International Student Conference on Next-Gen Computing: Application of AI, Big Data, Quantum Computing, Signal Processing and Cloud Innovations (ICNGC-2026)

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

Accident and Crime Detection from Surveillance Video Simulation Using YOLO

Mr. Kunal Kanchankar[1], Dr. Ravindra Kale[2]

[1]Student of M. Tech Final Year Department of computer Science Engineering G H Raisoni Skill Tech University, Nagpur, Maharashtra

[2]Assistant Professor & Guide, Department of computer Science Engineering, G H Raisoni Skill Tech University, Nagpur, Maharashtra

Abstract- The quick growth of the cities and smart cities made the need for surveillance systems that can automatically spot incidents and make the public safer even greater. People are in charge of watching video feeds in the traditional surveillance systems, which can be slow it and make mistakes when there is a lot of footage to watch. Recently advancements in (AI) artificial intelligence, Computer Vision, and deep learning have facilitated automated-systems in the analysis of real-time video streams, enhancing an accuracy of abnormal activity detection. This review paper looks at current research on the AI-based surveillance systems that are meant to find incidents in cities. It talks about the popular methods like (CNN), (RNN), and real-time object detection models like YOLO. The study also reviews available surveillance datasets, evaluation metrics, and implementation approaches used in previous work. A comparison of existing methods highlights their strengths and limitations. The paper also emphasizes the need for integrated multi-incident detection systems capable of improving safety in modern smart cities.

Keywords: Deep Learning, Smart City Surveillance, Incident Detection, Computer Vision.

I.INTRODUCTION

With the rapid growth of urban populations and the development of smart cities, ensuring public safety has become a major challenge for governments and the law enforcement agencies.

Modern cities deploy thousands of surveillance camera in public place such as roads, railway stations, airports ,shopping malls, and residential areas to monitor activities and maintain security. These surveillance infrastructures generate massive volumes of video data that must be analyze continuously to detect suspicious and dangerous events. However, manually monitoring of such large scale video data is extremely difficult and inefficient. Human operators often experience fatigue and may overlook important incidents occurring in real time, which can delay response and compromise public safety [8][14].

Recent advancements in AI, machine learning, and computer-vision have enabled the development of intelligent surveillance systems capable of automatically analyzing video streams and detecting incidents without continuous human supervision. Deep learning techniques, particularly (CNNs), have demonstrated amazing performance in photo and video analysis job such as object detection, activity recognition, and anomaly detection. These technologies enable automated systems to interpret visual information and identify events such as accidents, theft, suspicious activities, and abnormal crowd behavior in real time [1][4][6].

Among the various computer vision techniques, deep learning based object detection models such as the YOLO framework have become widely used for real-time surveillance analysis. These systems have the ability to identify items like vehicles, pedestrians, and suspicious activities with great velocity and precision, rendering them ideal for applications involving real-time observation. Several studies have successfully applied YOLO-based models for detecting objects and monitoring activities in surveillance environments [1][3].

In addition to object detection, sequential deep learning architectures such as Recurrent Neural Networks, commonly known as RNNs, along with Long Short-Term Memory networks, referred to as LSTMs, have been employed to examine time-related patterns in sequences of video footage. These models can identify abnormal activities or unusual behavioral patterns over time, which is particularly useful for detecting complex events such as crowd disturbances, suspicious movements, or violent behavior [5][10][11].

Many researchers have explored the application of AI-based surveillance systems for specific security applications. For example, intelligent surveillance models have been developed for crime detection, anomaly detection, and automated monitoring in public environments. These systems use machine learning and computer vision techniques to analyze video streams and identify unusual behaviors or potential threats, thereby improving security monitoring efficiency [2][6][12].

Similarly, AI-based surveillance systems have been applied in specialized domains such as banking security, high-security areas, transportation systems, and smart city infrastructures. For instance, automated systems have been developed to detect robbery attempts in banks using intelligent CCTV monitoring, while other research focuses on detecting traffic incidents and hazards in transportation environments using AI-based video analytics [7][13][15].

The IoT technologies with AI-based surveillance has further enhanced the capabilities of intelligent monitoring systems. IoT-enabled smart surveillance frameworks allow multiple cameras and sensors to collect and transmit data in real time, enabling centralized monitoring and intelligent decision making for smart city environments [5][9].

