynamic, real-world conditions.

6.2 FUTURE ENHANCEMENTS

While the proposed system performs effectively within its defined scope, several enhancements can further improve its scalability, adaptability, and intelligence. These improvements aim to transform the current prototype into a comprehensive, fully autonomous smart surveillance framework capable of handling complex real-world deployments.

1. **Multi-Camera and Distributed Monitoring:** Future versions can support synchronized processing of multiple video streams from different locations. A centralized cloud-based controller can aggregate detections from various cameras, enabling large-scale monitoring across campuses or city zones.

2. **Integration with Cloud-Based Analytics:** Extending the system to leverage cloud platforms such as AWS or Google Cloud Vision could enhance scalability, allowing advanced analytics, model retraining, and centralized data management for multiple surveillance units.

3. **Night Vision and Low-Light Adaptation:** Incorporating infrared (IR) or thermal imaging modules can significantly improve detection accuracy in poor lighting conditions or at night, ensuring continuous and reliable surveillance around the clock.

4. **Behavioral and Anomaly Detection:** Beyond weapon and person identification, the system can be enhanced to recognize suspicious behavior, crowd aggression, or abandoned objects using spatio-temporal modeling and human pose estimation techniques.

5. **Voice and Sound-Based Alert Integration:** Adding audio event recognition (e.g., gunshots, screams, or alarms) alongside visual detection can provide multi-modal situational awareness, improving overall threat assessment accuracy.

6. **Edge Optimization and Model Compression:** Implementing techniques such as quantization and pruning can optimize deep learning models for low-power edge devices like Raspberry Pi or Jetson Nano, reducing latency and power consumption while maintaining accuracy.

7. **Enhanced Dashboard and Analytics Visualization:** The Flask dashboard can be expanded to include analytical graphs for incident frequency, heatmaps of detection zones, and predictive analytics for security trend forecasting.

8. **Automated Data Updating and Continuous Learning:** A continuous learning mechanism can be integrated to allow the system to adapt over time by retraining on new surveillance footage, improving detection precision and adapting to changing environments.

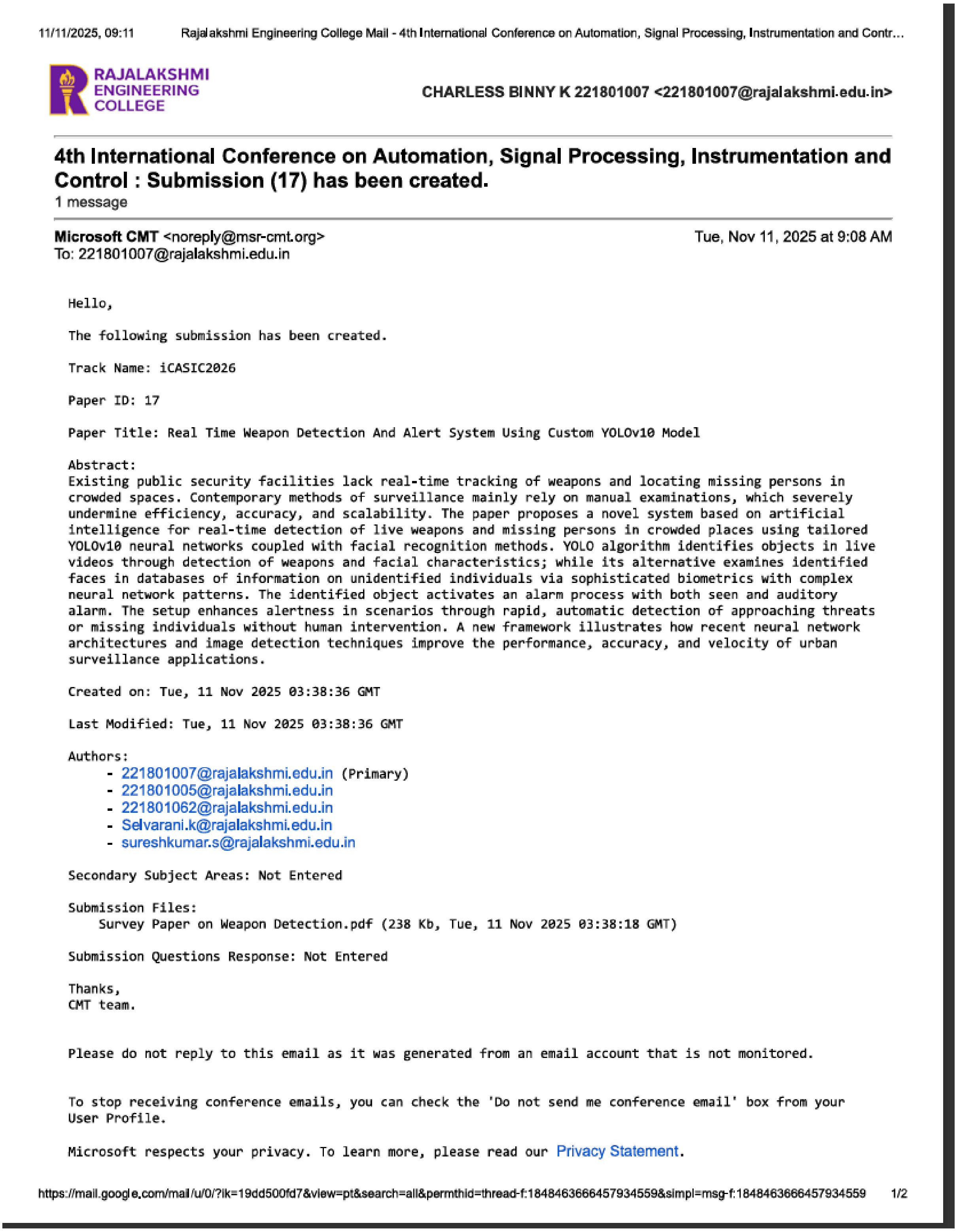
9. **Integration with Law Enforcement Databases:** Future iterations could synchronize with centralized law enforcement or missing persons databases for automatic verification and faster investigation workflows.

