RESEARCH ARTICLE

RSP Science Hub

International Research Journal on Advanced Science Hub
2582-4376
www.rspsciencehub.com
Vol. 08, Issue 06 June

Check for updates

http://dx.doi.org/10.47392/IRJASH.2026.029

Threat Sense-XAI: A Context-Aware Threat Intelligence Framework for Real-Time Weapon, Violence, and Behavioural Risk Assessment

Megha Saha[1], R. Velvizhi Ramya[2]
[1]*PG Scholar, Dept. of DS, AMC Engineering College, Bangalore, 560083, Karnataka, India*
[2]*Assistant Professor, Dept. of CSE, AMC Engineering College, Bangalore,560083, Karnataka, India*
Emails: *meghasaha1322@gmail.com[1], rvramya2021@gmail.com[2]*

Article history
Received: 05 June 2026
Accepted: 24 June 2026
Published: 24 July 2026

Keywords:
Artificial Intelligence; Smart Surveillance; Threat Detection; Threat Intelligence; Explainable AI.

Abstract
Classic approaches to security monitoring often encounter problems when dealing with detecting and resolving security threats in real time. This paper introduces a system that uses context-aware threat intelligence to detect weapons, recognize violent actions, analyze time-based patterns, and assess potential threats based on the situation, all to spot unusual behaviour in real-time video feeds. The proposed system utilizes YOLO-based deep learning models to detect weapons and classify human behaviour. Alternatively, a scoring engine determines the level of risks from the objects observed, the behaviour pattern, the location, and past occurrences. High-risk occurrences will automatically be accompanied by raising an alarm, collecting the evidence, reporting the occurrence, and sending emails. In order to foster openness in the process and increase the level of confidence, Explainable Artificial Intelligence (XAI) methods including SHAP (SHapley Additive exPlanations) and Grad-CAM are applied to enhance transparency. Also, a real-time dashboard for incidents is put up for monitoring and analysing incidents.

1. Introduction

There is a requirement for enhanced technology that will enable security surveillance networks to automatically recognize abnormal activities. Today's security surveillance networks are largely dependent on manual operation, which makes real-time surveillance impossible and also creates problems whenever an emergency occurs. Thanks to recent developments in deep learning, weapons and violent activities can be detected automatically, thus increasing the performance and reactivity of modern surveillance systems [6]. Current systems analyze incidents individually, thereby ignoring the background. To solve this problem, we propose a novel framework consisting of five components: weapon detection, violence recognition, temporal activity analysis, contextual threat evaluation, and Explainable AI. The framework will help assess any suspicious activities more accurately and prevent false alerts, making it ideal for use in sensitive areas [14].

1.1. Objectives

- Detection of weapons from videos through deep learning algorithms for object detection.
- Classification of violent vs non-violent activities of people in video frames.
- Estimation of threat level in relation to weapons, human behaviour, time, etc.
- Real-time alerts and reports of incidents for security purposes.

- Use of Explainable AI methods such as Grad-CAM, SHAP for increasing transparency in the system.

2. Literature Survey

Table 1 Comparison of Existing Studies

Ref.	Method	Advantages	Limitations
[5]	VD-Net violence detection	Real-time detection	No weapon context
[6]	DL-based weapon detection	Accurate detection	Weapon-only focus
[7]	YOLO weapon detection	Fast inference	No behaviour analysis
[9]	CNN–LSTM violence recognition	Temporal modelling	No weapon integration
[10]	Grad-CAM framework	Explainable results	Limited threat classification
[13]	Vision Transformer	High accuracy	High computation cost
[14]	Context-aware assessment	Uses scene context	No integrated detection

3. Research Gap

The existing surveillance systems concentrate on weapon detection or violence analysis separately. Although some research attempts to improve the detection rate and some include contextual information into their models, only a small number of systems implement weapon detection, violence analysis, threat assessment in time, the history of incidents, and explainable AI at the same time. Most current solutions do not address the problem of automation as well. All these factors create a need for smarter surveillance systems that can handle different aspects of threat evaluation at once.

