Crowd Management and Weapon Detection Using Existing CCTV Network

Greeshma Reddy, Mohammad Travadi, Vignesh G, Adithya R
School of Computer Science and Engineering
Presidency University
Bengaluru, India
Email: (mohammad.travadi, greeshma.reddy, adithya.r, vignesh.g) @presidencyuniversity.in

***Abstract*—Ensuring public safety in crowded urban environ- ments presents major challenges for traditional surveillance systems that rely on manual monitoring of CCTV footage. Such systems are often prone to delayed threat detection, human error, and limited scalability. This paper proposes a modular, AI- powered Security Management Suite designed to augment exist- ing CCTV infrastructure by enabling real-time crowd monitoring and weapon detection. The system leverages advanced object detection models, YOLOv4 and YOLOv8, integrated within a scalable Python-based GUI to automate surveillance operations. Key features include multithreaded processing, live event logging, multilingual alerts, and minimal hardware requirements. Experi- mental results demonstrate high detection accuracy, rapid threat recognition, and reliable multi-stream processing. The proposed solution offers a significant step forward in creating proactive, intelligent public safety systems that can adapt to evolving urban security challenges.**

***Keywords*—Crowd Management, CCTV Surveillance, YOLOv4, YOLOv8, Weapon Detection, Real-Time Monitoring, Computer Vision, Public Safety**

I. INTRODUCTION

The ubiquitous presence of CCTV cameras in urban settings forms a backbone for modern surveillance infrastructure. As cities grow and evolve, the accumulation of visual data from these networks holds immense potential for maintaining public safety. Nevertheless, the reliance on manual monitoring of CCTV footage poses significant challenges. Human operators are vulnerable to fatigue and oversight, which can lead to delayed threat detection and compromised security outcomes [1]

Amidst escalating urban population densities and the increasing sophistication of potential threats, including terrorist activities, mass gatherings, and public unrest, there is an urgent call for evolution in surveillance methodologies [2]

Traditional systems fall short in providing real-time alerts and comprehensive coverage due to the sheer volume of footage needing analysis. AI technologies, particularly advances in real-time object detection and deep learning, present an opportunity to redefine surveillance systems by automating detection and monitoring tasks [3].

Our work introduces an Artificial Intelligence-powered surveillance framework designed to enhance existing CCTV networks by empowering them with automation capabilities for real-time crowd and weapon detection. This system utilizes advanced YOLOv4 and YOLOv8 models integrated within a centralized GUI, thereby offering a cost-effective, scalable solution that can adapt to a variety of urban security demands. By transforming passive, hardware-bound infrastructure into smart, interconnected networks, this approach not only elevates immediate threat detection protocols but also strategically supports the development of long-term urban safety policies [1].

II. LITERATURE REVIEW

In recent years, computer vision and machine learning have made major strides, especially when it comes to surveillance. Not too long ago, most surveillance systems were completely manual—heavily dependent on human judgment and constant attention. That setup worked to a point, but it just couldn't scale or keep up with increasing demands [4]

The introduction of artificial intelligence and deep learning has changed the game. These technologies now make it possible to automatically detect objects, suspicious activities, and unusual behavior with much greater efficiency [1], [5]. Among the most influential tools in this space are the YOLO (You Only Look Once) models, known for their ability to recognize objects quickly without sacrificing accuracy [6] [8].

Earlier versions like YOLOv2 and YOLOv3 helped set the stage for real-time detection and showed how useful these models could be in real-world scenarios. Building on those foundations, researchers have fine-tuned YOLO for more specific tasks—weapon detection being a key one, given today's security concerns [2].

