IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141053

AI Surveillance and Crime Detection: A Literature Review

Mahima A[1], Pranamya K L[2], Shreya R[3], Siva Harshitha[4]

Dept. of CSE – ICB, K.S. Institute of Technology, Bengaluru, India[1-4]

Abstract: The growing incidence of street crimes such as theft, robbery, assault, harassment, and illegal weapon possession has highlighted the urgent need for real-time surveillance and rapid response systems. This paper proposes a mobile-based crime detection framework that transforms everyday smartphones into intelligent CCTV devices capable of identifying suspicious activity. Leveraging deep learning techniques, the system detects violent behavior, identifies weapons (e.g., knives and firearms), and performs facial recognition to detect known suspects. Upon threat detection, alerts are sent instantly to authorities along with supporting evidence such as images, timestamps, and location. The proposed system offers a scalable, affordable, and effec- tive solution for urban surveillance—especially in areas lacking traditional CCTV infrastructure—thus enabling quicker law enforcement response and contributing to safer, smarter cities.

Keywords: Crime Detection, Mobile Surveillance, Weapon Detection, Violence Recognition, Real-Time Alert System.

I. INTRODUCTION

Ensuring public safety in urban spaces has become increas- ingly important, especially as street crimes like theft, assault, harassment, and illegal weapon possession continue to rise. While traditional CCTV-based surveillance systems are widely deployed, they often fall short of preventing crime in real time. These systems typically rely on manual monitoring, have limited field of view, and lack intelligent analytics, which means critical incidents may go undetected or are discovered only after the damage is done. The rapid growth of artificial intelligence (AI), particularly in computer vision and deep learning, is now transforming how surveillance sys- tems operate. Modern AI-enabled solutions can analyze video streams in real time, recognize suspicious behavior, detect objects like weapons, and even identify individuals through facial recognition. These capabilities significantly improve the chances of preventing or quickly responding to crime. Most existing AI-based surveillance systems are designed for fixed cameras or dedicated edge devices like Raspberry Pi. However,a major gap exists in solutions tailored for mobile platforms. Smartphones—with their high-resolution cameras, increasing computational power, and widespread availability—offer a unique opportunity to bring smart surveillance to places where traditional setups are impractical or too expensive.In this paper, we explore the potential of using smartphones as portable, real-time crime detection units. We present a detailed literature survey covering recent advances in violence detection, weapon recognition, and alert systems, with a focus on mobile and edge-based implementations. By identifying limitations in cur- rent research and proposing a flexible, multi-module detection framework, this work aims to contribute to the development of affordable, scalable, and intelligent urban surveillance systems.

II. BACKGROUND AND MOTIVATION

Urban environments today face a growing challenge in maintaining public safety, as incidents of street crimes—ranging from petty theft and harassment to violent assaults and illegal weapon possession—continue to increase. Surveillance technologies like CCTV have long been used as a preventive measure, but their effectiveness is often limited. Traditional systems rely heavily on continuous human monitoring and are not capable of analyzing or responding to threats automatically. As a result, many criminal activities go unnoticed until after they've occurred, reducing the chances of timely intervention.

This limitation has driven interest in smarter surveillance solutions powered by Artificial Intelligence (AI). AI-based video analytics can process live footage in real time, iden- tify suspicious behavior, detect dangerous objects such as weapons, and even recognize faces. These intelligent features have the potential to not only detect crimes but also to help prevent them by enabling faster responses from law enforcement.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141053

However, most current implementations of AI-powered surveillance are based on fixed camera setups or resource-constrained edge devices like Raspberry Pi. While effective in certain scenarios, these systems lack the flexibility and scalability required for widespread use, especially in areas without existing infrastructure. Smartphones, on the other hand, offer a promising alternative. With their powerful pro- cessors, high-quality cameras, and near-ubiquitous availabil- ity, mobile phones can serve as portable surveillance tools. Leveraging mobile devices for AI-based crime detection brings both affordability and convenience, making smart surveillance more accessible—particularly in under-monitored or high-risk public areas.This project is motivated by the need to bridge the gap between advanced AI capabilities and real-world public safety challenges. By using smartphones as intelligent, mobile surveillance units, we aim to create a scalable, real-time crime detection system that can enhance urban safety and support faster law enforcement responses.