4. Proposed Methodology

4.1. System Architecture

Figure 3.1. processes real-time surveillance footage through person detection, weapon detection, and behaviour classification modules. The extracted information is analyzed by the threat intelligence module to assess threat levels and generate video evidence, email alerts, incident reports, alarm notifications, and dashboard updates for security monitoring.

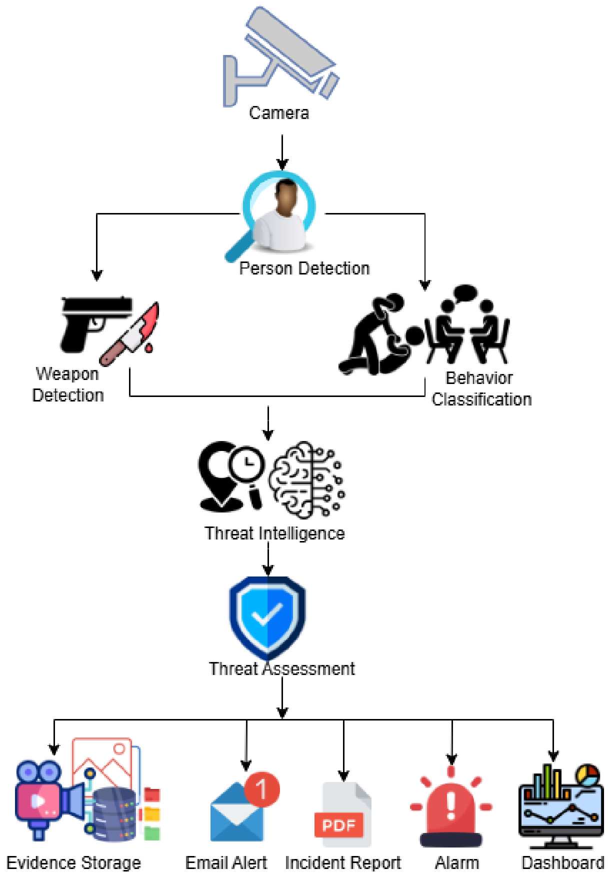


Figure 1 Architecture of the Proposed Threat Intelligence Framework

4.2. Workflow Diagram

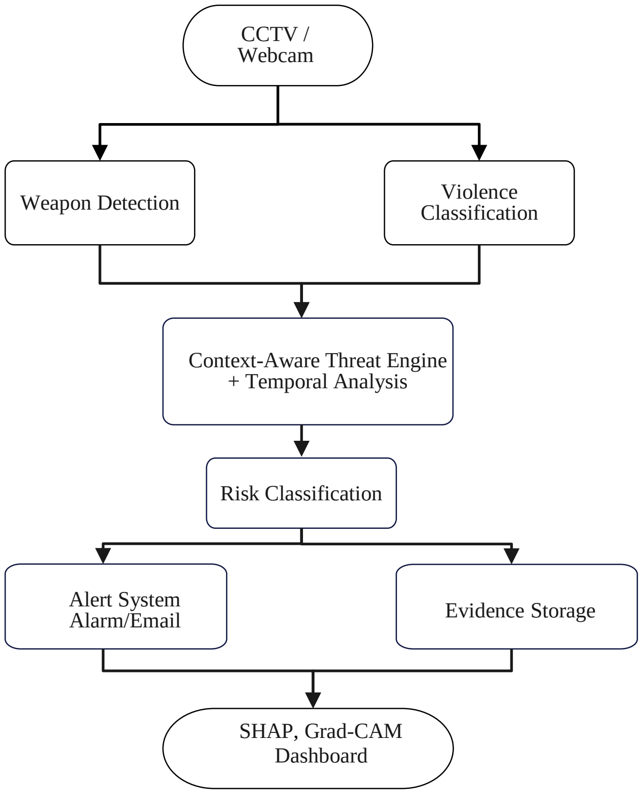


Figure 2 Proposed System Workflow

Figure 2. represents the complete flow chart of the system being proposed. The video frames obtained through the surveillance camera undergo processing for weapon detection and behaviour analysis. The obtained data is sent to the threat intelligence engine, where temporal analysis and context-based reasoning take place. The threat score computed from this process decides the threat level. In case of high risk, the system triggers alarms, logging, emailing, and dashboard management.