More recent work with YOLOv4 has highlighted how well it balances speed and precision, even in complex urban settings [9]. Meanwhile, YOLOv8 introduces anchor-free detection and better feature aggregation, which makes it especially effective in crowded or visually cluttered environments [3]. These improvements point to how crucial AI is in moving beyond passive surveillance, turning security systems into active tools that can respond to threats as they happen [7]

Thanks to smarter model designs and better processing power, surveillance technology is becoming faster and more accurate. By combining traditional CCTV setups with real-time analysis, we're moving toward a more predictive approach to security—one that focuses on preventing threats before they escalate [1].This shift is paving the way for intelligent systems that are better equipped to handle the growing demands of urban safety.

III. Methodology

This section details the approaches used in developing an AI-driven surveillance system using existing CCTV infrastruc- ture to provide real-time threat recognition and response [4]. The system's design is characterized by a combination of auto- mated machine learning techniques and user-centric interface applications, ensuring that real-time monitoring requirements are met efficiently and effectively.

A. System Overview

Our proposed system takes traditional CCTV monitoring to the next level by adding real-time, AI-powered threat detection capabilities [5]. It works by pulling video feeds from existing CCTV cameras using the Real-Time Streaming Protocol (RTSP). These live streams are then continuously analyzed using advanced object detection models—YOLOv4 and YOLOv8—which are equipped to measure crowd density and detect weapons with high accuracy [2]

The results are displayed through a user-friendly Python-based graphical interface, giving operators an intuitive way to manage surveillance tasks [8]. The system's modular design makes it highly scalable—it can easily handle more camera inputs or detection features without interrupting the current workflow [7]

One of the key advantages of this setup is that it requires minimal hardware changes, which makes it a cost-effective solution for upgrading existing surveillance systems with AI capabilities [1]. By blending AI seamlessly into current CCTV networks, we reduce resource waste and improve the monitoring of large, constantly changing urban areas. The system is built to scale with growing cities, staying responsive to evolving security needs [3]. Its ease of integration and flexible functionality not only extend the usefulness of existing infrastructure but also significantly boost overall security performance..

B. Data Preparation

The performance of any AI model depends heavily on the quality and diversity of the data it learns from. For our system, we used carefully curated datasets focused on both crowd and weapon detection. These include well-known sources like the ShanghaiTech Crowd Dataset and the Open Images Dataset V6 [8]]. Each dataset was thoroughly annotated, with bounding boxes marking people and weapons to help the models learn more effectively [4].

To make the system more adaptable to real-world scenarios, we applied various data augmentation techniques such as rotation, scaling, flipping, and adjusting contrast. These methods simulate different lighting and environmental conditions, helping the model generalize better [9]We also standardized all input images—resizing them to formats like 416×416 or 640×640 pixels—so they align with the input requirements of the YOLO models [7]. This detailed preparation ensures the system stays accurate even in complex or unpredictable settings [5]

Ultimately, the way data is annotated and augmented has a big impact on how well the model performs in live conditions. Urban surveillance environments are often chaotic and unpredictable, so maintaining high data quality during training helps ensure that the system operates reliably in the field [6]. Thanks to these robust preprocessing steps, our system can handle complicated visual scenes with confidence, delivering accurate results when it matters most [1].

C. Model Deployment

To achieve real-time responsiveness and dependable accuracy, our system relies on specialized deep learning models tailored to the unique demands of urban surveillance [2]

Crowd Detection: We use YOLOv8 to identify and count individuals within each video frame, allowing for accurate assessments of crowd density [4]. Its enhanced feature aggregation makes it especially effective in dealing with partially obstructed views and densely populated scenes [8]

Weapon Detection: For spotting weapons such as firearms and knives, YOLOv4 is employed. This model strikes a strong balance between speed and accuracy, making it well-suited for real-time weapon detection in various environmental conditions [3], [5]

Both models were refined using transfer learning techniques—leveraging pre-trained weights to better align with our specific datasets [1]. To further optimize performance, we implemented TensorRT acceleration, ensuring fast inference times on GPU-powered systems [6].The system displays detections using bounding boxes and confidence scores, and when these scores exceed predefined thresholds, alerts are triggered in real time and logged through the GUI interface [9]