III. LITERATURE SURVEY

Venkatesh et al. [1] present a lightweight real-time crime detection pipeline built for edge devices (Raspberry Pi class). Their approach combines Early-Stopping Multiple Instance Learning (MIL) with optical-flow based action features to prioritize short, informative video segments and reduce wasted computation. The paper shows promising accuracy on bench- mark action datasets while keeping inference latency low, but it focuses on fixed edge nodes and does not address object- level threats (for example weapon detection) or mobile-phone deployment scenarios, which limits its applicability in highly dynamic public spaces.

Gao et al. [2] propose a multi-model CNN pipeline using YOLO variants (v5/6/7) and Faster R-CNN to detect property-related crimes such as arson, vandalism and burglary. They complement the detection backbone with a web interface (Gradio) and automated alerting via Twilio, demonstrating a full detection-to-notification workflow. The system is ro- bust for property-crime scenarios and ensemble modeling helps reduce false positives, but it is primarily targeted at fixed-camera installations and lacks modules for interpersonal crimes (assault, harassment) or mobile edge deployment. Thakur et al. [3] designed a YOLOv8-based weapon de- tector for real-time identification of knives and firearms in surveillance streams. They trained and validated on curated firearm/knife datasets and report high mAP and low inference latency suitable for real-time monitoring. The work is valuable as a focused object-detection solution, yet it does not integrate behavioral or temporal context (i.e., it cannot tell if a weapon is being brandished versus casually carried) and lacks face- recognition or multi-modal corroboration that would reduce false alarms.

De Paula et al. [4] introduce *CamNuvem*, a robbery-centric dataset of hundreds of real CCTV robbery clips recorded in commercial settings. CamNuvem emphasizes real-world contextual cues and temporal dynamics, enabling weakly supervised anomaly detectors (RADS, WSAL, RTFM) to be evaluated on robbery detection. The dataset closes a gap for robbery-specific research, but its narrow crime focus and regional composition mean models trained on it may not generalize well to other crime types or environments without further augmentation.

Vijeikis et al. [5] present a compact violence-detection model combining U-Net style spatial encoders with LSTM temporal modules, optimized for embedded hardware. Their architecture targets one-second video snippets and achieves an attrac- tive speed/accuracy trade-off on datasets like RWF-2000 and Hockey Fights. While well suited to low-power deployments, the approach treats violence as a binary label and does not model weapons or identity — limiting forensic usefulness and nuance when differentiating between rough play and real threats.

Akdag et al. [6] propose TeG (Temporal Granularity), a transformer-based method that analyzes video at multiple temporal scales to detect both short bursts (e.g., sudden fights) and long-term anomalies (e.g., loitering then theft). TeG shows strong performance in crowded urban camera settings and improves recall for subtle or slow-developing events. The drawback is its computational heft — Swin-transformer components and multi-scale processing are resource-intensive, so direct mobile/edge deployment requires heavy pruning or server-side offload.

Ullah et al. [7] provide a comprehensive survey of vision- based violence detection approaches, cataloguing CNN, RNN/LSTM, and hybrid architectures and summarizing datasets and evaluation metrics. Their synthesis is helpful as a roadmap and emphasizes the importance of spatio-temporal fusion and lightweight models for edge scenarios. As a survey, it guides design choices but does not deliver an implementation— it highlights dataset and generalization issues that many later works still face.

Negre et al. [8] review deep-learning methods for violence detection in video, with attention to input modalities (RGB, optical flow, skeletons) and to hybrid two-stream or 3D-CNN designs. They stress that dataset imbalance, occlusion and scene variability limit cross-dataset generalization and recom- mend transfer learning and multimodal fusion as

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141053

remedies. The paper is useful for researchers selecting architectures but again is descriptive rather than prescriptive (no turnkey system).

Song and Nang [9] implement an edge–server hybrid archi- tecture where inexpensive edge sensors perform rule-based filtering and suspected events are escalated to server-side deep models. This architecture reduces bandwidth and server load while retaining high accuracy for complex inference. The trade-off is an added dependency on network availability and potential latency; the system primarily targets pedestrian abnormality detection and does not include dedicated weapon or identity modules.

Mukto et al. [10] describe a near end-to-end crime-monitoring system that integrates YOLOv5 for weapons, MobileNetV2 for violence classification, and LBPH for face recognition, combined with a dashboard and alert generation. Their work demonstrates the benefits of multi-module fusion for practical monitoring. However, the design assumes fixed CCTV sources and server-side processing — adaptation to mobile smartphone-based capture would require model compression and re-engineering of the pipeline.