4.3. Dataset Description

Training and Validation processes of our system are based on two different sets of data. The weapon detection dataset includes six classes, which are Knife, Holstered Handgun, Unholstered Handgun, Rifle on Person, Rifle off Person, and Explosive, with bounding box information for object detection tasks using the YOLO network [6]. The human behaviour classification dataset includes three activities: Normal, Grappling, and Striking, allowing the system to detect normal and violent human interactions [5]. To enhance the performance and to reduce overfitting, we employed several data augmentation strategies, such as rotation, flipping, zooming, and changes in brightness.

4.4. Weapon Detection Module

The weapon detection component detects weapons in the frames coming from the surveillance video in real time. The weapon detection task utilizes a YOLO-based object detection algorithm because of its higher level of detection efficiency along with real-time processing ability. Every frame that comes in is processed to identify whether any weapon exists in it and the confidence score associated with it. Weapons include knives, handguns, rifles, and explosives among others. Once the confidence score crosses the pre-defined threshold value, the detected weapon is passed to the threat intelligence module. YOLO-based object detection algorithms perform well in real-time surveillance systems [7].

4.5. Behaviour Classification Module

The behaviour classification model will analyze any actions carried out by people in the recorded surveillance video and classify them in various categories like Normal, Grappling, and Striking. A deep learning classifier will examine each image and score each behaviour class. The class having the highest score will be considered the prediction. There is also a confidence level required to be met in order to assign a particular class. This module will assist the application in detecting possible aggressive actions [5].

4.6. Threat Intelligence Engine

The Threat Intelligence Engine serves as the decision-making component of the proposed system by analyzing the combination of data obtained from weapon detection, behaviour classification, temporal analysis, and contextual analysis to determine the criticality of an event. The threat score is obtained through the summation of weapon, behaviour, temporal, and contextual scores ensuring a thorough risk assessment beyond the isolated identification of a single event [14].

$$\begin{aligned} Threat\ Score = {} & Weapon\ Score \\ & + Behavior\ Score \\ & + Temporal\ Score \\ & + Context\ Score \end{aligned}$$

Table 2 Threat Level Classification

Threat Score Range	Risk Level
0 – 20	Low
21 – 70	Medium
71 – 110	High
> 110	Critical

Table 3.1. shows score obtained from the threat intelligence engine, each event is assigned a risk level, which could be one of the following: Low Risk (0-20), Moderate Risk (21-70), High Risk (71-110), or Critical Risk (>110). Table 3.1. shows a summary of the risk levels.

4.7. Explainable AI

To increase transparency in the proposed surveillance system, techniques of explainable artificial intelligence are used. Heatmaps are generated using the Grad-CAM method that highlights the parts of an image responsible for classification of the observed behaviour [10]. Additionally, SHAP technique is used to explain how weapon detection, behaviour assessment, timing factors, and the context influence the process of calculating the threat score [14].

4.8. Real-Time Alert Generation

In case of a threat score that is higher than the present level, an automatic alarm is triggered, and the response mechanism kicks into action. The proposed system takes pictures of the evidence, captures a video snippet of what happened, stores the gathered evidence, logs the details, produces an incident report, and sends emails to the security team members [11].

