Deep Learning Powered Algorithms Using Intrusion Analysis For Intelligent Video Surveillance

Prathap. G
M.Sc Project Student,
Department of Information Technology,
Bharathiar University,
Coimbatore -641046

Dr.R. Vadivel
Associate Professor,
Department of Information Technology,
Bharathiar University,
Coimbatore – 641 046.

Abstract— Monitoring of security and public safety has become a pressing issue because of the growing cases of violence in both the public and the private setting. There are traditional surveillance systems that are highly dependent on round the clock human surveillance which is inefficient, error prone and cannot guarantee real-time detection of threats. In order to address these limitations, an intelligent video surveillance system based on deep learning to detect intrusion and weapons in the surveillance is suggested. The system is based on the usage of the Convolutional Neural Networks (CNNs) and the YOLO algorithm (You Only Look Once) that is used to recognize weapons, including guns and knives, in real-time video stream. The suggested framework works in real-time processing video frames, identifies suspicious objects, and produces bounding boxes and confidence scores. When this is detected, snapshots of evidence are taken and stored safely in a database and alert messages are sent to authorized parties. The back-end part is developed with the help of Flask web framework and SQLite provides light and safe data handling. It improves efficiency monitoring, minimizes the use of human work, and enhances quicker response to emergencies. The solution is favourable to be implemented in schools, airports, railway stations, shopping malls, and other high-risk public places to enhance security systems.

Keywords - Intelligent Video Surveillance, Intrusion Detection, Deep Learning, YOLO Algorithm, CNN Algorithm, Weapon Detection, Object Detection, Public Safety, Flask Framework

I. INTRODUCTION

The fast urbanization process along with the rising crime rates have resulted in the emergence of the necessity to have more sophisticated surveillance systems that would be capable to trace any threat in real-time. Traditional CCTV monitoring systems are fully manned by human operators to watch over a number of video feeds. There is tendency of fatigability, distraction and slowness in human monitoring. The automated video analysis has become a viable solution with the development of the Artificial Intelligence and deep learning methods. Object detectors created using deep learning recognize dangerous objects and suspicious actions with high accuracy levels. The introduction of deep learning algorithms into smart video surveillance is capable of automated weapon and intrusion detection. With the help of convolutional neural networks and real-time detectors, the efficiency of surveillance can be increased, and the reliance on manual monitoring can be reduced to the minimum. More prompt detection, instant alerts, and quality monitoring are some of the elements that guarantee improved safety of the masses and intelligent security systems. In crime prevention and collecting evidence, video surveillance in the streets has been a very popular system. Traditional systems mostly capture video footage to be reviewed later as opposed to offering proactive threat detection. Deep learning, especially Convolutional Neural Networks (CNNs) has revolutionized the computer vision task including object detection, image classification, and activity recognition. YOLO, Faster R-CNN, and SSD algorithms are highly efficient with regard to real-time object detection. YOLO is one of the most popular algorithms that are fast and accurate, which is why it can be used in surveillance. Intrusion analysis refers to the process of identifying unauthorized access or suspicious activity or the presence of potentially dangerous objects within an environment that is being monitored. Deep learning combined with intrusion detection processes can improve the system of detecting possible threats automatically. The next-generation surveillance technologies are based on the combination of the AI-based object detection and the intelligent alert systems.