A general survey on Visual Anomaly Detection (VAD) [11] highlights the major categories (reconstruction-based, prediction-based, and scoring-based) and the scarcity of labeled anomaly data, especially for rare or safety-critical events. The paper argues for few-shot and semi-supervised approaches and proposes evaluation metrics tailored to anomaly tasks. This viewpoint is useful for crime detection where labeled crime events are sparse, but direct application to multi-class crime detection needs targeted dataset collection. The "Quantum-AI Empowered Intelligent Surveillance" study [12] explores hybrid quantum-classical architectures (Quantum-RetinaNet) as a future direction for faster feature extraction and contraband detection. The authors also intro- duce a large Street Crimes Arms Dataset (SCAD) curated for under-represented weapon types. While imaginative and promising for large-scale acceleration, QAI solutions are still experimental and depend on nascent quantum hardware; prac- ticality for near-term mobile deployment is limited.

A broad survey of video surveillance systems in smart cities (Electronics) [13] examines the integration of edge computing, blockchain for data integrity, and deep learning for analytics. It outlines applications (traffic, environmental monitoring, pub- lic safety) and emphasizes multimodal sensing and privacy- preserving design. The review is conceptually valuable for po- sitioning surveillance within smart-city stacks, but it does not resolve low-level deployment constraints for mobile devices or real-time alert latencies.

An industry-facing work on enhancing real-time weapons detection [14] proposes practical strategies: scale-matching to boost small-object detection, a temporal second-stage classifier to reduce false positives, and a large DISARM CCTV-style weapon dataset. Their pipeline reports significant improve- ments in small-weapon mAP and false-positive reduction. This engineering-focused study demonstrates what dataset engineering and post-processing can achieve, yet its reliance on heavy back-end compute limits immediate on-device use. A real-time pedestrian abnormal behavior system [15] couples YOLOv4 detection with KCF tracking and a 3D-ResNet action recognizer to identify loitering, falls, intrusion, and violence. Evaluated on the KISA dataset, the system achieves high precision and recall and validates RTSP-based live-stream operation. Its modular design supports real-world CCTV use, but it requires tuning for camera angle variability and lacks native weapon-detection or identity resolution modules.

Argus++ [16] proposes a robust cube-proposal approach for activity detection in unconstrained, untrimmed video streams. By generating overlapping spatio-temporal cubes and dedu- plicating outputs, Argus++ is resilient to fragmentary events and broad field-of-view scenes. The method performs well on benchmarks but still depends on solid detection/tracking backbones and may need adaptation to run on constrained edge hardware used in mobile scenarios.

A second survey on video surveillance systems for smart cities [17] reiterates the importance of edge-cloud balance, privacy-aware storage, and multi-sensor fusion; it also notes open issues like false alarm mitigation and unsupervised anomaly discovery. The paper maps potential application domains and future trends such as drones and embedded analytics, which informs design choices, but again leaves practical mobile-inference details to implementers.

Deep Learning-Based Anomaly Detection in Video Surveil- lance [18] categorizes approaches (reconstruction, prediction, classification) and highlights the role of transfer learning and multimodal fusion in improving robustness. The survey suggests practical evaluation protocols and stresses the need for real-time feasible architectures – a direct concern for mobile deployments – while acknowledging that many high- performing models remain too heavy for on-device use without optimization.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⵘ Peer-reviewed & Refereed journal ⵘ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141053

A study on lightweight surveillance (Haar/SVM → L- CNN) [19] traces the evolution from classical detectors to compact CNNs optimized for edge inference. The authors demonstrate that with depthwise separable convolutions and careful pruning, viable real-time human detection on micro- devices is possible. The limitation is that these lightweight detectors typically handle only presence detection and require additional modules to infer intent or weapon possession. MSD-CNN for abnormal object detection [20] introduces a two-branch lightweight CNN that classifies frames into normal and abnormal subclasses (e.g., guns, knives). The approach reports strong accuracy on multi-view inputs and proposes a Detection Time per Interval (DTpI) metric to quantify real- time responsiveness. MSD-CNN strikes a pragmatic balance between speed and accuracy, but it still faces challenges on very low-resolution or heavily occluded feeds typical of some urban cameras.

Another survey of video surveillance systems [21] again synthesizes the state of the art, compiling use-cases, key algorithms (action recognition, tracking, anomaly detection) and deployment patterns. Its broad outlook supports interdis- ciplinary system design, however specific model-level imple- mentation guidance for mobile inference is limited.

