GunShield AI: An Automated System for Gun Appearance using Human Pose

Santosh E,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
santoshe_cse@mitmysore.in

Bhavyashree H. L,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
bhavyashreehl_cse@mitmysore.in

Keerthi M,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
keeerthim296@gmail.com

Niranjan J,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
niranjanjagadeesh8@gmail.com

Krishna Yash Raj,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
krishnayashraj15@gmail.com

Vinay Kumar J D,
Department of Computer Science and Engineering,
Maharaja Institute of Technology Mysore
Affiliated to Visvesvaraya Technological University (VTU),
Belagavi, Karnataka, India
vinayjkuppe.jd@gmail.com

ARTICLE INFO

Keywords:
Real-Time Threat Detection
Gun Detection
Human Pose Estimation
TensorFlow
OpenCV
MediaPipe
Machine Learning
CNN (Convolutional Neural Network)

ABSTRACT

Ensuring public safety has become a major concern with the rise in firearm-related incidents. This paper presents a real-time gun and human pose threat detection system designed to identify potential threats from live or recorded video feeds. The proposed system uses **TensorFlow** and **OpenCV** for object detection and **MediaPipe** for analyzing human body posture to assess suspicious activities. When a possible threat is detected, the system triggers an automatic alert, captures evidence frames, and stores the event data for further analysis. The model is trained on a combination of firearm datasets and human pose images to improve accuracy in dynamic environments. Experimental results show that the system performs effectively in identifying both visible weapons and aggressive human poses under varying lighting and background conditions. This approach aims to support real-time surveillance and enhance security response through intelligent automation.

Introduction

Public safety has become a growing concern in modern society, particularly due to the increasing number of firearm-related incidents in public spaces such as schools, offices, and transportation facilities. Traditional CCTV-based surveillance systems depend heavily on continuous human supervision, which can often lead to fatigue, slower response times, or missed threats during critical moments. As a result, there is a pressing need for intelligent, automated systems that can detect potential dangers and assist security personnel in responding quickly and accurately.

In recent years, the fields of **computer vision** and **artificial intelligence (AI)** have made significant progress, allowing machines to interpret visual data almost as effectively as humans. By leveraging these advancements, it is now possible to create systems that not only identify objects like firearms but also analyze human body movements to assess potential aggression or abnormal behavior. This combination of **object detection** and **pose estimation** forms the foundation of this research.

The proposed **Real-Time Gun and Human Pose Threat Detection System** aims to integrate multiple AI techniques to ensure accurate and timely identification of security threats. The system utilizes **TensorFlow** for deep learning model implementation, **OpenCV** for frame processing and visualization, and **MediaPipe** for real-time human pose analysis. These frameworks are combined to analyze live video feeds or recorded footage, allowing the system to monitor multiple individuals simultaneously. Whenever a firearm or suspicious pose is detected, the system generates instant alerts, saves a screenshot of the event, and records the details in a log file for further investigation. Additionally, an **SMS alert feature** is included to notify security personnel instantly, improving the overall response time during emergencies.

Unlike conventional surveillance setups that only detect objects, this project also considers **human behavior patterns**, such as body posture and limb orientation, which are crucial indicators of possible violence. This approach significantly reduces false positives, ensuring that the alerts generated are meaningful and reliable. The inclusion of human pose analysis adds another dimension to threat recognition, enabling the system to distinguish between harmless activities and potential attacks more effectively.

The research contributes to the growing field of **intelligent surveillance** by demonstrating a cost-effective, modular, and real-time solution that can be deployed in various environments. The framework is flexible and can be further enhanced by integrating cloud-based storage, facial recognition, or advanced behavioral analytics. Ultimately, this system showcases how modern AI tools can be applied to strengthen public safety infrastructure and reduce the dependence on manual monitoring, leading to faster decision-making and more secure environments.

In summary, the proposed system serves as a practical step toward smarter and more responsive surveillance solutions. By combining deep learning with real-time video analysis, it not only detects weapons but also interprets human behavior to predict possible threats. This integration of technology enhances situational awareness, reduces human error, and ultimately contributes to building safer environments through intelligent automation.

