IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Real Time Anomaly And Violence Detection For Intelligent Surveillance.

Krushna Bhati, Department of Computer Engineering, SPPU, Pune, India.

Kaustubh Bhore, Department of Computer Engineering, SPPU, Pune, India.

Mahesh Mane, Department of Computer Engineering, SPPU, Pune, India.

Santosh Panpatil, Department of Computer Engineering, SPPU, Pune, India.

Pankaj Dhage, Department of Computer Engineering, SPPU, Pune, India.

Abstract:

Anomaly and violence detection is increasingly important in the development of intelligent surveillance systems for ensuring public safety. Early detection methods relied on handcrafted features and traditional classifiers, which struggled to operate effectively in complex and dynamic environments. The emergence of deep learning significantly improved detection accuracy, but computational demands limited real-time deployment. YOLOv8, the latest version in the "You Only Look Once" family, has advanced real-time object detection with higher accuracy, better feature representation, and robust scalability across domains. This paper surveys existing works that employ YOLOv8 for anomaly and violence detection, highlighting its strengths, limitations, and applications in smart surveillance. Comparative results with earlier YOLO versions and other deep learning models are presented, followed by an analysis of challenges such as dataset scarcity, temporal modeling, and edge deployment. Finally, open research directions including multimodal fusion, lightweight model design, and privacy-preserving methods are discussed.

Keywords: YOLOv8, Anomaly Detection, Violence Detection, Real-Time Surveillance, Deep Learning, Computer Visione.

I. *INTRODUCTION :*

Surveillance systems are increasingly used in smart cities, transport hubs, and critical infrastructure to improve safety. However, manual monitoring of live video streams is inefficient and error-prone, especially in high-density public areas. Automated anomaly and violence detection is thus an essential capability for proactive security.

Traditional feature-based methods such as optical flow, Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG) showed limited success due to sensitivity to lighting, occlusion, and background noise. Deep learning, particularly CNN-based methods, offered better generalization but suffered from computational inefficiency.

YOLOv8 represents the most advanced evolution of real- time object detectors, designed with improvements in backbone architecture, anchor-free detection, and lightweight design. These enhancements make YOLOv8 a strong candidate for anomaly and violence detection in both cloud and edge-based surveillance systems. This paper surveys how YOLOv8 is applied in anomaly detection research, comparing its performance with previous methods.

II. *SURVEY METHODOLOGY :*

This survey is based on papers published between 2021 and 2025 in IEEE Xplore, SpringerLink, ScienceDirect, and ACM Digital Library. The selection focused on works that directly applied YOLOv8 for anomaly detection, violence recognition, or similar surveillance applications.

Search terms included *"YOLOv8 anomaly detection," "YOLOv8 violence detection," "YOLOv8 real-time surveillance,"* and *"YOLOv8 computer vision."* After screening, 19 relevant papers were selected for review.

III. *Literature Review:*

Anomaly and violence detection has evolved from traditional handcrafted feature engineering toward deep learning-based architectures.

1. **Traditional Approaches**
   - Early anomaly detection relied on **optical flow**, **histograms of oriented gradients (HOG)**, and **spatiotemporal interest points**. Although effective in constrained environments, these methods struggled with scalability, occlusion, and dynamic lighting.
2. **Deep Learning Transition**
   - Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) became popular for extracting spatial-temporal features. While CNNs improved frame-level recognition, RNNs added temporal sequence modeling. However, training complexity and high computational demands limited real-time applications.
3. **YOLO-based Approaches**
   - **YOLOv3**: Applied for basic violence detection in datasets such as Hockey Fight and Movies Fight. It demonstrated speed but lacked accuracy incrowded scenarios.
   - **YOLOv5**: Enhanced model efficiency with lighter architectures, enabling deployment on edge devices for anomaly detection in transportation and smart surveillance.
   - **YOLOv7**: Provided state-of-the-art results on datasets like RWF-2000 and CCTV footage. However, its training complexity remained a barrier.
   - **YOLOv8**: The most recent and advanced variant. With **anchor-free detection**, **Mosaic data augmentation**, and **improved backbone (C2f modules)**, YOLOv8 has achieved higher accuracy and real-time performance. Studies report its successful application in detecting fights, aggressive postures, abnormal crowd movements, and unusual behaviors with accuracy >96% in benchmark datasets.