Choosing the right models and fine-tuning them for speed and precision reflects the real-world urgency of urban monitoring needs [7]. Our approach addresses both the demand for rapid detection and the need for accurate identification, enabling immediate threat notifications and faster interventions. This thoughtful integration of AI brings traditional surveillance infrastructure closer to becoming an intelligent, responsive safety net for the public [4], [8]

IV. System Architecture

The proposed surveillance system employs a modular, scal- able architecture that enhances detection precision, data man- agement, and user interaction [2]. The system's architecture builds upon best practices in system design by integrating novel AI technologies with robust infrastructural elements, achieving effective results [7]. The architecture comprises five core layers, each serving a distinct and essential function:

A. Input Layer

The input layer captures real-time video streams from exist- ing CCTV cameras using RTSP or USB interfaces, supporting both standard-definition and high-definition feeds for adapt- ability across multiple surveillance setups . Preprocessing techniques, such as frame resizing and noise filtering, areutilized to preserve input quality for subsequent AI processing
[6] .

The diversity of input compatibility speaks to the system's flexibility in integrating with the array of available surveillance tools . Through standardizing input feed dimensions and a meliorating noise, this layer serves as the initial catalyst for the detection processes that follow, enabling precise AI inference from diverse source conditions [8]. This preparatory approach ensures the system's resilience and readiness to accommodate evolving surveillance technologies, effectively future-proofing the architecture [3].

B. AI Processing Layer

At the heart of the system, the AI processing layer utilizes YOLO-based models to perform rapid, frame-by-frame analysis of incoming video streams [5]. This layer leverages:

- **Crowd Detection:** : Leveraging YOLOv8 for individual detection and counting amidst video frames enables real- time crowd density assessment [2]
- **Weapon Detection:** Utilizing YOLOv4, the system iden- tifies weapons such as firearms and knives with high precision [4]

Optimized for low-latency inference, this layer effectively processes video streams with minimal delay, making it the cen- trepiece of real-time threat detection and surveillance system reliability [9]. Advanced AI architectures, such as these, ensure not only detection accuracy but also operational efficiency necessary to support uninterrupted comprehensive surveillance needs. This architecture supports continuous innovation and adaptation of newer, superior detection methodologies into the existing framework [6]

C. Control Layer

The control layer coordinates various subsystems using multithreaded processing architectures to ensure independent yet synchronized execution of video acquisition, AI inference, alert generation, logging, and GUI updates [3]Through mul- tithreading, this design diminishes bottlenecks, prevents GUI freezing, and capacitates the concurrent handling of multiple video feeds [8].

By orchestrating task execution across threads, the control layer maintains operational fluidity and efficiency, crucial for processing relentless streams of information inherent in urban surveillance networks [7]. Consequently, security personnel gain timely access to a continuous flow of actionable intel- ligence, free of any operational backlog or data latency [5]. This strategic integration of system functions provides a solid foundation for expanding operational scopes and enhancing functional versatility [9].

D. GUI Layer

Developed with Python's Tkinter and CustomTkinter li- braries, the graphical user interface operates as the central nerve of system control and observation [8] . It offers live visualization of detection results, displays event logs, sup- ports multilingual configurations, and enables manual ad- justments of detection modules [6]. The real-time feedback features—encompassing alert popups and screenshot view- ers—heighten situational awareness for operators [5].

A well-designed GUI is indispensable for transforming machine-generated insights into accessible information inter- faces [2]. By offering enriched visual feedback and operational control, this layer promotes proactive threat management, equipping operators with up-to-the-minute information that guides informed response decisions essential for effective surveillance [3]. Dynamic interface features foster user en- gagement and prompt efficient operational maneuvers, suc- cessfully bridging computational outputs with human inter- pretative skills [7].