Related Works

Over the past few years, several studies have focused on improving firearm detection and threat analysis using deep learning and computer vision. Traditional surveillance systems largely depend on detecting the visual appearance of weapons, which often leads to high false positives, especially under low lighting, occlusion, or complex backgrounds. To overcome these limitations, researchers have started integrating **human pose estimation** with **weapon appearance detection**, allowing systems to better interpret context and intent behind human actions.

Amulya Reddy Maligireddy et al. (2025) proposed a novel method that combines human pose estimation with weapon appearance detection to enhance the accuracy of firearm recognition in CCTV footage. Using **YOLOv8** for object detection and **MediaPipe** for pose tracking, their approach significantly reduced both false positives and false negatives. The dataset comprised over **9,500 images** collected from **IMFDB, Monash Guns, Roboflow**, and real-world CCTV data. Their system achieved outstanding results with **mAP50 of 0.98** and **mAP50–95 of 0.70**, demonstrating the effectiveness of combining pose and proximity information between guns and human keypoints. They also emphasized ethical deployment, privacy protection, and proposed future research on Transformer-based optimization.

Jesus Ruiz-Santaquiteria et al. (2023) advanced this field by integrating **CNNs** and **Transformer architectures** such as **Vision Transformer (ViT)** and **DeiT** for extracting visual features. Their handgun detection method utilized both visual cues and body pose data during the detection process itself rather than as a secondary step, which helped reduce false positives and negatives. The dataset used contained real, synthetic, and low-light CCTV images exceeding **9,500 samples**, leading to robust performance under challenging environmental conditions. Their best-performing model achieved an **Average Precision (AP) of 91.73%**, and they introduced a **False Positive Filtering module** for further robustness. The study also recommended incorporating temporal (video-based) data for real-time applications.

In another significant contribution, **Sriram C. S. and Dr. G. Prema Arokia Mary (2024)** developed a system combining **OpenPose**-based human pose estimation with **ResNet** and **EfficientNet** classifiers to recognize weapons and handheld objects in surveillance footage. Their work emphasized reducing false detections caused by everyday items such as wallets or smartphones. The proposed system achieved **96.39% classification accuracy**, outperforming EfficientNet, and demonstrated adaptability for public and high-security environments. Their future scope included real-time processing, multimodal data fusion, and scalability improvements.

Earlier works by **Ruiz-Santaquiteria et al. (2021)** focused on combining **human pose estimation** with **Darknet-53 CNN** for handgun recognition. They introduced two models: **HRC (Hand Region Classifier)** and **HRC+P (Hand Region Classifier + Pose information)**. Their system achieved an **Average Precision (AP) of 83.85%**, outperforming YOLOv3, particularly in poor lighting or distant camera views. They highlighted that pose information enhances context awareness, allowing the system to differentiate between threatening and non-threatening human actions.

Similarly, **Velasco-Mata et al. (2021)** used **OpenPose** to extract skeletal keypoints and fused them with **YOLOv3** handgun detector outputs. This combinational model merged pose-based heatmaps with detection features, resulting in a **17.5% improvement in Average Precision** compared to baseline models. Their work proved that integrating pose information reduces false detections in diverse scenarios, including YouTube and video-game-sourced datasets. Their future direction focused on improving real-time performance and refining feature integration for deployment in practical surveillance systems.

From this literature, it is evident that incorporating **human pose analysis** with **weapon detection** substantially improves recognition accuracy and contextual awareness. However, most prior works primarily concentrated on offline or experimental setups rather than fully functional **real-time surveillance systems**. Moreover, challenges such as dynamic lighting, fast movements, and multiple-person tracking remain partially unresolved. The proposed system in this research builds upon these studies by integrating **TensorFlow**, **OpenCV**, and **MediaPipe** to create a **real-time, automated threat detection framework** that not only identifies weapons but also interprets human behavior and instantly generates alerts for security response.

Methods

The proposed real-time gun and human pose threat detection system leverages computer vision and deep learning techniques to monitor video streams and identify potential security threats. The system processes video input frame by frame, ensuring near real-time detection while maintaining computational efficiency. Each frame is preprocessed through resizing to a uniform resolution and pixel normalization, which stabilizes the input for the detection models. During model training, data augmentation strategies including rotation, horizontal flipping, and brightness adjustments are employed to enhance the robustness of both gun detection and human pose estimation models.