A horizontal review on video surveillance for smart cities [22] emphasizes edge device capabilities, common datasets, and application trends (people counting, HAR, vehicle tracking). The review recommends tiered architectures and highlights the vendors providing on-device analytics (e.g., FLIR, Bosch). It reinforces the necessity to design surveillance models that respect latency, bandwidth, and privacy constraints when tar- geting city-scale deployments.

Event Detection in Surveillance Videos: A Review [23] fo- cuses on methods for detecting events in long, untrimmed videos. It discusses spatio-temporal feature extraction, tem- poral segmentation, and 3D-CNN frameworks, and stresses the importance of temporal localization precision. For crime- detection tasks where events can be brief and ambiguous, the survey's recommendations on temporal modeling and multi- target tracking are directly applicable.

A Survey on Visual Anomaly Detection: Challenge, Approach, and Prospect [24] revisits VAD taxonomy and emphasizes few- shot, zero-shot, and data augmentation approaches to combat label scarcity. It underlines promising directions — foundation models, multi-modal fusion, and semantic-aware scoring — which are relevant for building generalized crime detectors, but implementing these for real-time mobile inference is an ongoing research challenge.

Finally, a second real-time pedestrian abnormal behavior paper [25] (appearing again in the report) consolidates KCF/YOLOv4 tracking and dedicated classifiers to achieve high F1 scores on KISA-like datasets. The practical testing with RTSP streams demonstrates deployability; still, like many systems, it needs extension for weapon detection, identity tracking, and privacy-preserving logging to be production- ready in sensitive public deployments.

IV. RESEARCH GAP AND PROBLEM STATEMENT

The literature review of 25 recent works in AI-based crime detection highlights significant progress in areas such as violence detection, weapon recognition, anomaly detection, and integrated surveillance systems. However, critical limita- tions remain unaddressed. Most existing approaches are de- signed for fixed CCTV infrastructures or server-based deploy- ments, which limits their applicability in dynamic, resource- constrained environments. Only a small subset of the sur- veyed methods support deployment on mobile or low-power edge devices, and even fewer offer an integrated multi-threat detection capability covering violence, weapons, and suspect identification simultaneously.

Another common gap is the lack of real-time alerting systems. While several works achieve high detection accuracy, they fail to bridge the last mile by promptly notifying relevant authorities or stakeholders. Furthermore, the datasets used in most studies lack diversity, particularly in challenging conditions such as poor lighting, dense crowds, occlusions, and varying cultural or environmental contexts. This limits the generalizability of models when deployed in diverse urban settings.

To address these shortcomings, our proposed work aims to design and implement a lightweight, mobile-optimized AI surveillance system capable of real-time multi-threat detection and integrated alerting. By leveraging advances in compact deep learning models, optimized inference pipelines, and smartphone hardware capabilities, this system will bridge the gap between high-accuracy AI crime detection and portable, affordable deployment.

V. PROPOSED METHODOLOGY

Based on the findings from the literature survey, this work proposes a smartphone-based AI surveillance framework de- signed for real-time, multi-threat detection. The framework emphasizes lightweight model deployment, low latency,

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141053

and integrated alerting mechanisms, making it suitable for environments without fixed CCTV infrastructure.

System Architecture
The proposed system consists of the following core modules:

1) **Video Acquisition**
   - Continuous video capture using the device's high- resolution camera.
   - Target frame rate: 25–30 FPS for smooth analysis.
2) **Preprocessing**
   - Frame resizing and normalization for model input.
   - Background noise reduction and selective frame skipping to optimize computation.
3) **Detection Modules**
   - **Weapon Detection:** YOLOv8-tiny model trained on firearm and knife datasets for rapid detection.
   - **Violence Recognition:** CNN-LSTM hybrid model to capture both spatial and temporal action features.
   - **Face Identification:** Mobile FaceNet for low-latency recognition of persons of interest from a secure local database.
4) **Decision Fusion**
   - Outputs from detection modules combined using a weighted voting mechanism.
   - Threshold-based decision to minimize false alarms.
5) **Alert Generation**
   - Real-time push notifications to security personnel.
   - Optional SMS/email alerts via integrated APIs.
   - Storage of snapshots and short video clips for evidence.

Key Advantages

- Fully portable and deployable in non-CCTV zones.
- Optimized inference pipeline for resource-limited de- vices.
- Integrated multi-threat detection and alerting in a single framework.

VI. EVALUATION CRITERIA AND FUTURE WORK

Although the current stage is focused on the literature survey and framework design, the implementation phase will follow a defined evaluation plan to ensure real-world readiness.