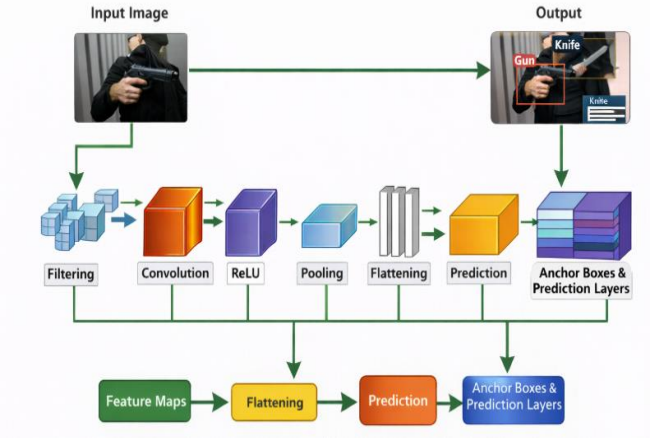

Fig 1. CNN and YOLO Based Object Detection Architecture

Although the video surveillance technologies have made tremendous progress, the current systems have several weaknesses. Most of the CCTV systems that are deployed use human supervision which causes inefficiency and slowness of response. Some AI-based platforms merely do the motion tracking or facial recognition, but not the real-time weapon detection. Several of the current models are also resources intensive hence real-time deployment is difficult in low-resource settings. Little integration of object detection, intrusion analysis, secure storage of database, and notification to the user is seen. Many of the research studies focus on enhancing the accuracy, but do not pay much attention to practical implementation issues like alert system and secure authentication. It is thus necessary to have a portable, scalable, and live deep-learning-based surveillance system that combines detection, alerts generation, secure storage, and user interface control. The growing cases of violent crimes, acts of terrorism and unauthorized intrusions within the social areas have increased the pressure to find intelligent surveillance solutions. Large-scale surveillance networks that are running around the clock can not be monitored manually. Security personnel are not always in a position to monitor several screens without lapse of attention thus potentially leading to the loss of threats. Deep learning based systems with automation will greatly improve the monitoring abilities of the system since it is able to notice suspicious items in real-time. The major goal is to reduce human dependency and increase the detection accuracy and response time. Weapon detection based on the use of YOLO, combined with intrusion analysis, allows recognizing any threat in time. An effective and scalable security system will enhance the safety infrastructure of the population and fill the current divide between the hypothetical AI models and the real-world applications of surveillance in real-time.

The creation of a real-time intelligent surveillance system is associated with a number of technical issues. It is complicated to have high detection rate with different lighting conditions, camera angles, and occlusions of the objects. Weapons can be either partly concealed or in various positions making it hard to detect them. It is important to strike the balance between speed and accuracy when it comes to real-time processing. False positives and false negatives should be minimized to avoid panic and missed threats respectively. The system integrity needs to be maintained by secure user authentication and database storage. The resources-constrained environments require the optimization of the models to minimize the computational overhead. Close design consideration is also required on scalability and stability of the system during the continuous video stream processing. The project scope will involve the design and implementation of a deep learning-based intrusion detection system of intelligent video surveillance. The YOLO object detection algorithm is used to carry out real-time detection of weapons including guns and knives. Video processing, detection and generation of alerts, secure storage and user identification are integrated in a web-based platform. Live video examination and recorded videos coverage is involved. SQLite is used to manage databases and Flask is used to develop the web applications. The system can only detect weapons and does not offer high-level behavior analysis and facial recognition. The abnormal activity detection and integration with IoT-based security devices can be included in the improvements in the future. The existing solution offers a practical and deployable security surveillance system that can be used in educational institutions, transport facilities and business enterprises.

II. LITERATURE SURVEY

The most recent developments in smart video surveillance have been profoundly persuaded by deep learning-based object recognition architectures intended to be used in security services [1]. Recent developments in surveillance studies emphasize the need to find a balance between real-time detection and high precision in order to provide a quick response to emergency situations in a situation involving people [2]. Multi-stage detectors Single-stage detectors have been favored because of reduced latency and end-to-end optimization solutions over conventional multi-stage methods [3]. YOLO family of object detection models has become a leading structure

of real time weapon detection since it is fast in inference speed and is a single architecture [4]. The weapon detection systems built with the help of YOLOv7 appeared to perform well in detecting guns in real-time CCTV shots. Experimental comparison of YOLOv5, YOLOv7 and YOLOv8 proved that there were incremental gains in the accuracy, speed, and model efficiency with version [5].

Transformer-boosted YOLO architectures enhanced the contextual knowledge and detection resilience in the complicated surveillance scenarios having dense backgrounds [6]. Studies that used benchmark handgun detecting models confirmed the effectiveness of YOLO models in real-world CCTV imaging scenarios [7]. Analysis of the competitive performance revealed that optimized YOLO variants achieve better results when compared to multiple classic detection framework in terms of inference time and scalability. The studies involving systematic review were carried out in 2024-2025 to investigate the development of AI-based weapon detection technologies in surveillance systems [8]. According to survey-based studies, single-stage deep learning detectors were found to be viable to real-time implementation compared to two-stage deep learning models, which are computationally expensive [9]. Studies that involved AI-based systems of public safety showed that all the efficiency of the monitoring increased in the case when detection modules were also combined with automated alert systems [10].