5. Experimental Results and Performance Analysis

5.1. Normal Behaviour Detection

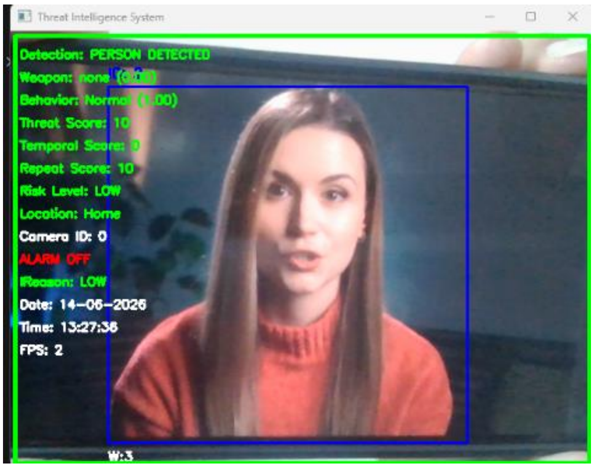


Figure 3 **Normal Behaviour Detection Result**

Figure 4.1. shows a normal human activity sample processed by the proposed Threat Intelligence System. The behaviour classification module correctly identified the activity as Normal with high confidence and assigned a low threat score. Since no suspicious behaviour or weapon was detected, the system maintained a low-risk status and no alarm was triggered.

5.2. Grappling Behaviour Detection Result

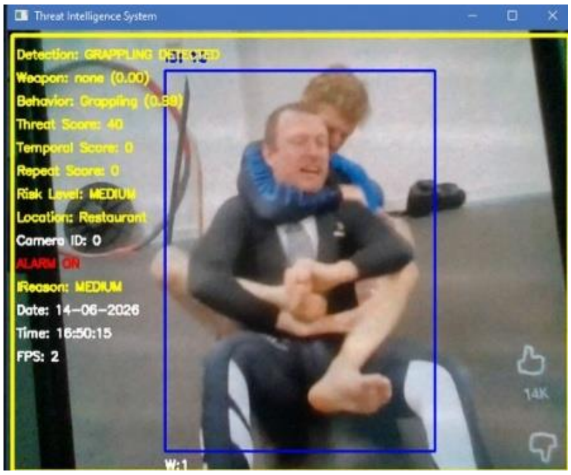


Figure 4 **Grappling Behaviour Detection Result**

Figure 4.2. shows an example of the grappling activity detection carried out by the proposed framework. The model correctly identifies physical engagement between persons and classifies it as a Grappling event. Based on these results, the behaviour classification, the threat intelligence engine increases the score of threats and categorizes the event as one having medium-level risks.

5.3. Weapon Detection Result

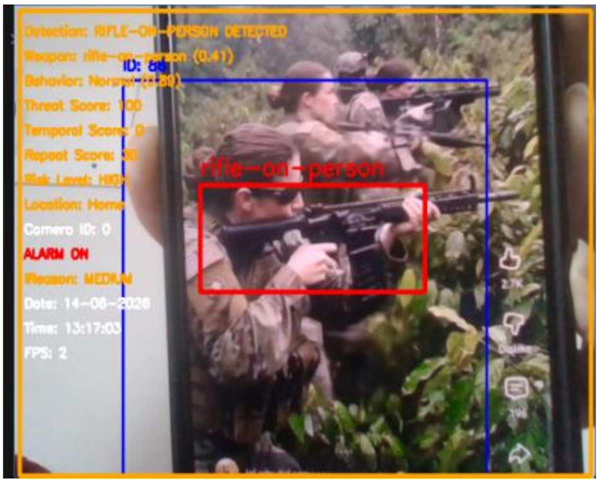
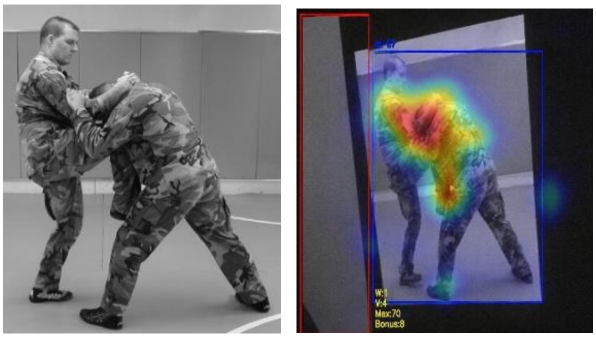


Figure 5 **Rifle Detection and Threat Assessment**

Figure 5 is an output from the weapon detection module. The proposed framework correctly identifies the location of the weapon through a bounding box while also calculating the high threat score because of the possible threat posed by the object. The threat intelligence engine thus classifies the event as high risk and sets off the alarm system.