Despite these advancements, many existing surveillance systems focus on detecting only a single type of incident, such as anomaly detection, crowd monitoring, or theft detection. These systems often lack the capability to identify multiple types of incidents simultaneously within a unified framework. Additionally, some approaches require high computational resources, which makes large scale real-time deployment challenging in smart city environments [8][14].

To address these limitations, researchers are increasingly focusing on developing intelligent multi-incident detection systems capable of identifying multiple types of events within a single surveillance framework. Such systems can automatically detect incidents such as accidents, robbery, kidnapping, harassment, suspicious activities, and abnormal crowd behavior from surveillance footage and generate immediate alerts for authorities. These intelligent monitoring solutions can significantly enhance urban safety and improve emergency response times in modern smart city infrastructures [5][9][16].

II. LITERATURE REVIEW

AI and advanced deep learning methods have greatly enhanced the functionality of contemporary monitoring systems. Traditional CCTV monitoring systems depend heavily on human operators, which can lead to delays and errors in detecting suspicious events. Recent research focuses on automated surveillance using computer vision, machine learning and technologies related to the IoT to facilitate immediate observation and recognition of incidents.

Proposed a real-time surveillance system using the YOLO deep learning framework to detect objects in video streams efficiently. The model was trained on surveillance video datasets and optimized to handle problems such as low lighting conditions, occlusion, and crowded environments, achieving high detection speed suitable for real-time security applications [1].

Agarwal. developed an AI-based smart surveillance system that detects anomalies and suspicious activities in monitored environments. Their approach analyzes behavioral patterns using artificial intelligence models to automatically identify abnormal events and improve public safety monitoring [2].

Manu. presented an AI-powered smart surveillance camera capable of detecting unusual activities through computer vision algorithms embedded within the camera system. The system performs local data processing, reducing network load and enabling faster response times in security monitoring environments [3].

Suggested a surveillance system powered by artificial intelligence that combines image analysis and automated technology. learning techniques for automatic monitoring of video feeds. Their system focuses on identifying suspicious behaviors and reducing the dependency on human supervision in security systems [4].

Villegas. explored the role of artificial intelligence in intelligent IoT monitoring systems. Their study highlighted how AI models integrated with IoT devices enable real time analysis, prediction, and monitoring of events in smart city environments [5].

Developed an AI-based surveillance system aimed at crime control and security enhancement. The system analyzes video feeds using machine learning techniques to detect criminal activities and send alerts to authorities for quick response [6].

Introduced an AI-based surveillance model specifically designed for robbery and theft detection in banking environments. The system uses smart CCTV cameras and AI algorithms to analyze suspicious movements and trigger alerts when abnormal activity is detected [7].

Saranya carried out a study on intelligent CCTV monitoring systems that utilize artificial intelligence, emphasizing the application of technologies including face identification, movement detection, and behavior assessment. Their work emphasized the growing importance of automated surveillance for improving monitoring efficiency in large-scale environments [8].

Proposed an AI-enabled Industrial Internet of Things (IIoT) framework for smart city monitoring. The system integrates surveillance cameras and sensors to monitor events such as traffic congestion, accidents, and public gatherings in real time [9].

Qaraqe. developed a secure surveillance framework called PublicVision for recognizing crowd behavior in public spaces. The system securely transmits CCTV video to a central hub where deep learning models analyze crowd size and violence levels using advanced architectures such as Swin Transformer [10][11].

Ahmed and Nyarko introduced Smart-Watcher, a surveillance system utilizing artificial intelligence and Internet of Things technology, specifically crafted for small and medium-sized properties. The system integrates sensors and intelligent video analysis to monitor activities and detect potential security threats automatically [12].

Overall, the reviewed literature demonstrates that artificial intelligence and computer vision technologies have significantly improved surveillance systems by enabling automated object detection, anomaly detection, and behavior analysis. However, most existing systems focus on specific tasks such as crowd monitoring, anomaly detection, or theft detection. There is still a need for an integrated surveillance framework capable of detecting multiple types of incidents—including accidents, violence, fires, and abnormal events—in real time within a unified intelligent monitoring system. This gap motivates the development of advanced multi-incident detection systems for smart city surveillance.

III. METHODOLOGY

This section explains how different components of the system interact with each other, starting from video input to final incident detection and alert generation. Typically, the architecture includes modules such as video acquisition from CCTV cameras, frame preprocessing, object detection, motion analysis, incident evaluation, and visualization through a monitoring dashboard. Presenting the system architecture helps readers clearly understand how the proposed framework processes surveillance data step by step.