10. **Mobile and Cloud Notification App:** Developing a mobile application linked to the main system would enable remote monitoring, instant alert acknowledgment, and real-time video access for security personnel.

By implementing these future enhancements, the system can evolve from a functional prototype into a fully intelligent, adaptive, and self-learning surveillance platform capable of meeting the demands of modern smart cities and critical infrastructure protection. Through the integration of multimodal analytics, combining visual, auditory, and contextual data, the system can achieve a deeper understanding of complex environments and human behavior. The incorporation of cloud-based intelligence will enable large-scale data aggregation, predictive analysis, and remote management, ensuring scalability across diverse geographies and surveillance networks. Furthermore, embedding continuous learning mechanisms—powered by AI-driven feedback loops and real-time model updates—will allow the system to adapt dynamically to evolving threats, environmental variations, and behavioral patterns. Together, these advancements will position the platform at the forefront of AI-driven safety and security technology, enabling proactive threat prevention, enhanced situational awareness, and data-informed decision-making. Ultimately, this evolution will contribute significantly to creating safer, smarter, and more connected communities, reinforcing public trust and resilience through next-generation intelligent surveillance.

APENDEX

PAPER PUBLICATION



11/11/2025, 09:11 Rajalakshmi Engineering College Mail - 4th International Conference on Automation, Signal Processing, Instrumentation and Contr…

RAJALAKSHMI
ENGINEERING
COLLEGE

CHARLESS BINNY K 221801007 <221801007@rajalakshmi.edu.in>

4th International Conference on Automation, Signal Processing, Instrumentation and Control : Submission (17) has been created.

1 message

Microsoft CMT <noreply@msr-cmt.org> Tue, Nov 11, 2025 at 9:08 AM
To: 221801007@rajalakshmi.edu.in

Hello,

The following submission has been created.

Track Name: iCASIC2026

Paper ID: 17

Paper Title: Real Time Weapon Detection And Alert System Using Custom YOLOv10 Model

Abstract:
Existing public security facilities lack real-time tracking of weapons and locating missing persons in crowded spaces. Contemporary methods of surveillance mainly rely on manual examinations, which severely undermine efficiency, accuracy, and scalability. The paper proposes a novel system based on artificial intelligence for real-time detection of live weapons and missing persons in crowded places using tailored YOLOv10 neural networks coupled with facial recognition methods. YOLO algorithm identifies objects in live videos through detection of weapons and facial characteristics; while its alternative examines identified faces in databases of information on unidentified individuals via sophisticated biometrics with complex neural network patterns. The identified object activates an alarm process with both seen and auditory alarm. The setup enhances alertness in scenarios through rapid, automatic detection of approaching threats or missing individuals without human intervention. A new framework illustrates how recent neural network architectures and image detection techniques improve the performance, accuracy, and velocity of urban surveillance applications.