5.4. Explainable AI

5.4.1. Grad-CAM Visualization

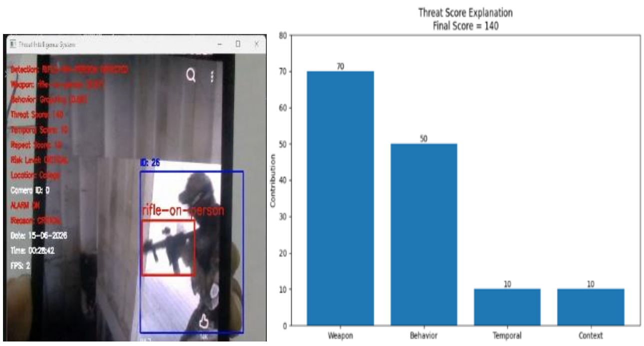


(a) Original Image (b) Grad- CAM Image

Figure 6 Grad-CAM Visualization for Grappling Activity Detection: (a) Original Image; (b) Grad-CAM Heat Map Highlighting Important Regions

Figure 6 shows the data obtained from the grad-CAM analysis of grappling activity detection. In Figure 4.4(a), we see the input image whereas Figure 4.4(b) depicts the heat map produced by grad-CAM. From Figure 4.4. the data proves that the highlighted areas mainly consist of the physical interaction between people.

5.4.2. SHAP Visualization



(a) Threat detection result (b) SHAP contribution analysis

Figure 7 SHAP-Based Threat Explanation: (a) Threat Detection Result; (b) SHAP Contribution Analysis.

Figure 7 depicts the explanation for the process of threat assessment by means of SHAP analysis. The threat is depicted in Figure 4.5(a). Figure 4.5(b) explains how weapon, behaviour, time, and context contribute to the final threat score.

5.5. Behaviour Classification Performance

5.5.1. Classification Report

Table 3 shows the classification behaviours attained an accuracy of 95%, which reflects successful classification of behaviours, both normal and violent.

Table 3 Behaviour Classification Performance

Class	Precision	Recall	F1-Score
Grappling	0.93	0.99	0.96
Normal	1.00	1.00	1.00
Striking	0.84	0.52	0.64
Overall Accuracy			**95%**

5.5.2. Confusion Matrix

Figure 8 Confusion Matrix of Behaviour Classification

From Figure 4.6, we can see that the model was able to accurately classify all samples from Class 0 and Class 2, although there is a misclassification for one sample of Class 1, which was erroneously identified as belonging to Class 2.

5.6. Weapon Detection Performance

Table 4 Weapon Detection Performance

Metric	Value
Precision	69.5%
Recall	44.3%
mAP@50	51.3%
mAP@50-95	31.6%

Table 4 indicates that the mAP@50 of the system model was 51.3%, which suggests that the model is able to classify different weapon categories during surveillance.

5.7. 4.7 Ablation Study

Table 5 Ablation Study of Threat Intelligence Engine

Configuration	Threat Score	Risk Level
Weapon Only	0	LOW
Behaviour Only	40	MEDIUM
Weapon + Behaviour	70	MEDIUM
Full Threat Engine	130	CRITICAL

Table 5 represents the ablation study, whereby it is evident that using weapon detection, behaviour classification, and threat intelligence scoring improves architectural performance of the system

Conclusion and Future Scope

In the paper, an idea of developing an intelligent surveillance system was considered. The system under consideration uses the ability of weapon detection, behavioural analysis, contextual data, and explainable AI in order to estimate threats. Specifically, the system can detect potential threats based on its ability to generate dynamic risk scores considering detected behaviour and context. According to the experimental results, the proposed solution works effectively and has higher efficacy than alternative solutions in relation to behavioural analysis and weapon detection tasks.