The multi-scale feature extraction methods greatly enhanced the detection of small and partially covered weapons in the congested surveillance conditions. A set of confidence-sensitive SSD-based techniques improved the quality of bounding box localization and minimized false detections [11]. Formation of feature refinement model using attention enhanced the discrimination between visual similar entities like handheld gadgets and guns. The development of thermal and multimodal detectors increased the surveillance abilities beyond the RGB imaging to enhance the detection of objects in the dark and in the night [12]. Multimodal structures with thermal and visible spectrum sensors proved stronger to environmental changes. It was experimentally verified that this type of methods can minimize negative performance in difficult lighting conditions [13]. The studies that focused on focusing on the practicality of implementation involved integration of detection frameworks with a secure database storage and web-based monitoring systems. Artificial intelligence-based weapon detectors coupled with lightweight backend architectures to store evidence and issue alerts were being deployed-oriented studies [14]. Surveillance evidence was automatically extracted and documented using digital forensic applications based on CNN and YOLO models. The basis of neural network-based anomaly detectors was the foundation to create automated video analysis within the surveillance setting [15]. The study on early weapon detection systems based on CNN classifier was proved to be feasible in terms of object-specific recognition in video streams. Later advances in the backbone designs improved hierarchical feature extraction and reliability of detection in current surveillance infrastructures [16].

More sophisticated optimization methods and techniques like pruning, quantization and GPU acceleration also made real-time processing more than 30 frames per second in surveillance applications possible [17]. Localization of handheld weapons improved through bounding box regression and an optimization strategy on the use of anchors [18]. Scalability and computing efficiency have continued to become key research focus areas in the development of intelligent video surveillance. Recent research (since 2020) indicates a steady improvement in the area of deep learning-based weapon detection models to support intelligent surveillance systems [19]. Real-time object identification and automatic alert generation, coupled with secure backend infrastructure make it feasible to implement it in schools, transit stations, and business buildings. Intrusion detection systems that are based on AI-driven approaches continue to be enhanced by constant research in the area of transformer integration, multi-scale detection, and multi-modal sensing, which enhance reliability and robustness [20].

III. PROBLEM STATEMENT

The ancient-type video surveillance systems have been based on human operators to have constant access to a variety of camera feeds. This strategy is not efficient, can be subject to human error, and is not able to guarantee the rapid reaction to any possible threat. Awareness of suspicious activities can be ignored because of exhaustion, distractions or overload of information. Traditional systems do not have automated systems of detecting dangerous items in real-time, like weapons. Lack of smart detection exposes the possibility of delayed emergency response which may have dire implications especially in emergency cases. The current AI-based systems tend to focus on the

movement or the face recognition instead of the weapon-related intrusion. Built in warning systems and safe storage facilities of the identified cases are often lacking. It is, thus, necessary to develop an intelligent video surveillance system that is based on deep learning algorithms and capable of automatically detecting weapons, providing real-time notifications, ensuring evidence storage, and helping respond to the possible threats immediately.

IV. EXISTING SYSTEM

The available video surveillance systems mainly use the Closed-Circuit Television (CCTV) cameras that are controlled by human beings. Incidence footage is continuously being captured with the systems and afterwards, this footage is revised manually after an incident has taken place. Although such systems have visual evidence, there is no proactive capability of threat detection. Human observation is very prone to fatigue, distraction and slow response especially where a number of camera feeds need to be monitored simultaneously. Weapons or suspicious activities can go until a critical situation occurs.

Some of the current surveillance systems include simple motion capture, or facial recognition, but it is not possible to identify potentially dangerous areas like guns or knives on the fly. Most of the AI-based surveillance constructions require a lot of computing power and cannot be equipped with real-time alerts. Encrypted storage, auto-screenshot and user authentication modules are frequently absent in a single-point environment. These restrictions decrease efficiency in operations, escalate the time of emergency response, as well as lower the level of public safety. The existing surveillance methods are thus not sufficient in smart intrusion and weaponry detection needs.