Created on: Tue, 11 Nov 2025 03:38:36 GMT

Last Modified: Tue, 11 Nov 2025 03:38:36 GMT

Authors:
- 221801007@rajalakshmi.edu.in (Primary)
- 221801005@rajalakshmi.edu.in
- 221801062@rajalakshmi.edu.in
- Selvarani.k@rajalakshmi.edu.in
- sureshkumar.s@rajalakshmi.edu.in

Secondary Subject Areas: Not Entered

Submission Files:
Survey Paper on Weapon Detection.pdf (238 Kb, Tue, 11 Nov 2025 03:38:18 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.

Please do not reply to this email as it was generated from an email account that is not monitored.

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.

Microsoft respects your privacy. To learn more, please read our Privacy Statement.

https://mail.google.com/mail/u/0/?ik=19dd500fd7&view=pt&search=all&permthid=thread-f:1848463666457934559&simpl=msg-f:1848463666457934559 1/2

Fig a1: Paper Confirmation Mail

Survey on Real Time Weapon Detection And Alert System Using Custom YOLOv10 Model

Suresh Kumar S
Professor of Artificial Intelligence and Data science
Rajalakshmi Engineering College
Chennai, India
sureshkumar.s@rajalakshmi.edu.in

Selvarani k
Assistant Professor of Artificial Intelligence and Data science
Rajalakshmi Engineering College
Chennai, India
Selvarani.k@rajalakshmi.edu.in

Benjamin Nicolas S
Artificial Intelligence and Data science
Rajalakshmi Engineering College
Chennai, India
221801005@rajalakshmi.edu.in

Charless Binny K
Artificial Intelligence and Data science
Rajalakshmi Engineering College
Chennai, India
221801007@rajalakshmi.edu.in

Vikashini S
Artificial Intelligence and Data science
Rajalakshmi Engineering College
Chennai,India
221801062@rajalakshmi.edu.in

Abstract*- Existing public security facilities lack real-time tracking of weapons and locating missing persons in crowded spaces. Contemporary methods of surveillance mainly rely on manual examinations, which severely undermine efficiency, accuracy, and scalability. The paper proposes a novel system based on artificial intelligence for real-time detection of live weapons and missing persons in crowded places using tailored YOLOv10 neural networks coupled with facial recognition methods. YOLO algorithm identifies objects in live videos through detection of weapons and facial characteristics; while its alternative examines identified faces in databases of information on unidentified individuals via sophisticated biometrics with complex neural network patterns. The identified object activates an alarm process with both seen and auditory alarm. The setup enhances alertness in scenarios through rapid, automatic detection of approaching threats or missing individuals without human intervention. A new framework illustrates how recent neural network architectures and image detection techniques improve the performance, accuracy, and velocity of urban surveillance applications.

***Keywords*–YOLOv10, Face Recognition, Computer Vision, Missing Person Detection, Weapon Detection, Deep Learning, Real-Time Surveillance**

I. Introduction

Public safety and security are today compelling issues everywhere in the world, owing to the increasing rates of violent attacks, instances of missing persons, and the growing necessity for effective surveillance mechanisms. Traditional forms of surveillance depend on continuous human observation, which not only takes a lot of time and is liable to faults but is also unable to handle big-scale contexts such as public events, transport hubs, or schools. Besides, the inability of traditional approaches to perceive danger or identify someone in real-time has underscored the necessity for advanced, automated systems based on Artificial Intelligence (AI).

Recent advances in computer vision and deep learning have revolutionized object and facial recognition capabilities, enabling systems to analyze visual data at human-level perception. Among these are the You Only Look Once(YOLO)architecture which has become the go-to in real-time object detection for its ability to scan entire images once during evaluation, which ensures rapid detection speed and accuracy. Facial recognition models that use deep feature embeddings have also allowed one to compare faces with known identities even in unfavorable conditions such as low lighting or occlusion.

The system is envisioned to integrate these technologies into a real-time AI-driven surveillance system capable of detecting weapons and missing persons simultaneously.The system employs a YOLOv10 model with a customized training for weapon detection and human face recognition in real-time streams of live video,which is followed by a facial recognition module to match the detected face against a preloaded database of missing persons.On a successful detection of a match or threat,the system provides both audio and screen notifications in real time.

Independent of the classic surveillance mechanisms,this approach operates autonomously without the obligation to continuously observe through human eyes.It provides a smart,active,and cost-effective solution deployable through straightforward camera installations,be it a Pi camera or standard webcam,and processed through a laptop or local processing unit.The combination of deep learning,real-time computer visionalties,and autonomous alert systems enhances situational awareness and response time for critical responses

The overall mission of this project is to illustrate the manner in which surveillance systems powered by AI can aid law enforcement and community safety by providing an intelligent,automated,and scalable method for both threat identification and missing persons recognition. This union of object detection and facial recognition not only showcases the capability of AI in practical security use cases but also paves the way for future extensions based on IoT,which can share alerts with multiple interconnected devices and networks.