Future Scope

The future areas of development of this architectural model include the development of an integrated intelligent surveillance system that incorporates other functions apart from what currently exists. This would include functions such as detection of anomalies, crowd analytics, person re-identification, and prediction of threats. This would contribute to better situational awareness of the threat intelligence system. Other functions that include the use of deep learning models, edge computing, and multimodal sensor systems will also be important in improving the efficiency of the system.

References

[1]. S. N. S. Nair et al., "Image Detection and Segmentation Using YOLOv5 for Surveillance," Proc. Int. Conf. Software Engineering and Machine Learning, 2023.

[2]. R. Chatterjee, R. R. Choudhury, M. K. Gourisaria, S. Banerjee, S. Dey, M. Sahni and E. León-Castro, "Temporal-Aware Transformer Approach for Violence Activity Recognition," IEEE Access, vol. XX, pp. XX–XX, 2025.

[3]. N. Hnoohom et al., "Weapon Detection Using Faster R-CNN Inception-V2 for a CCTV Surveillance System," IEEE ICSEC, 2021.

[4]. R. Ramya et al., "An Intelligent Surveillance System for Weapon Detection Based on EfficientDet Algorithm," IEEE ICIRCA, 2023.

[5]. M. Khan et al., "VD-Net: An Edge Vision-Based Surveillance System for Violence Detection," IEEE Access, 2024.

[6]. M. M. Hassan, M. H. Rahman and M. S. Hossain, "A Deep Learning-Based Smart Surveillance Framework for Weapon Detection in Public Areas," IEEE Access, vol. 12, pp. 45231-45245, 2024.

[7]. A. S. Alqahtani, F. Alenezi and M. Alotaibi, "Real-Time Weapon Detection in CCTV Surveillance Using YOLO-Based Deep Learning Models," 2024 International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 201-206, 2024.

[8]. P. Kumar, R. Sharma and S. Gupta, "Intelligent Video Surveillance System for Violence Detection Using Deep Neural Networks," IEEE, ICCCIS, pp. 118-123, 2023.

[9]. S. Roy, A. Chatterjee and K. Bhattacharya, "Human Activity Recognition for Violence Detection in Surveillance Videos Using CNN-LSTM Architecture," IEEE Access, vol. 11, pp. 125630-125642, 2023.

[10]. Y. Wang, J. Li and H. Zhang, "Explainable Artificial Intelligence for Video Surveillance: A Grad-CAM-Based Threat Analysis Framework," IEEE Access, vol. 13, pp. 18754-18768, 2025.

[11]. S. S. Chavan, A. S. Patil and R. R. Joshi, "Real-Time Smart Surveillance System for Suspicious Activity Detection Using Deep Learning," Proceedings of the IEEE International Conference on Intelligent Systems and Machine Vision, pp. 145-150, 2024.

[12]. M. A. Albahar, A. Alharbi and M. Alqahtani, "Deep Learning-Based Weapon

Detection for Public Safety Surveillance Applications," IEEE Access, vol. 12, pp. 78120-78135, 2024.

[13]. Y. Li, H. Chen and X. Wang, "Vision Transformer-Based Human Violence Recognition in Surveillance Videos," IEEE Access, vol. 13, pp. 24567-24580, 2025.

[14]. P. Singh, S. Verma and A. Kumar, "Context-Aware Threat Assessment Framework for Intelligent Video Surveillance Systems," Proceedings of the IEEE International Conference on Artificial Intelligence and Security, pp. 88-94, 2024.

[15]. J. Rodriguez, E. Leon-Castro and R. Martinez, "Explainable Deep Learning for Security Surveillance Using Class Activation Mapping Techniques," IEEE Access, vol. 13, pp. 32410-32425, 2025.