Another important component is the dataset description. This section explains the datasets used for training, testing, or evaluating the surveillance system. Information such as the source of the dataset, number of images or videos, object classes, and annotation methods should be described. In many surveillance applications, publicly available datasets such as COCO, Open Images, or specialized traffic and surveillance datasets are used. In some cases, researchers also create custom datasets using CCTV footage to train models for specific incident detection tasks.

The data preprocessing stage is also an essential part of the methodology. Before training or analyzing video data, preprocessing steps are applied to improve data quality and ensure compatibility with deep learning models. These steps may include frame extraction from videos, resizing images, normalization, and methods of data augmentation include brightness adjustment, flipping, and rotation. By making the system more resistant to changes in lighting conditions, camera angles, and environmental variables, data preprocessing enhances model performance.

The model training procedure describes how the deep learning model is trained to detect objects and incidents. This section includes details about the chosen algorithm, such as YOLOv8 for object detection, as well as training variables such as batch size, epoch count, and learning rate, and optimizer settings. It may also describe the computational environment used for training, such as GPU-based systems or cloud platforms. Providing these details ensures that the experimental process is transparent and reproducible.

After objects are detected in each video frame, the system analyzes various features such as object movement, speed, direction, and interactions between objects. These features are used to identify abnormal patterns that may indicate incidents such as accidents, suspicious activities, or abnormal crowd behavior. The system applies predefined rules or machine learning models to classify whether the detected behavior represents a normal situation or a potential incident.

International Student Conference on Next-Gen Computing: Application of AI, Big Data, Quantum Computing, Signal Processing and Cloud Innovations (ICNGC-2026)
International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

The implementation environment should be described to provide technical details about the tools and technologies used in developing the system. This may include the programming language (such as Python), deep learning frameworks like PyTorch or TensorFlow, object detection models like YOLOv8, and visualization tools such as Flask dashboards or graphical interfaces. Information about hardware specifications, including CPU, GPU, and memory, can also be included.

a) Description of System Architecture The proposed framework presents a deep learning-based surveillance system developed to detect incidents from video footages in real time. The system analyzes surveillance videos through several steps, starting with the initial video preparation, object detection, motion analysis, and incident evaluation. The entire operation process of the system is depicted in the suggested architectural diagram in (fig1.1)

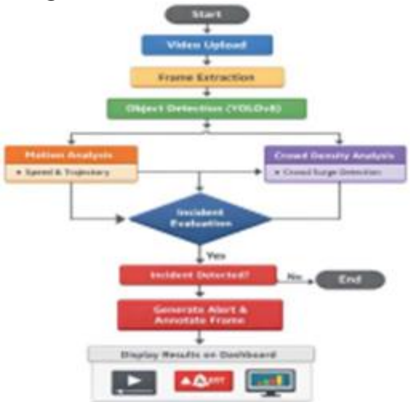


Figure 1(a)flowchart

i) 3.1.1. Video Upload

The steps starts from the input of surveillance video data. The system accepts video streams captured from CCTV cameras, traffic monitoring systems, or uploaded video files.
These videos serve as the primary input for further processing and analysis.

ii) Frame Extraction

Once the video is uploaded, the system performs frame extraction, where the video is divided into individual frames. Each frame represents a single image extracted from the video sequence. This step is necessary because most computer vision algorithms analyze visual data frame-by-frame.
Frame extraction enables the system to process video data efficiently and prepare it for object detection and behavioral analysis.

iii) Object Detection using YOLOv8

After frame extraction, each frame is processed using the YOLOv8 object detection model. YOLOv8 is a modern deep learning algorithm known for its high detection accuracy and real
time performance.

In this stage, the model detects and classifies objects present in the frame, such as:

- Vehicles
- Pedestrians
- Motorcycles
- Emergency vehicles
- Other relevant objects

The identified items are highlighted with rectangular outlines and descriptions, which help the system track object movements and interactions in subsequent steps.

iv) Motion Analysis

The system then performs motion analysis to evaluate the behavior and movement of detected objects. This step focuses on analyzing parameters such as:

- Speed of moving objects
- Trajectory patterns
- Direction of movement

By analyzing motion patterns, the system can identify abnormal behaviors such as sudden stops, collisions, or unusual movement patterns that may indicate incidents such as accidents or suspicious activities.