Datasets

- **Violence Detection:** RWF-2000, Hockey Fight Dataset.
- **Weapon Detection:** DISARM Dataset, Open Images Dataset V7 (filtered).
- **Face Recognition:** LFW dataset (aligned subset) and custom local dataset of offenders.

Performance Metrics

- **Accuracy / Precision / Recall:** For classification and detection reliability.
- **F1-score:** Balanced measure for imbalanced datasets.
- **Mean Average Precision (mAP):** For evaluating object detection accuracy.
- **Inference Latency:** Frame processing time with target

≤50 ms on mid-range smartphones.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141053

TABLE I
COMPARISON OF RELATED WORKS IN CRIME DETECTION AND SURVEILLANCE

Reference	**Methodology / Models**	**Crime Types Covered**	**Real-Time**	**Mobile /Edge-Friendly**	**Alert Mech-anism**
Venkatesh et al. [1]	Early-Stopping MIL + Optical Flow on Edge Devices	Violence (physical fights)	Yes	Yes (edge devices)	No
Gao et al. [2]	YOLOv5/6/7 + Faster R-CNN	Arson, vandalism, burglary	Yes	No	Yes (Twilio)
Thakur et al. [3]	YOLOv8 Object Detection	Firearms, knives	Yes	Partial	No
De Paula et al. [4]	Weakly-supervised models (RADS, WSAL, RTFM)	Robbery	No	No	No
Vijeikis et al. [5]	U-Net + LSTM	Violence	Yes	Yes	No
Akdag et al. [6]	Swin Transformer (TeG)	Short + long-term anomalies	Yes	No	No
Ullah et al. [7]	Survey of NN/RNN/Hybrid Mod- els	Violence detection (review)	N/A	N/A	N/A
Negre et al. [8]	Review of deep learning methods	Violence detection (review)	N/A	N/A	N/A
Song & Nang [9]	Edge–server hybrid processing	Pedestrian anomalies	Yes	Partial	No
Mukto et al. [10]	YOLOv5 + MobileNetV2 + LBPH	Weapons, violence, faces	Yes	No	Yes
Li et al. [11]	Visual Anomaly Detection Survey	Anomalies (review)	N/A	N/A	N/A
Sharma et al. [12]	Quantum–AI + SCAD dataset	Weapons	Yes	No	No
Khan et al. [13]	Smart city surveillance survey	Public safety (review)	N/A	N/A	N/A
Ahmed et al. [14]	DISARM dataset + detection pipeline	Weapons	Yes	No	No
Kim et al. [15]	YOLOv4 + KCF + 3D-ResNet	Pedestrian abnormal behavior	Yes	No	No
Lee et al. [16]	Argus++ cube proposals	General activity detection	Yes	No	No
Ali et al. [17]	Smart city surveillance survey	Various crimes (review)	N/A	N/A	N/A
Patel et al. [18]	Anomaly detection review	Anomalies (review)	N/A	N/A	N/A
Zhang et al. [19]	Haar cascades → lightweight CNN	Person detection	Yes	Yes	No
Wang et al. [20]	MSD-CNN multi-stream	Weapons, abnormal objects	Yes	Yes	No
Omar et al. [21]	Video surveillance review	Multiple (review)	N/A	N/A	N/A
Fernandez et al. [22]	Surveillance review	Multiple (review)	N/A	N/A	N/A
Liang et al. [23]	Event detection review	Events in surveillance	N/A	N/A	N/A
Kumar et al. [24]	Visual anomaly detection review	Anomalies (review)	N/A	N/A	N/A
Park et al. [25]	YOLOv4 + KCF	Pedestrian abnormal behavior	Yes	No	No

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025

DOI: 10.17148/IJARCCE.2025.141053

Testing Scenarios
- Indoor environments with staged violence and weapon displays.
- Outdoor public spaces under varying lighting and crowd densities.
- Low-network or offline mode testing for local alert capa- bilities.

Future Enhancements
- Integration with municipal IoT surveillance networks.
- On-device continual learning for improving detection over time.
- Multi-camera coordination for event corroboration.