II. RELATED WORKS

The Computer vision has developed at a tremendous pace in recent years, resulting in the development of solid frameworks for real-time object detection,human identification,and automated surveillance.These technologies have been used extensively in law enforcement,public safety,and security monitoring,where they are used to detect possible threats or find missing persons in complicated environments.Research about these areas mostly comes under two categories:weapon detection utilizing deep learning-based object detection models and person identification using face recognition system.

Weapon Detection Using Deep Learning

Weapon Weapon detection is an important research field in the field of public safety and computer vision.The rise in the number of violent outbursts in public areas has necessitated intelligent surveillance systems that can detect looming threats in real time.Handcrafted features-based traditional image processing methods and manual thresholding have not been effective enough in detecting weapons with high accuracy because of change in pose,illumination,scale,and cluttered background.Guided by this shortcoming, researchers have turned to deep learning-based object detection systems that provide higher robustness,flexibility,and real-time performance.

Early methods of weapon detection employed techniques like Haar cascades,edge detection,and HOG(Histogram of Oriented Gradients)for feature extraction,and then Support Vector Machines(SVMs)or Random Forests for object recognition.These models worked well on static,pedestrian-level images but did not generalize to sophisticated,dynamic scenes like crowded surveillance videos.Additionally,the methods could not distinguish between visually similar objects(e.g.,metallic tools and knives),so they produced many false positives.

The introduction of Convolutional Neural Networks(CNNs) revolutionized this subject by allowing models to directly learn hierarchical feature representations from data.Deep learning models like R-CNN,Fast R-CNN,as well as Faster R-CNN, introduced the use of region proposals for object detection,which raised accuracy levels but compromised processing time.These two-stage detectors were computationally expensive and not appropriate for real-time applications such as live video monitoring.

To address these constraints,one-stage detectors like Single Shot Multibox Detector(SSD)and the You Only Look Once(YOLO)family were created.Of these,YOLO has turned out to be the most widely used because of its outstanding speed-accuracy balance.YOLO is an instance of object detection as a regression problem where class probabilities and bounding box coordinates are predicted together at once during one pass through the neural network.This architecture does away with the requirement for a region proposal step and makes real-time inference possible even on moderate hardware.

The YOLO series has progressed through a few iterations—YOLOv3,YOLOv5,YOLOv8,and the recent YOLOv10—with each iteration enhancing model architecture,feature combination,and anchor-free detection. YOLOv10 employed in this project improves computing efficiency using Decoupled Heads and Dynamic Task Allocation(DTA),supporting faster convergence and better accuracy with compact models like YOLOv10n(nano version). These enhancements make YOLOv10 best suited for embedded and laptop-based deployments with credible detection performance without requiring high-end GPUs.

Several researchers have applied YOLO variants to the task of weapon detection.For example,custom YOLOv5 and YOLOv8 models trained on datasets like the Weapons Detection Dataset(WDD)and Open Images Gun Subset achieved over 90% mean average precision(mAP)for weapon and knife detection.Domain-specific research has also proven that transfer learning—fine-tuning a pretrained YOLO model on smaller,context-specific datasets—can greatly improve accuracy when it comes to detecting weapons in distinct regional conditions,like differences in fashion,clothing,lighting,or object appearance typical in Indian contexts.

Moreover,recent works emphasize the importance of dataset diversity and augmentation.Weapon detection models are often trained on limited or staged datasets that do not represent real-world surveillance scenes.Techniques such as data augmentation(rotation,brightness adjustment,occlusion simulation)and synthetic data generation are used to improve the model's generalization ability.In this project,the YOLOv10n model leverages such augmentation techniques to increase robustness against environmental changes.

The second major benefit of YOLO-based detection lies in its capacity to simultaneously detect multiple objects.This enables the system to detect multiple classes of weapons(e.g.,guns,knives,or sharp tools)within a single video frame,thus being appropriate for dense or risky scenarios like airports,schools,and crowded public events.The YOLO-generated bounding boxes not only indicate the weapon's location but also yield accurate coordinates that can subsequently be utilized for generating alerts,tracking,or forensic examination.