Figure 1(b) Block Diagram

v) Crowd Density Analysis

In addition to motion analysis, the system performs crowd density analysis. This module evaluates the number of people present in a given area and monitors crowd behavior.

The purpose of this stage is to detect events such as:

- Crowd congestion
- Crowd surge
- Abnormal crowd movements

Detecting crowd density patterns helps identify potential risks such as crowd panic situations or abnormal gatherings in public areas.

vi) Incident Evaluation

The results obtained from motion analysis and crowd density analysis are combined in the incident evaluation module. This module analyzes all extracted information and determines whether an incident has occurred.

The evaluation process considers multiple factors such as:

- Sudden changes in object speed • Collision patterns
- Abnormal crowd behavior
- Unusual movement trajectories

Based on these observations, the system classifies whether the situation represents a normal scenario or a potential incident.

vii) Alert Generation and Frame

Annotation

When an incident is identified, the system automatically generates an alert. The alert mechanism may include:

- Highlighting the detected object
- Displaying warning messages
- Annotating the frame with bounding boxes or incident labels

This step helps authorities quickly identify the location and nature of the detected incident.

International Journal of Science, Engineering and Technology ISSN: 2348-4098, P-ISSN: 2395-4752

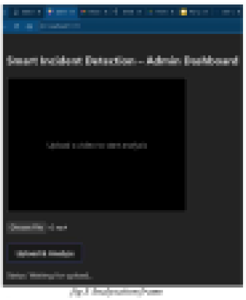
viii) Dashboard Visualization

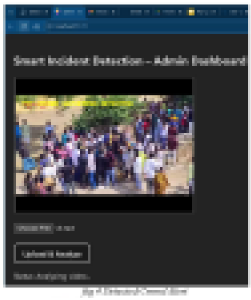
Finally, the results are displayed on the monitoring dashboard. The dashboard provides:

- Real-time video visualization
- Incident alerts
- Analytical charts and monitoring statistics

This interface allows operators or authorities to easily monitor incidents and respond promptly.

IV. RESULTS





The examination of the chosen research articles indicates that advancements in artificial intelligence and deep learning methods have greatly enhanced the efficiency of contemporary surveillance systems. Many studies have successfully used computer vision models like CNN, YOLO-based object detection models, and sequential models like Recurrent Neural Networks and Long Short Term Memory networks to analyze surveillance videos and detect abnormal events. These techniques allow automated systems to detect objects, monitor activities, and identify suspicious behaviors in real time. The studies evaluated indicate detection rates that fall roughly between 85% and 93%, indicating the effectiveness of deep learning methods in surveillance applications.

However, several challenges still exist. Most current systems are designed to detect only a single type of incident, such as accidents, violence, or abnormal crowd behavior. In addition, issues such as high computational requirements, environmental variations, and false alarms can affect system performance. Therefore, developing integrated multi-incident detection systems is essential for improving real-time surveillance in smart city environments.

V. CONCLUSION

The rapid growth of urban populations and the development of smart city infrastructure have exacerbated the demand for smart surveil lance. systems that can effectively ensure public safety and monitor large public spaces. Traditional surveillance systems depend heavily on human operators to observe video feeds from numerous CCTV cameras. However, manually monitoring such a large amount of video data is inefficient and can lead to fatigue, delayed responses, and human errors. As Artificial Intelligence, Computer Vision, and Deep Learning technologies advance, automated monitoring systems can now evaluate video footage. and detecting abnormal activities more accurately and efficiently.

Numerous methods in deep learning like Convolutional Neural Networks, Recurrent Neural Networks, Long Short-Term Memory, models based on transformers, and object detection algorithms have been widely applied for analyzing surveillance videos. Researchers have also used datasets like UCF-Crime, UCSD Pedestrian, and ShanghaiTech to evaluate the performance of different detection models. Many of these systems have achieved detection accuracies between approximately 85% and 93%, indicating the strong potential of deep learning methods in surveillance applications.

Despite these advancements, many existing systems focus on detecting only one specific type of incident, such as violence detection, traffic accidents, or abnormal crowd behavior. In real-world environments, multiple incidents can occur simultaneously, making it necessary to develop more comprehensive surveillance frameworks. Other challenges identified in the literature include high computational requirements, scalability issues, false alarm rates, and difficulties in handling complex real world conditions.