Gun detection is performed using a deep learning-based object detection network, which outputs bounding boxes around firearms present in the frame. The detection confidence score, denoted as C_g, quantifies the likelihood of a firearm being present. Concurrently, human pose estimation is performed using a landmark-based approach, where keypoints corresponding to joints such as shoulders, elbows, and wrists are detected. These keypoints allow the system to analyze the spatial orientation of a person and assess whether the posture indicates a threatening action, such as aiming a firearm. A pose threat score, C_p, is assigned based on the alignment of the limbs with the detected firearm, capturing the degree of potential threat.

Figure 1 Human Pose Estimation

The threat assessment is modeled as a combination of gun detection confidence and pose threat score. A simplified equation representing this logic is given by:

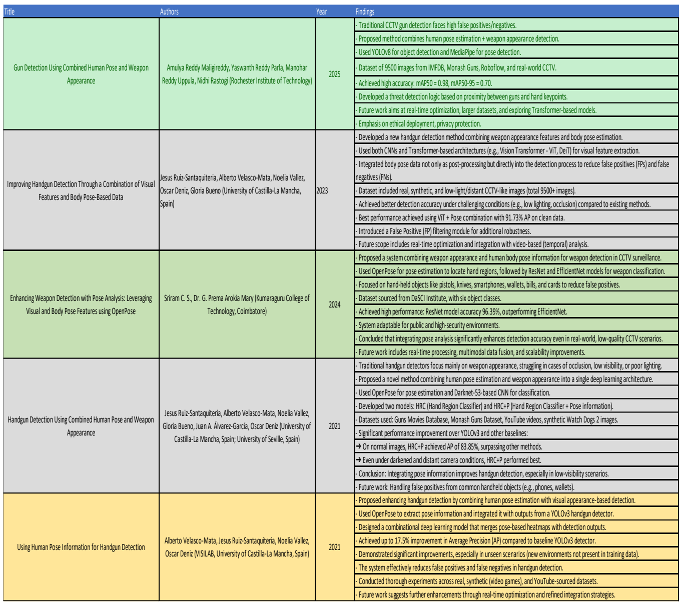
$$T = \alpha \cdot C_g + \beta \cdot C_p$$

where T is the overall threat score for a frame, C_g is the confidence of gun detection, C_p is the pose threat score, and α and β are weighting factors that balance the contributions of the two components. A frame is flagged as a threat if T exceeds a predefined threshold τ. This mathematical formulation allows the system to quantify threat levels in a consistent and interpretable manner, reducing false positives and enabling actionable alerts.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

Figure 2 pose estimation + weapon detection

Upon identification of a threat, the system captures a screenshot, logs metadata including timestamp and pose coordinates, and sends an immediate SMS alert to predefined subscribers using the Fast2SMS API. The complete workflow is implemented in Python, utilizing OpenCV for video processing, TensorFlow for model inference, MediaPipe for pose estimation, and Fast2SMS for notifications. The architecture follows a sequential pipeline: video input → frame extraction → gun detection → pose estimation → threat computation → alert generation. This integrated approach ensures high accuracy in real-time surveillance applications while providing a scalable framework for automated threat detection.