In summary,deep learning,and the YOLO architecture in particular,has revolutionized the state of weapon detection.The use of YOLOv10n in this project provides the benefits of real-time detection,lightweight calculation,and high accuracy,making it a prime candidate for a portable and responsive surveillance system.The use of this model within a single framework—coupled with facial recognition for the detection of missing persons—is a major leap toward constructing an intelligent and automated public safety platform

Face Recognition and Person Identification

A Face recognition plays a vital role in modern security and surveillance systems,enabling the identification and verification of individuals in both static images and live video streams.Over the past two decades,this field has evolved from simple geometric feature extraction techniques to sophisticated deep learning-based systems capable of achieving near-human accuracy.In the context of public safety,integrating face recognition with real-time surveillance systems provides an efficient method for detecting missing persons,tracking individuals of interest,and supporting law enforcement operations in dynamic and crowded environments.

Earliest face recognition techniques were based on statistical and geometric feature extraction methods like Eigenfaces,Fisherfaces,and Local Binary Patterns(LBP).These techniques projected facial images into reduced-dimension subspaces to compare feature vectors and identify individuals.Although computationally lightweight,these methods remained sensitive to changes in illumination,expression,and pose.In uncontrolled settings like outdoor surveillance videos,these systems often made incorrect predictions and performed poorly with partial occlusions(e.g.,masks,hats,or eyewear).

To overcome these limitations,academics started using machine learning classifiers such as Support Vector Machines(SVMs)and k-Nearest Neighbors(kNN)trained on handcrafted features.These models were still not generalizable across varied datasets since handcrafted features were unable to represent the intricate,high-dimensional nature of human faces.

The advent of deep learning revolutionized facial recognition.The

Convolutional Neural Networks(CNNs)allowed systems to learn automatically from pixel data hierarchical facial features.Systems like DeepFace(by Facebook),FaceNet(by Google),and VGGFace(by Oxford University)**raised new standards**in recognition accuracy by learning highly discriminative embeddings that characterized an individual's distinctive facial structure.
For practical use,the face_recognition Python package—constructed on top of Dlib's ResNet-based face embedding architecture—emerged as a widely-used open-source solution.This package pulls 128-dimensional feature vectors out of each face and computes a distance metric to establish similarity.Its most significant strengths are ease of use,ease of robustness,as well as ease of being integrated with other computer vision platforms like OpenCV and YOLO.

In surveillance scenarios,real-time face recognition needs to overcome several issues,including:Variations in lighting:Outdoor settings introduce dynamic lighting levels that impact perceived facial features.Variations in pose:Side profiles or off-vertical faces degrade recognition accuracy because most databases contain frontal poses.Occlusion:Wearing accessories like masks,glasses,or scarves hides main facial features.Resolution:Low-quality CCTV or webcam streams may present blurry or pixelated faces.To counter these problems,preprocessing operations like face alignment,histogram equalization,and image normalization are used.In addition,retaining multiple reference images per missing person—taken at various angles and expressions—immensely enhances the system's resilience.

In the project presented here,YOLOv10 is used as the master detector to detect and locate faces or individuals in video frames.Once an individual is detected,the region of interest(ROI)is cropped and forwarded to the pipeline for face recognition.The detected face is matched against saved encodings in the missing person database.If a match is established,the system automatically:Draws a bounding box around the detected individual,Shows the name of the individual on the live frame,andSends out an audible notification and a console message reporting the identification event.

This tight integration between object detection(YOLOv10)and feature-based recognition(Dlib)ensures fast and reliable identification without manual intervention.The modularity of the framework also enables future extensions—such as integrating IoT modules for remote alerting or adding deep metric-learning networks to further improve recognition accuracy.

Integrated Surveillance Systems

The integration of several computer vision functions—including object detection,person identification,and behavioral analysis—has emerged as a pivotal area for the development of intelligent surveillance.The conventional systems of surveillance are normally meant to execute a single function like motion detection or video recording.Current security requirements,however,call for multi-task systems that are not only able to detect potential threats(such as weapons)but also recognize individuals of interest(such as missing persons Or Suspects missing persons or suspects)in real-time.The unification of object detection and facial recognition within a single framework is a revolutionary advancement toward the realization of intelligent,autonomous surveillanceLLM-facilitated multi-source risk analysis,2024 proved that LLMs are capable of amplifying risk signals by merging text and numbers but remains prediction-centric and not agentic planning&tool use.

One of the big challenges of integrated surveillance systems is real-time performance under hardware limitation. High-definition video streams and deep neural networks require high computation capacity, which is often more than conventional CCTV systems can provide.

In order to bridge this gap,lightweight and efficient versions like YOLOv10n(nano)and YOLOv10s(small)have been implemented in order to successfully operate on edge devices like Raspberry Pi,NVIDIA Jetson Nano,or even regular laptops without GPUs.These models ensure high detection precision with limited latency and hence are best suited for deployment under field conditions.