Overall, AI-powered surveillance systems have the potential to transform how cities manage security and public safety. Continued research and technological improvements will help create more reliable, scalable, and efficient monitoring solutions capable of supporting safer and smarter urban environments.

REFERENCE

1. Sanjeev Kukreti; RVS Praveen; Saloni Bansal; Hemanth Raju; Navdeep Singh; Rowsonara Begum "Object Detection in Real-Time Surveillance Using Deep Learning-Based YOLO Framework" in 2025 International Conference on Computational, Communication and Information Technology (ICCCIT), February 2025
2. Varad Agarwal, Yuvraj Singh Rathore, Shyam Singh, Indra Kishor "Smart Surveillance System Using AI for Anomaly Detection" in 2025 International Journal of Innovative Research in Technology (IJIRT), Volume 11 Issue 11 Apr 2025.
3. Manu, Y. M., S. V. Shashikala, N. Darshan, A. P. Dheeraj, N. J. Hemanth, and B. Nishanth Gowda. "Smart Surveillance Camera Using AI." In 2024 Second International Conference on Advances in Information Technology (ICAIT), vol. 1, pp. 1-6. IEEE, 2024.
4. Maheshwari, K.K., Hajano, S., Baloch, S., Zafar, Z. and Fatima, T., 2025, April. AI Based Surveillance System. In 2025 International Conference on Emerging Technologies in Electronics, Computing, and Communication (ICETECC) (pp. 1-6). IEEE.
5. Villegas-Ch, W., García-Ortiz, J. and Sánchez-Viteri, S., 2024. Toward intelligent monitoring in IoT: AI applications for real-time analysis and prediction. IEEE Access, 12, pp.40368- 40386.
6. Siddiqui, Muhammad Zulqarnain, Kamran Khan, and Abdul Jameel Khan. "Smart Surveillance and Crime Control Using an AI-Based Model." Journal of Computing & Biomedical Informatics 6, no. 01 (2023): 216-221.
7. Kakadiya, R., Lemos, R., Mangalan, S., Pillai, M. and Nikam, S., 2019, June. Ai based automatic robbery/theft detection using smart surveillance in banks. In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 201-204). IEEE.
8. Saranya E, Vanitha A, Selvanandhini A, Soundarya VN, Vinisha M. A Survey on AI Based Smart CCTV Surveillance System. In2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS) 2024 Mar 14 (Vol. 1, pp. 1-5). IEEE.
9. Rahman, Md Abdur, M. Shamim Hossain, Ahmad J. Showail, Nabil A. Alrajeh, and Ahmed Ghoneim. "AI enabled IIoT for live smart city event monitoring." IEEE Internet of Things Journal 10, no. 4 (2021): 2872-2880.
10. Qaraqe M, Elzein A, Basaran E, Yang Y, Varghese EB, Costandi W, Rizk J, Alam N. PublicVision: A secure smart surveillance system for crowd behavior recognition. IEEE Access. 2024 Feb 16;12:26474-91.

11. Qaraqe, M., Elzein, A., Basaran, E., Yang, Y., Varghese, E.B., Costandi, W., Rizk, J. and Alam, N., 2024. PublicVision: A secure smart surveillance system for crowd behavior recognition. IEEE Access, 12, pp.26474-26491.
12. Ahmed, Ahmed Abdelmoamen, and Bernard Nyarko. "Smart-watcher: An ai-powered iot monitoring system for small-medium scale premises." (2024).
13. Afreen H, Kashif M, Shaheen Q, Alfaifi YH, Ayaz M. IoT-based smart surveillance system for high security areas. Applied Sciences. 2023 Aug 3;13(15):8936.
14. Ardabili, B.R., Pazho, A.D., Noghre, G.A., Neff, C., Bhaskararayuni, S.D., Ravindran, A., Reid, S. and Tabkhi, H., 2023. Understanding policy and technical aspects of ai-enabled smart video surveillance to address public safety. Computational Urban Science, 3(1), p.21.
15. Sikora P, Malina L, Kiac M, Martinasek Z, Riha K, Prinosil J, Jirik L, Srivastava G. Artificial intelligence based surveillance system for railway crossing traffic. IEEE Sensors Journal. 2020 Oct 16;21(14):15515-26.
16. Jun H. Smart Drone Police System: Development of Autonomous Patrol and Real-time Activation System Based on Big Data and AI. International Journal of Internet, Broadcasting and Communication. 2024;16(4):168-73.