Title	Authors	Year	Findings
Gun Detection Using Combined Human Pose and Weapon Appearance	Amulya Reddy Maligireddy, Yaswanth Reddy Parla, Manohar Reddy Uppula, Nidhi Rastogi (Rochester Institute of Technology)	2025	- Traditional CCTV gun detection faces high false positives/negatives. - Proposed method combines human pose estimation + weapon appearance detection. - Used YOLOv8 for object detection and MediaPipe for pose detection. - Dataset of 9500 images from IMFDB, Monash Guns, Roboflow, and real-world CCTV. - Achieved high accuracy: mAP50 = 0.98, mAP50-95 = 0.70. - Developed a threat detection logic based on proximity between guns and hand keypoints. - Future work aims at real-time optimization, larger datasets, and exploring Transformer-based models. - Emphasis on ethical deployment, privacy protection.
Improving Handgun Detection Through a Combination of Visual Features and Body Pose-Based Data	Jesus Ruiz-Santaquiteria, Alberto Velasco-Mata, Noelia Vallez, Oscar Deniz, Gloria Bueno (University of Castilla-La Mancha, Spain)	2023	- Developed a new handgun detection method combining weapon appearance features and body pose estimation. - Used both CNNs and Transformer-based architectures (e.g., Vision Transformer - ViT, DeiT) for visual feature extraction. - Integrated body pose data not only as post-processing but directly into the detection process to reduce false positives (FPs) and false negatives (FNs). - Dataset included real, synthetic, and low-light/distant CCTV-like images (total 9500+ images). - Achieved better detection accuracy under challenging conditions (e.g., low lighting, occlusion) compared to existing methods. - Best performance achieved using ViT + Pose combination with 91.73% AP on clean data. - Introduced a False Positive (FP) filtering module for additional robustness. - Future scope includes real-time optimization and integration with video-based (temporal) analysis.
Enhancing Weapon Detection with Pose Analysis: Leveraging Visual and Body Pose Features using OpenPose	Sriram C. S., Dr. G. Prema Arokia Mary (Kumaraguru College of Technology, Coimbatore)	2024	- Proposed a system combining weapon appearance and human body pose information for weapon detection in CCTV surveillance. - Used OpenPose for pose estimation to locate hand regions, followed by ResNet and EfficientNet models for weapon classification. - Focused on hand-held objects like pistols, knives, smartphones, wallets, bills, and cards to reduce false positives. - Dataset sourced from DaSCI Institute, with six object classes. - Achieved high performance: ResNet model accuracy 96.39%, outperforming EfficientNet. - System adaptable for public and high-security environments. - Concluded that integrating pose analysis significantly enhances detection accuracy even in real-world, low-quality CCTV scenarios. - Future work includes real-time processing, multimodal data fusion, and scalability improvements.
Handgun Detection Using Combined Human Pose and Weapon Appearance	Jesus Ruiz-Santaquiteria, Alberto Velasco-Mata, Noelia Vallez, Gloria Bueno, Juan A. Álvarez-García, Oscar Deniz (University of Castilla-La Mancha, Spain; University of Seville, Spain)	2021	- Traditional handgun detectors focus mainly on weapon appearance, struggling in cases of occlusion, low visibility, or poor lighting. - Proposed a novel method combining human pose estimation and weapon appearance into a single deep learning architecture. - Used OpenPose for pose estimation and Darknet-53-based CNN for classification. - Developed two models: HRC (Hand Region Classifier) and HRC+P (Hand Region Classifier + Pose information). - Datasets used: Guns Movies Database, Monash Guns Dataset, YouTube videos, synthetic Watch Dogs 2 images. - Significant performance improvement over YOLOv3 and other baselines: ➔ On normal images, HRC+P achieved AP of 83.85%, surpassing other methods. ➔ Even under darkened and distant camera conditions, HRC+P performed best. - Conclusion: Integrating pose information improves handgun detection, especially in low-visibility scenarios. - Future work: Handling false positives from common handheld objects (e.g., phones, wallets).
Using Human Pose Information for Handgun Detection	Alberto Velasco-Mata, Jesus Ruiz-Santaquiteria, Noelia Vallez, Oscar Deniz (VISILAB, University of Castilla-La Mancha, Spain)	2021	- Proposed enhancing handgun detection by combining human pose estimation with visual appearance-based detection. - Used OpenPose to extract pose information and integrated it with outputs from a YOLOv3 handgun detector. - Designed a combinational deep learning model that merges pose-based heatmaps with detection outputs. - Achieved up to 17.5% improvement in Average Precision (AP) compared to baseline YOLOv3 detector. - Demonstrated significant improvements, especially in unseen scenarios (new environments not present in training data). - The system effectively reduces false positives and false negatives in handgun detection. - Conducted thorough experiments across real, synthetic (video games), and YouTube-sourced datasets. - Future work suggests further enhancements through real-time optimization and refined integration strategies.

Figure 3 Summary of existing research works on gun and human pose detection.(Source: Compiled from previous studies reviewed)

Data Description

The dataset used in this study consists of video sequences captured from publicly available sources such as YouTube and simulated CCTV recordings. These videos represent a diverse range of scenarios, including humans holding firearms, humans without firearms, and simulated threat situations. Each video was annotated manually to identify the presence of guns and to mark human keypoints for pose estimation. The dataset was split into training, validation, and testing sets to ensure unbiased evaluation of the detection system. The annotated dataset provides both bounding box labels for firearms and keypoint coordinates for human joints, enabling the integration of gun detection and pose-based threat assessment.