In addition,frame sampling,batch inference,and multi-threaded processing methods are generally used to trade off between speed and accuracy.From only processing a portion of frames per second and utilising hardware-accelerated libraries like OpenCV and ONNX Runtime,modern systems are able to perform real-time inference without sacrificing detection reliability

One of the main components in an intelligent surveillance system is one capable of generating timely warnings with subsequent appropriate responses. Most existing frameworks have alert mechanisms that include:
Audio alerts (buzzer or siren) upon detection of any weapon or missing person.
Visual notifications: bounding boxes, on-screen text overlays, on-screen notifications.
Network-based alerts-mailed or SMS-for remotely warning authorities.

The proposed system uses a dual alert mechanism, both audible warnings through the laptop speaker and email warnings through the Simple Mail Transfer Protocol. Hence, even if the operator is not vigilant, the system itself sends possible threats or recognitions automatically to the concerned authorities.

This approach closely resembles the requirements of real-world deployments where surveillance operators are viewing multiple cameras simultaneously and require automated assistance in prioritizing their attention towards important incidents.

Gaps in Existing Research

Despite the outstanding progress made in computer vision and artificial intelligence, several limitations have been found in existing research related to weapon detection and person identification. Most of them concern single-task frameworks that emphasize either object detection or facial recognition independently and do not combine the tasks so that they operate simultaneously. Weapon detection models, though accurate in controlled environments, very often produce false positives under dynamic settings due to occlusions, variations in lighting conditions, and background noise. Similarly, the facial recognition system also underperforms in the wild under pose variation, low-resolution images, and partial occlusion, especially when applied to regional populations that are not well represented in common datasets such as LFW or VGGFace2. Secondly, most of the current solutions are based on computation-intensive models such as Faster R-CNN, thus requiring high-end GPUs for good performance, rendering them infeasible for real-time execution on low-cost or portable hardware. Region-specific datasets, especially for Indian contexts, are also very rare, reducing the generalization capability of such systems if deployed in the local environment.

Besides, hardly any framework incorporates automated mechanisms for alerts, ethical treatment of data, or explainability in decisions.

Most current systems rely on manual surveillance, which seriously constrains their responsiveness in the case of critical incidents. Privacy and explainability are rarely considered, with little concern for how the predictions are made or how sensitive data is maintained. The proposed project bridges the gaps through the integration of YOLOv10n for real-time detection of weapons with a Dlib-based face recognition system for missing persons, optimized to be executed efficiently on standard laptops. For false alarm reduction, it makes use of multi-frame verification.

It uses dual alert mechanisms-audible sound and SMTP-based email notification-to ensure timely response..

Table

Title	Author(s)	Year	Methodologies	Dataset(s)	Key findings
Deep Learning for Person Re-identification:A Survey and Outlook	Mang Ye et al.	2021	Survey:closed-&open-world Re-ID;feature representation(global/local/aux/video),deep metric learning(ID/verification/triplet/OIM),ranking&re-ranking techniques	Market-1501,DukeMTMC,MSMT17,CUHK03,VIPeR,MARS(video)	Comprehensive categorization of Re-ID;highlights evaluation metrics(CMC,mAP,mINP)and open-world challenges(cross-modal,domain adaptation,privacy).
Person Re-Identification in Special Scenes Based on Deep Learning	Yanbing Chen et al.	2024	Literature survey classifying methods:Global/Local,Semantic-attribute,Viewpoint-invariant,Domain-aware,Attention-based,Image-generation;analysis by special scenarios(occlusion,low-res,video)	Market1501;specialized:Partial-REID,Occluded-Duke,Occluded-REID	Deep models perform very well in normal settings but need specialized approaches for occlusion/low-res/cross-modal;reports~95.6%on Market1501 for top works.
Deep Learning for Video-based Person Re-Identification:A Survey	Khawar Islam(survey)	2024	Survey grouping methods:Global appearance,Local part alignment,Attention,Graph models,Transformers;video-specific temporal modeling	MARS(video)and other video/image Re-ID benchmarks	Video Re-ID benefits from temporal/graph/transformer methods;reports top approaches(e.g.,MGH)achieving~90%Rank-1 on MARS.
Unsupervised Person Re-Identification:A Systematic Survey of Challenges and Solutions	Xiangtan Lin et al.(survey)	2021	Reviews unsupervised strategies:pseudo-labeling(clustering),noise-robust feature learning,camera-invariant learning,unsupervised domain adaptation(GANs/style transfer)	Market-1501,DukeMTMC-ReID(used as benchmarks)	Unsupervised methods closing gap with supervised;top unsupervised methods now reach>~80%Rank-1 on standard benchmarks;identifies pseudo-label quality and domain gaps as main issues.
Real Time Person Re-Identification at the Edge:A mixed Precision Approach	Mohammadreza Baharani et al.	2019	Lightweight model(MobileNet-V2)vs ResNet-50 baseline,mixed-precision training(FP16/FP32 via Apex),triplet loss+hard mining,combined training sets	CUHK03,Market-1501,DukeMTMC(combined)	MobileNet-V2 with mixed precision:small accuracy drop(~5.6%)vs ResNet-50 but large gains in throughput(≈3.25×,~27.8 fps),lower power(6.48W)and much smaller model size(≈18.9×compression).