E. Notification Layer

The notification layer manages the distribution of alerts to both local and remote endpoints [1] . When anomalies such as overcrowding or the presence of weapons are detected, the system activates instant visual alerts within the GUI and optionally dispatches notifications through external APIs, email services, or push notification platforms [5]. This ensures critical information is rapidly communicated to the relevant personnel, awareness extending beyond the confines of the control room [8].

The effectiveness of alert dissemination mechanisms lies in their speed and versatility [6]. By facilitating immediate notification pathways, irrespective of geographical barriers, the system ensures rapid dissemination of critical event alerts, thereby enhancing the response capability of security teams actively working to neutralize threats before they escalate [4]. This comprehensive, layered disposition ensures contin- uous situational alertness, accommodating rapid operational demands and upholding system integrity [9].

V. Results and Discussion

Our Our testing showed that the AI-powered surveillance system significantly improves detection accuracy and response times when compared to traditional monitoring approaches. The combined use of YOLOv8 and YOLOv4 models proved particularly effective—delivering strong performance even in complex scenarios involving poor lighting or dense crowds [7].

Scalability was another key strength. The system maintained stable performance while handling multiple video streams simultaneously, demonstrating its readiness for deployment in larger, more demanding environments [3]. Feedback from users interacting with the Python-based GUI also highlighted ease of use and overall satisfaction, validating the system's intuitive design and operational effectiveness [6].

By integrating real-time analytics and automated threat detection, the project signals a shift toward proactive surveillance—one capable of addressing threats before they escalate [7]. Looking ahead, future iterations could expand the system's capabilities to detect unattended objects, abnormal behavior, or unauthorized access, further increasing its value in diverse security contexts [8]. Additionally, incorporating self-correcting machine learning models may lead to an even more adaptive and efficient system—laying the foundation for a resilient, future-proof solution for public safety [4]

CONCLUSION

In this paper, we presented the design and deployment of an AI-driven surveillance solution aimed at bolstering public safety via real-time crowd control and weapon detection, using existing CCTV networks [2]. By incorporating YOLOv8 and YOLOv4 deep learning models into a modular architecture, our system overcomes traditional manual monitoring limitations [5].

Experimental findings demonstrated exceptional detection accuracy, expedient alert generation, and resilient performance, even during concurrent multi-stream processing [6]. The system's multilingual GUI positively impacts operational usability, granting security operators the efficacy to address threats decisively [3].

Future system enhancements could see: the integration of predictive analytics for early congestion trend detection, adaptive alert sensitivity tuned to environmental conditions, and the expansion of detection capabilities to include unattended objects, abnormal motion patterns, and unauthorized area intrusions [1]. These advancements lay the foundation for safer, smarter urban environments driven by AI-powered automation.[8].

REFERENCES

[1] S. Patil, "Weapon Detection Using YOLO V3 for Smart Surveillance System," 2021. [Online]. Available: https://www.researchgate.net/publication/351538173_Weapon_Detection_Using_YOLO_V3_for_Smart_Surveillance_System.

[2] A. Chandra and P. Sharma, "Weapon detection system for surveillance and security," 2021. [Online]. Available: https://www.researchgate.net/publication/370148091_Weapon_detection_system_for_surveillance_and_security.

[3] S. Mehta, "Weapon Detection Using YOLOv4 CNN," 2021. [Online]. Available: https://www.ijraset.com/research-paper/weapon-detection-using-yolov4-cnn.

[4] H. Kiani and A. Seyfi, "Weapon detection using deep learning and web application implementation," Heliyon, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S240584402410953X.

[5] R. Harisha, "Smart Surveillance System using YOLOv4 and Deep Learning," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10616365.

[6] S. Murthy and A. P. Verma, "AI-enabled surveillance: Efficiency and applications," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10910535.

[7] K. Nair, "Real-time analytics for security systems enhancement," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/9057329.

[8] J. K. Singh, "AI methodologies in modern surveillance," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10099733.

[9] A. Khandelwal, "Advanced threat detection in dynamic environments," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10574710.