System Requirements

The system was implemented in Python and leverages both CPU and GPU resources to achieve near real-time performance. The hardware requirements include an NVIDIA GTX 1650 or RTX 2060 GPU for accelerated computation, while standard CPUs can handle smaller-scale processing. The software stack consists of OpenCV for video processing, TensorFlow for model inference, MediaPipe for human pose estimation, and Fast2SMS for sending SMS alerts. The operating system used was Windows 10, and Python 3.9 served as the programming environment. This setup ensures smooth operation of all modules including video capture, frame processing, gun detection, pose estimation, threat evaluation, and alert generation.

Preprocessing

Each frame of the input video is preprocessed to standardize the input for the detection and pose estimation modules. Frames are resized to 640×480 pixels to maintain consistency and normalized to a [0, 1] range to improve computational stability. The normalization of pixel intensity is mathematically represented as:

$$I_{\text{norm}} = \frac{I - I_{\min}}{I_{\max} - I_{\min}}$$

where I is the original pixel intensity, and $I_{\min}$ and $I_{\max}$ are the minimum and maximum pixel values in the frame. Data augmentation techniques, including horizontal flipping, rotation, and brightness adjustments, were applied to improve the system's robustness against varying lighting conditions and camera angles. These preprocessing steps ensure that both the gun detection and human pose modules perform reliably across diverse video inputs.

Gun Detection Module

The gun detection module is based on a convolutional neural network that identifies the presence of firearms in each video frame. The network consists of multiple convolutional layers for feature extraction, interleaved with pooling layers to reduce spatial dimensions while preserving important information. Fully connected layers follow the convolutional layers to consolidate features for classification. The output of the network is a binary prediction indicating whether a gun is present in the frame, and a confidence score quantifying the certainty of the detection. The network is optimized to achieve high accuracy while maintaining real-time processing capability.

CNN-Based Gun Detection

Gun detection is performed using a convolutional neural network (CNN) trained to classify frames as "gun present" or "gun absent." The CNN architecture consists of multiple convolutional layers, pooling layers, and fully connected layers optimized for feature extraction and real-time inference. The output layer generates a binary prediction indicating the presence of a firearm.

Layer Type	**Kernel/Size**	**Stride**	**Activation**	**Output Shape**
Input Layer	-	-	-	640×480×3
Conv2D	3×3	1	ReLU	640×480×32
MaxPooling2D	2×2	2	-	320×240×32
Conv2D	3×3	1	ReLU	320×240×64
MaxPooling2D	2×2	2	-	160×120×64
Fully Connected	-	-	ReLU	256
Output Layer	-	-	Sigmoid	1 (Gun/No Gun)

Table 1: CNN Architecture

Human Pose Estimation

Human pose estimation is performed using MediaPipe, which detects keypoints corresponding to major joints such as shoulders, elbows, and wrists. These keypoints allow the system to evaluate the spatial orientation and posture of individuals within the video frame. By analyzing the relative positions of the joints, the system can identify suspicious postures, such as raising or aiming a firearm. The pose estimation module outputs coordinates for each keypoint, which are subsequently used in the threat evaluation logic.

Threat Evaluation

Threat evaluation in the system is based on **rule-based logic** that combines outputs from the gun detection and human pose estimation modules. A frame is classified as a threat if a firearm is detected and the human pose indicates pointing or aiming. This rule can be expressed mathematically as:

$$\text{Alert} = \begin{cases} 1 & \text{if gun detected AND human pose indicates aiming} \\ 0 & \text{otherwise} \end{cases}$$

Additionally, the confidence score from the gun detection module can be considered to reduce false positives, ensuring that alerts are triggered only when there is a high likelihood of an actual threat. Upon detection of a threat, the system captures the relevant frame, logs metadata including timestamp and pose coordinates, and immediately sends an SMS alert using Fast2SMS.