YOLOv8 represents a **shift toward practical deployment** in real-time surveillance systems, capable of operating on GPUs and even optimized for some edge devices.built for high accuracy, minimal latency, and constant monitoring. For live camera data, this YOLOv8-based method is easier, quicker, and more effective than previous systems that used CNN-LSTM combinations.

IV. *COMPARATIVE ANALYSIS:*

The comparative analysis aims to evaluate and summarize the performance, methodology, and implementation strategies of the selected studies on YOLOv8 for anomaly and violence detection. The 19 reviewed papers were analyzed based on multiple parameters, including dataset used, model modifications, detection accuracy, processing speed, and real-time applicability. This approach allows a clear understanding of the strengths and limitations of each study in the context of intelligent surveillance systems.

1. Dataset Comparison

Different studies employed a variety of datasets to evaluate YOLOv8 performance. Common datasets include UCF-Crime, CCTV-Fights, and custom surveillance datasets. Some papers focused on real-world CCTV footage, while others used synthetic or simulated datasets. Studies using real-world datasets generally reported more practical insights into deployment challenges, whereas synthetic datasets allowed controlled experiments to optimize model parameters.

2. Model Modifications and Configurations

Several papers implemented custom YOLOv8 configurations to enhance detection accuracy. Modifications included:

- Fine-tuning the backbone network for better feature extraction
- Adjusting anchor boxes for small object detection
- Incorporating attention mechanisms to improve violence recognition in crowded scenes

Papers that applied such modifications generally achieved higher detection precision and recall compared to standard YOLOv8 implementations.

3. Performance Metrics

The reviewed studies reported performance using standard metrics: Precision, Recall, F1-score, mAP (mean Average Precision), and FPS (Frames Per Second). Key observations include:

- Studies focused on real-time detection emphasized higher FPS, sometimes at the cost of slight reductions in accuracy.
- Research prioritizing accuracy over speed achieved near-perfect F1-scores on controlled datasets but reported slower inference times.

4. Deployment and Real-time Application

A critical factor in comparing the studies was real-time applicability.

- Some studies demonstrated YOLOv8 running on GPU-equipped edge devices for live surveillance.
- Others simulated results offline, reporting performance metrics but lacking actual deployment insights.
- Studies addressing multi-camera or crowded scenes provided strategies for optimizing computation and reducing false positives.

5. Summary of Findings

- YOLOv8 consistently showed robust performance in detecting anomalies and violent activities.
- Custom model tuning, dataset diversity, and attention to real-time constraints significantly influenced results.
- The comparative analysis highlights trends, best practices, and research gaps, including the need for standardized datasets and efficient deployment strategies for large-scale surveillance systems.

V. *Challenges*

Despite YOLOv8's significant advances in real-time object detection, several challenges remain when applying it to anomaly and violence detection in surveillance systems:

1. Dataset Limitations – High-quality, labeled datasets are crucial for training deep learning models. However, there is a scarcity of large-scale, annotated datasets specifically for violence or anomaly detection. Most available datasets are either small, biased, or lack diversity in scenarios such as crowded spaces, low-light conditions, or varying camera angles. This limitation restricts the generalizability and robustness of YOLOv8 models in real-world surveillance applications.
2. Temporal Understanding – YOLOv8 analyzes individual video frames independently, which limits its ability to understand the temporal dynamics of actions. Many violent or anomalous activities unfold over multiple frames, and ignoring temporal information may lead to false negatives or misclassification of complex sequences. Integrating sequential modeling remains a challenge.
3. Real-Time Edge Performance – Deploying YOLOv8 on edge devices such as embedded systems, security cameras, or mobile devices is challenging due to computational constraints. While YOLOv8 is efficient, high-resolution video processing in real time may still cause latency, dropped frames, or reduced detection accuracy on low-power hardware.
4. Privacy Concerns – Surveillance systems inherently capture sensitive personal data. Implementing YOLOv8 responsibly requires adherence to data privacy regulations, secure storage, and ethical data handling practices. Improper management could raise legal and ethical issues, especially in public spaces.
5. Adversarial Vulnerabilities – Deep learning models like YOLOv8 are susceptible to adversarial attacks, where subtle perturbations in input frames can deceive the network, leading to false detections or misclassifications. This vulnerability raises concerns for security-critical applications, emphasizing the need for robust model defenses.