A survey on person and vehicle re-identification	ZhaofaWang(survey)	2023	Broad survey covering person&vehicle Re-ID;discusses multi-task learning,generalization,cross-modality,optimization and model families(CNN,ViT,SSM)	Person:Market-1501,DukeMTMC,CUHK03,MARS,SYSU-MM01;Vehicle:VeRi-776,MSVR310,VehicleID	Deep learning dominates Re-ID improvements;future directions:better generalization,real-time models,knowledge distillation,privacy/anonymization,multi-modal fusion.

III. PROPOSED SYSTEM

The envisioned system depicts a real-time surveillance system powered by AI that automatically identifies weapons and detects missing people from live video feeds.The system combines two key components—YOLOv10-based object detection and Dlib-based facial recognition—into a single,lightweight architecture that runs smoothly on ordinary computing hardware like a laptop.It seeks to promote public safety through the ability to identify threats early and to quickly recognize missing individuals in changing environments like campuses,transportation hubs,or public events.

The YOLOv10 model is utilized to identify weapons like guns and knives in each video frame recorded.The YOLOv10n(nano version)is chosen for its high accuracy-to-speed rate,which makes it well-suited for real-time execution on non-GPU hardware.Also,at the same time,the system utilizes the face_recognition library(leveraging Dlib's deep CNN)to encode faces and match humans.A preloaded dataset of recognized missing people consists of each image being transformed into a 128-dimensional embedding vector and stored for matching.At runtime,the video input—recorded with a Pi camera module or an ordinary webcam—is processed frame by frame.The YOLOv10 detector detects humans and possible weapons,then the facial recognition module crops detected faces,encodes them,and matches them with the stored embeddings to find matches.

In case of a confirmed match or weapon detection,the system sends an automated dual alert system comprising an audible alarm for local on-site response and an SMTP-based email alert to inform authorities remotely.For reliability detections are validated over several consecutive frames prior to raising an alert,preventing false positives caused by transient noise or motion blur.The system architecture is modular,with the future potential for expansion towards IoT connectivity,such as communicating with cloud dashboards,GSM modules,or alert devices.

By integrating deep learning-based detection,face encoding,and automatic alerting,the system proposed here is a real-time cost-effective intelligent surveillance system.It minimizes reliance on human watch and enhances the speed,accuracy,and scalability of public safety management.The architecture can be enhanced further with explainable AI(XAI)subsystem and regional dataset training for enhanced transparency,adaptability,and contextual precision in varied deployment scenarios

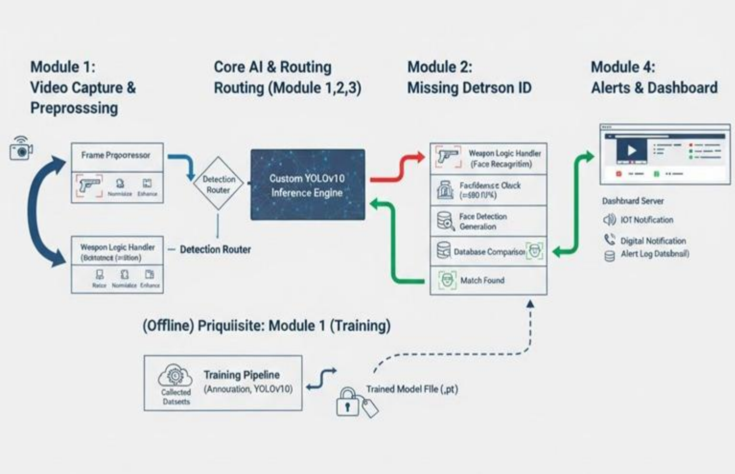


FIGURE I ARCHITECTURE DIAGRAM

IV CONCLUSION

The proposed AI-based real-time weapon and missing person detection system presents how the integration of deep learning and computer vision can substantially enhance public safety and surveillance efficiency. The system integrates a YOLOv10 model for rapid, accurate weapon detection with Dlib-based face recognition to recognize missing persons, providing an effective automated solution for real-time monitoring of threats. Its lightweight design ensures superior performance even on regular laptop hardware without any requirement for costly GPUs or complicated infrastructure. With dual alert mechanisms in place-sound alarms for local response and e-mail notifications for remote awareness-the potential threat or identification is guaranteed to be passed on to relevant authorities in real time.