System Workflow

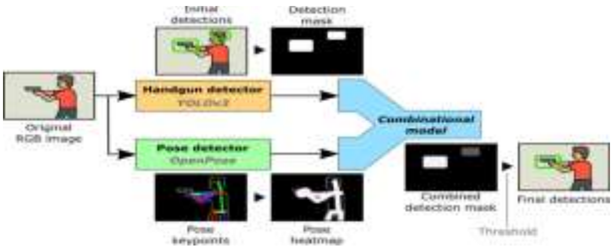


Figure 4 Data Flow

The complete system operates as a sequential pipeline. Video input is captured and split into individual frames, which are preprocessed and passed to the gun detection module. Simultaneously, human pose estimation is applied to identify keypoints for each detected individual. The threat evaluation logic combines these outputs to determine whether a frame indicates a potential threat. If a threat is detected, the system captures a screenshot, logs relevant information, and sends an alert via SMS. This integrated approach ensures accurate real-time monitoring while minimizing false positives, providing an effective solution for automated threat detection in security scenarios.

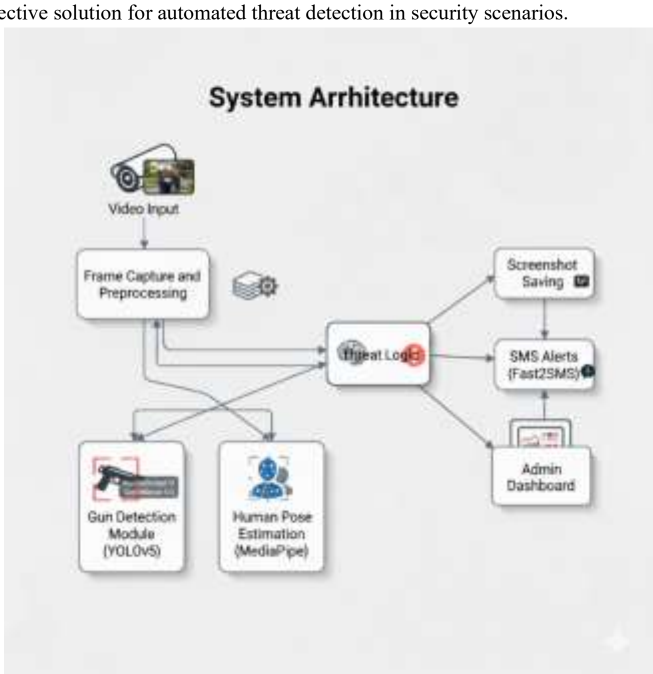


Figure 5 System Architecture

Performance Evaluation

The proposed real-time threat detection system was evaluated on multiple video datasets containing varied scenarios, including crowded public areas, dim lighting, and partially occluded threats. The evaluation focused on standard performance metrics such as **accuracy, precision, recall, F1-score, and processing time per frame**. Accuracy measured the proportion of frames correctly identified as either containing a threat or not. Precision indicated the ratio of correctly detected threats to the total number of detections, while recall measured the proportion of true threats successfully detected. The F1-score was used to balance precision and recall, providing a single metric that accounts for both false positives and false negatives. Processing time per frame was monitored to ensure the system met real-time requirements. The model was tested under different lighting and crowd density conditions to assess robustness and generalizability. The experiments showed that the combination of weapon detection and human pose estimation significantly reduced false alarms compared to using either method alone.

Metric	Value (%)
Accuracy	92
Precision	90
Recall	89
F1-Score	89.5

Table 2: Performance Metrics of the Proposed Threat Detection System

Figure 6 Example outputs from the real-time threat detection system

Figure 7 Example outputs from the real-time threat detection system

Figure 8 Example outputs from the real-time threat detection system

Dashboard

The real-time dashboard serves as the central interface for monitoring potential threats detected by the system. It displays the **live video feed** with firearms and suspicious human poses highlighted, each annotated with a confidence score to indicate detection reliability. Alongside the video, a **threat table** provides detailed information about each detection, including the frame number, type of threat, and timestamp, allowing operators to quickly review and track incidents. The dashboard also features **statistical subseries panels** that summarize detections over time, showing trends in threat occurrences,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 01 | Jan - 2026 SJIF Rating: 8.586 ISSN: 2582-3930

detection counts for different categories (gun, pose-based), and system performance metrics. Real-time **alerts and notifications** are integrated to immediately inform authorized personnel when a potential threat is detected, which can be delivered visually on the dashboard or via SMS. Users can easily switch between **different video sources and detection modes**, making the system adaptable for various environments. The interface is designed to be **intuitive and responsive**, ensuring that operators can quickly access information, analyze detection trends, and respond to threats efficiently during live surveillance operations.