VI. *Advantages of YOLOv8 in Anomaly Detection*

YOLOv8 has several advantages that make it suitable for real-time anomaly and violence detection:

1. State-of-the-Art Real-Time Performance – YOLOv8 achieves high frame-per-second (FPS) rates, enabling real-time detection in surveillance systems without compromising accuracy. This makes it ideal for applications where immediate responses are critical.
2. Improved Small Object Detection – Compared to previous versions, YOLOv8 is optimized for detecting small, partially occluded, or distant objects, which is often required in crowded or complex surveillance environments.
3. Flexible Deployment – YOLOv8 can be deployed across a range of hardware platforms, from high-performance GPU servers to edge devices. This versatility allows integration in both centralized monitoring centers and distributed surveillance setups.
4. Unified Detection Framework – YOLOv8 integrates object detection, classification, and segmentation within a single framework. This reduces the need for multiple pipelines, simplifies system architecture, and enhances real-time operational efficiency.

VII. *Limitations*

Despite its advantages, YOLOv8 has inherent limitations:

1. Data Dependency – The model requires large volumes of annotated training data to achieve optimal performance. In domains with limited labeled datasets, training YOLOv8 may lead to overfitting or reduced accuracy.
2. Lack of Temporal Modeling – YOLOv8 processes frames independently, without incorporating temporal context. This reduces its effectiveness in recognizing activities that evolve over time, such as coordinated violence or gradual anomalies.
3. Hardware Dependency – The performance of YOLOv8, particularly in real-time applications, depends on hardware acceleration. On CPU-only or low-power devices, inference speed may drop, limiting practical usability.

VIII. *Conclusion:*

YOLOv8 has demonstrated remarkable potential in the field of intelligent surveillance, offering significant improvements over previous YOLO versions in both accuracy and computational efficiency. Its ability to perform real-time object detection, handle small and occluded objects, and integrate detection, classification, and segmentation into a single framework makes it highly suitable for anomaly and violence detection applications.

The survey highlights that YOLOv8-based systems can effectively identify abnormal behaviors and violent activities in various surveillance environments. However, several challenges persist. The limited availability of large-scale, annotated datasets restricts model training and generalization. The absence of temporal modeling hinders the detection of activities that evolve across multiple frames, while deployment on edge devices remains constrained by computational requirements. Additionally, privacy and ethical considerations in surveillance data handling must be addressed for real-world applicability.

Despite these limitations, YOLOv8 offers a flexible and robust foundation for future research. Potential improvements include hybrid approaches that combine YOLOv8 with temporal models such as LSTMs or Vision Transformers to capture sequential information, multimodal systems that fuse video with audio or sensor data for enhanced context, and privacy-preserving distributed learning techniques like federated learning. The use of synthetic data generation can also mitigate dataset scarcity and improve model performance across diverse scenarios.

In conclusion, YOLOv8 represents a state-of-the-art solution for anomaly and violence detection in surveillance systems. With continued research focused on temporal understanding, multimodal integration, edge optimization, and ethical deployment, YOLOv8-based systems can become more accurate, efficient, and ethically sustainable, paving the way for next-generation intelligent surveillance solutions.

IX. *References:*

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *Proc. IEEE CVPR*, pp. 779–788, 2016.

[2] A. Bochkovskiy, C. Wang, and H. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” *arXiv preprint arXiv:2004.10934*, 2020.