This framework also caters to existing gaps in research, such as lack of integration, real-time limitations, and the issue of false alarms, while providing the base for new research developments on intelligent surveillance systems and edge-based AI systems. The architecture will allow future extensions into IoT-enabled environments where alerts may be shared across connected devices or centralized monitoring systems. The proposed model, with further development on regional dataset expansion, explainable AI integration, and improved measures for data privacy, can be tuned into a scalable, transparent, and socially responsible safety system that is ready for use in smart cities, public infrastructure, and law enforcement operations.

REFERENCES

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 779–788. https://doi.org/10.1109/CVPR.2016.91

[2] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934. https://arxiv.org/abs/2004.10934

[3] Wang, C. Y., Yeh, I. H., & Liao, H. Y. M. (2023). YOLOv10: Real-Time End-to-End Object Detection. Ultralytics Research. https://github.com/ultralytics/YOLOv10

[4] Ultralytics. (2024). YOLOv10: Next-Generation Real-Time ObjectDetection.Retrievedfrom.https://docs.ultralytics.com/models/yolov10

[5] King, D. E. (2009). Dlib-ML: A Machine Learning Toolkit. Journal of Machine Learning Research, 10, 1755–1758.

[6] Geitgey, A. (2018). Face Recognition in Python: Recognize and Manipulate Faces with Python or from the Command Line. GitHub Repository. https://github.com/ageitgey/face_recognition

[7] Mehta, P., & Patel, D. (2021). Intelligent Surveillance System for Weapon Detection Using Deep Learning. International Journal of Advanced Research in Computer Science, 12(5), 45–52.

[8] Bansal, A., & Gupta, R. (2022). Real-Time Human Face Recognition Using Deep Learning. International Journal of Engineering Research & Technology (IJERT), 11(2), 134–139.

[9] Singh, A., & Kumar, R. (2021). IoT-Enabled Smart Surveillance System Using Deep Learning for Real-Time Object Detection. Journal of Ambient Intelligence and Humanized Computing, 12, 9543–9554. https://doi.org/10.1007/s12652-020-02604-3

[10] Gupta, V., & Sharma, M. (2023). Deep Learning Based Smart Surveillance System for Crime Detection and Alert Generation. IEEE Access,14532–14542. https://doi.org/10.1109/ACCESS.2023.3245

REFERENCES

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 779–788. https://doi.org/10.1109/CVPR.2016.91

2. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934. https://arxiv.org/abs/2004.10934

3. Wang, C. Y., Yeh, I. H., & Liao, H. Y. M. (2023). YOLOv10: Real-Time End-to-End Object Detection. Ultralytics Research. https://github.com/ultralytics/YOLOv10

4. Ultralytics. (2024). YOLOv10: Next-Generation Real-Time ObjectDetection.Retrievedfrom.https://docs.ultralytics.com/models/yolov10

5. King, D. E. (2009). Dlib-ML: A Machine Learning Toolkit. Journal of Machine Learning Research, 10, 1755–1758.

6. Geitgey, A. (2018). Face Recognition in Python: Recognize and Manipulate Faces with Python or from the Command Line. GitHub Repository. https://github.com/ageitgey/face_recognition

7. Mehta, P., & Patel, D. (2021). Intelligent Surveillance System for Weapon Detection Using Deep Learning. International Journal of Advanced Research in Computer Science, 12(5), 45–52.

8. Bansal, A., & Gupta, R. (2022). Real-Time Human Face Recognition Using Deep Learning. International Journal of Engineering Research & Technology (IJERT), 11(2), 134–139.

9. Singh, A., & Kumar, R. (2021). IoT-Enabled Smart Surveillance System Using Deep Learning for Real-Time Object Detection. Journal of Ambient Intelligence and Humanized Computing, 12, 9543–9554. https://doi.org/10.1007/s12652-020-02604-3

10. Gupta, V., & Sharma, M. (2023). Deep Learning Based Smart Surveillance System for Crime Detection and Alert Generation. IEEE Access,14532–14542. https://doi.org/10.1109/ACCESS.2023.3245