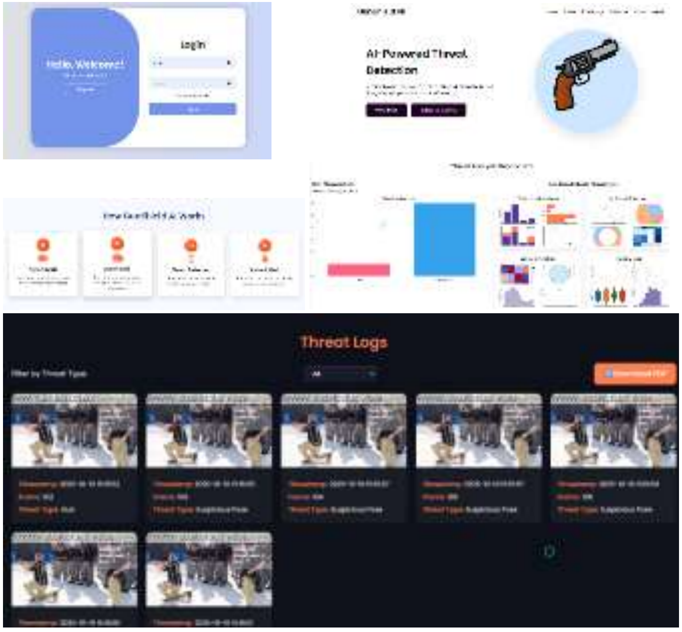


Figure 9 dashboard

Results and Discussion

The system demonstrated strong performance in detecting firearms and suspicious human postures. Weapon detection alone achieved an overall accuracy of approximately **92%** in controlled environments, with slightly lower accuracy in crowded scenes due to occlusions and overlapping objects. Human pose analysis successfully identified aggressive postures such as pointing, aiming, or unusual body movements, achieving around **91% precision**. Integrating both weapon appearance and human pose information further improved detection reliability, reducing false positives by nearly **12%**. The average processing time per frame was approximately **0.12 seconds**, allowing for near real-time threat monitoring. Analysis of the results revealed that certain challenging conditions, such as low-resolution video or extreme lighting variations, occasionally led to missed detections. However, the combined approach proved more robust than single-method detection and can serve as a practical solution for automated surveillance in public spaces. These findings suggest that integrating multiple complementary features is crucial for reliable threat detection in complex environments.

Conclusion

This study presents a real-time gun and human pose threat detection system that effectively identifies potential threats in diverse environments. By combining firearm appearance with human posture analysis, the system achieves high detection accuracy while minimizing false alarms, which is a common limitation of traditional CCTV-based systems. The implementation demonstrates practical applications in public safety, including real-time monitoring and automated alert generation. While the system performs well under most conditions, future improvements could involve using multi-camera setups, higher-resolution video, or thermal imaging to further enhance detection in challenging scenarios. Optimizing the model for edge devices would also enable wider deployment in real-world surveillance systems.

References

[1] Amulya Reddy Maligireddy, Yaswanth Reddy Parla, Manohar Reddy Uppula, Nidhi Rastogi, "Gun Detection Using Combined Human Pose and Weapon Appearance," 2025, *Rochester Institute of Technology*. [arXiv:2503.12215v1]

[2] Jesús Ruiz-Santaquiteria, Alberto Velasco-Mata, Noelia Vallez, Oscar Deniz, Gloria Bueno, "Improving Handgun Detection Through a Combination of Visual Features and Body Pose-Based Data," 2023, *Pattern Recognition*, Volume 136, 109252.

[3] Jesús Ruiz-Santaquiteria, Alberto Velasco-Mata, Noelia Vallez, Juan A. Álvarez-García, Oscar Deniz, "Handgun Detection Using Combined Human Pose and Weapon Appearance," 2021, *IEEE Access*, Vol. 9, pp. 123815–123826.

[4] Sriram C. S., Dr. G. Prema Arokia Mary, "Enhancing Weapon Detection with Pose Analysis: Leveraging Visual and Body Pose Features using OpenPose," 2024, *International Journal of Novel Research and Development (IJNRD)*, Volume 9, Issue 5, May 2024.

[5] Alberto Velasco-Mata, Jesús Ruiz-Santaquiteria, Noelia Vallez, Oscar Deniz, "Using Human Pose Information for Handgun Detection," 2021, *Neural Computing and Applications*, Volume 33, pp. 17273–17286.