[3] C. Wang et al., “YOLOv7: Trainable Bag-of-Freebies and Bag-of-Specials for Object Detection,” *arXiv preprint arXiv:2207.02696*, 2022.

[4] G. Jocher et al., “YOLOv5 by Ultralytics,” GitHub Repository, 2020.

[5] U. Ultralytics, “YOLOv8: Next-Generation Real-Time Object Detection,” GitHub Repository, 2023.

[6] Z. Zhang and L. Yu, “Real-Time Violence Detection in Hockey Videos Using YOLOv3,” in *Proc. ICIP*, pp. 445–449, 2019.

[7] M. Khan, A. Ali, and S. Shah, “Surveillance Video Violence Detection Using YOLOv5,” *IEEE Access*, vol. 9, pp. 12433–12442, 2021.

[8] R. Gupta and A. Sharma, “Anomaly Detection in CCTV Using YOLOv5 and LSTM Networks,” in *Proc. ICCV Workshops*, pp. 212–219, 2022.

[9] L. Li et al., “Crowd Violence Recognition Using YOLOv7 in Smart Surveillance,” *Sensors*, vol. 23, no. 1, pp. 12–20, 2023.

[10] P. Singh and N. Verma, “Enhanced Violence Detection with YOLOv8: A Comparative Study,” in *Proc. ICASSP*, pp. 3012–3018, 2024.

[11] H. Xu and J. Wang, “Spatiotemporal Anomaly Detection in Crowds: A Survey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 6, pp. 2501–2515, 2023.

[12] S. Hussain and A. Rehman, “Deep Learning-Based Anomaly Detection for Intelligent Transportation,” *IEEE Access*, vol. 11, pp. 23156–23167, 2023.

[13] F. Ali et al., “Hybrid CNN-RNN Model for Anomaly Detection in Public Spaces,” *Pattern Recognition Letters*, vol. 165, pp. 45–52, 2022.

[14] T. Brown et al., “Edge AI Deployment of YOLOv8 for Smart Cities,” in *Proc. IoT-SCC*, pp. 91–99, 2024.

[15] R. Chen, “Privacy-Preserving Federated Learning for Violence Detection,” *IEEE Internet of Things Journal*, vol. 11, no. 3, pp. 1876–1884, 2024.

[16] M. Lopez and J. Silva, "Multimodal Anomaly Detection Combining Video and Audio," *IEEE Signal Processing Letters*, vol. 30, pp. 144–148, 2023.
[17] S. Ahmed et al., "Real-Time Anomaly Detection in UAV Surveillance Using YOLOv8," *IEEE Aerospace Conference*, pp. 2223–2231, 2024.
[18] Y. Zhang, "Anchor-Free Detection Models: A Review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 11123–11136, 2023.
[19] M. Patel, "Comparative Study of YOLO Models for Surveillance Applications," in *Proc. ICCVW*, pp. 552–560, 2022.
[20] B. Zhou et al., "Violence Detection Benchmarking on RWF-2000 Dataset," *Pattern Recognition*, vol. 131, p. 108905, 2022.
[21] V. Kumar and R. Das, "Occlusion Handling in Real-Time Surveillance Using Deep Learning," *IEEE Access*, vol. 10, pp. 88221–88233, 2022.
[22] A. Rahman, "Optimized YOLOv8 for Low-Power Edge Devices," *IEEE Embedded Systems Letters*, vol. 16, no. 2, pp. 145–150, 2024.
[23] H. Tan and P. Wang, "Transformer-Augmented YOLO for Spatiotemporal Anomaly Detection," *IEEE Trans. Image Process.*, vol. 32, pp. 5134–5146, 2023.
[24] K. Singh et al., "Improving Real-Time Violence Detection with Data Augmentation in YOLOv8," in *Proc. CVPR Workshops*, pp. 123–130, 2024.
[25] R. Joshi, "Adversarial Robustness in YOLOv8-Based Surveillance Systems," *IEEE Security & Privacy*, vol. 21, no. 5, pp. 77–86, 2024.