VII. CONCLUSION

This literature survey provides a comprehensive review of 25 key works in AI-based surveillance, encompassing approaches for violence detection, weapon recognition, anomaly analysis, and integrated monitoring systems. The analysis reveals clear research trends toward deep learning-based object detection, multimodal data fusion, and real-time edge inference, but also identifies notable gaps in portability, multi-threat coverage, and rapid response mechanisms.Our findings underscore the need for a unified framework that not only achieves high detection accuracy but also ensures deployment feasibility in real-world environments without extensive infrastructure. The proposed direction—developing a smartphone-based, lightweight, multi- threat detection system with built-in alerting—directly re- sponds to these identified gaps. Such a system has the potential to democratize access to intelligent surveillance, enabling scal- able, responsive, and cost-effective public safety solutions in both developed and resource-limited settings.By bridging the gap between AI research and practical deployment, this work aims to contribute toward the realization of smart, connected, and safer cities.

REFERENCES

[1]. Venkatesh, A., et al., ”Real-time violence detection on edge devices using Early-Stopping MIL and optical flow features,” *Proceedings of the International Conference on Computer Vision Systems*, 2022.
[2]. Gao, Y., et al., ”Multi-model YOLO and Faster R-CNN pipeline for property crime detection with real-time alerts,” *IEEE Access*, 2024.
[3]. Thakur, R., et al., ”Real-time detection of knives and firearms using YOLOv8,” *Journal of Visual Communication and Image Representation*, 2023.
[4]. De Paula, A., et al., ”CamNuvem: A real-world robbery video dataset for anomaly detection,” *Pattern Recognition Letters*, 2022.
[5]. Vijeikis, R., et al., ”Lightweight violence detection on embedded devices using U-Net and LSTM,” *IEEE International Conference on Image Processing*, 2021.
[6]. Akdag, E., et al., ”TeG: Multi-scale temporal granularity transformer for anomaly detection,” *Expert Systems with Applications*, 2023.
[7]. Ullah, H., et al., ”A survey on vision-based violence detection in video,”*Artificial Intelligence Review*, 2021.
[8]. Negre, A., et al., ”Deep learning for violence detection in video: A review,” *Pattern Analysis and Applications*, 2020.
[9]. Song, J., and Nang, C., ”Edge–server collaborative pedestrian anomaly detection in surveillance video,” *IEEE Internet of Things Journal*, 2021.
[10]. Mukto, M., et al., ”A hybrid deep learning approach for crime moni- toring with weapon, violence, and face recognition,” *Applied Sciences*, 2023.
[11]. Li, Z., et al., ”A survey on visual anomaly detection: Challenges, approaches, and prospects,” *ACM Computing Surveys*, 2021.
[12]. Sharma, P., et al., ”Quantum–AI empowered intelligent surveillance with SCAD dataset,” *Future Generation Computer Systems*, 2023.
[13]. Khan, M., et al., ”Video surveillance systems in smart cities: A survey,”*Electronics*, 2022.
[14]. Ahmed, S., et al., ”DISARM: A CCTV weapon detection dataset and improved detection pipeline,” *IEEE Transactions on Information Forensics and Security*, 2024.
[15]. Kim, S., et al., ”Real-time pedestrian abnormal behavior detection using YOLOv4, KCF, and 3D-ResNet,” *Sensors*, 2021.
[16]. Lee, D., et al., ”Argus++: Robust cube proposals for activity detection in untrimmed video,” *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2020.
[17]. Ali, T., et al., ”A survey of video surveillance systems for smart cities,” *Sustainable Cities and Society*, 2021.
[18]. Patel, R., et al., ”Deep learning-based anomaly detection in video surveillance: A review,” *Image and*

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 10, October 2025
DOI: 10.17148/IJARCCE.2025.141053

Vision Computing, 2022.
[19]. Zhang, Q., et al., "From Haar cascades to lightweight CNNs: A path to efficient surveillance detection," *Pattern Recognition*, 2020.
[20]. Wang, L., et al., "MSD-CNN: Multi-stream lightweight CNN for abnor- mal object detection," *Neurocomputing*, 2022.
[21]. Omar, N., et al., "A comprehensive survey on video surveillance: Applications and trends," *Multimedia Tools and Applications*, 2021.
[22]. Fernandez, J., et al., "A horizontal review on video surveillance for smart cities," *IEEE Access*, 2022.
[23]. Liang, Y., et al., "Event detection in surveillance videos: A review," *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.
[24]. Kumar, A., et al., "A survey on visual anomaly detection: Challenges, approaches, and prospects," *Pattern Recognition*, 2023.
[25]. Park, J., et al., "Real-time pedestrian abnormal behavior detection with KCF and YOLOv4," *Applied Intelligence*, 2021.