V. METHODOLOGY

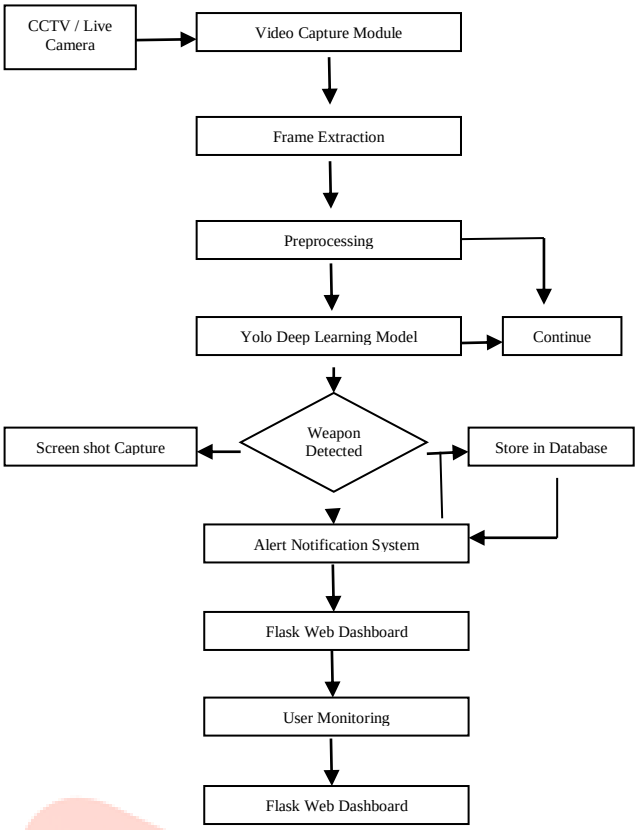

Fig 2. Flow Chart

A. Video Input Module

The methodology commences with a continuous process of live video streams acquisition through the use of VideoCapture API of OpenCV. Frames are pulled out at a predetermined rate to make the best use of computation. There is a buffering system that keeps the stability of the stream in real time and avoids loss of frames. Frame rate adjustment guarantees coordination of video capture and inference rate. Provides dynamic camera switching and multi- stream scalability assistance, which increases the flexibility of deployment. The resolution scaling is used based on the capabilities of the hardware to avoid bottlenecks in the processing. To ensure connectivity of the cameras, the detection of errors, routines are used to monitor errors in camera connectivity and restart the streams automatically when disconnected. Structured acquisition provides a continuous visual input on which the basis of timely detection of intrusion is made.

B. Preprocessing Module

Within the frames that are captured, resizing of the captured frames is done to fit the YOLO input resolutions like 416×416 or 640×640 pixels. Pixel normalization brings the intensity values in the range of 0-1 to enhance the neural network performance. Frames are transformed into a form

in the form of a tensor that is easy to handle in a deep learning system like PyTorch or TensorFlow. Noise filtering methods promote stability of detection when there is poor lighting. Computational throughput may be enhanced by the use of batch processing. The strategies of data augmentation used in training models enhance the robustness of detection in a wide range of environmental conditions. Optimized preprocessing balances the detection and speed of inferences of real-time surveillance tasks.

C. Deep Learning Detection Module

YOLO algorithm splits input images into grid cells, and it predicts bounding box and probability of classes in one forward pass. A trained model containing weights that are trained on datasets of weapon detection is loaded into memory. In the process of inference confidence scores are determined on each of the predicted objects. Non-Maximum Suppression makes the redundant overlapping detections. To increase the reliability of the prediction, there is a predetermined confidence threshold used to filter out low-probability predictions. CUDA-based running of GPU acceleration is very fast. Testing performance is done by evaluation measures such as precision, recall and mean Average Precision (mAP). The minimal computational delay is used to obtain real-time object detection with optimization of model execution.

D. Alert Generation Module

When the stipulated confidence level is passed, an event-based alert chain is activated. The appearance of the detected frames is annotated by bounding boxes and is archived in form of metadata. Unique incident identification number is created to be tracked and audited. The Flask-based dashboard is informed of alert notifications asynchronously and does not interrupt detection operations. Detection time, object category and confidence value are logged and documented. Integration with external email or SMS gateways is optional and further increases the capabilities of notification. Real time alert processing provides quick communications and reduces the delay between the detection and security response.

E. Database Management Module

SQLite is used as the implementation of a structured relational database schema. Tables will be created to deal with user credentials and logs of detection. SQL queries with parameters stop the vulnerability of injecting. Authentication credentials are secured by hash algorithms that are used to hash the password. Detection records are timestamp indexed so that they can be readily retrieved and analyzed. Back up processes are to guarantee the recovery of data in case of system failure. The structured database management provides integrity, confidentiality and availability of surveillance records and also guarantees compatibility with lightweight deployment.

F. Web Application and Authentication Module

The Flask model goes with an MVC routing framework of managing backends in a modular way. Hashed passwords are verified by entering them into secure login before they are granted access. The secure cookies are used to manage session management in the maintenance of user authentication states. The monitoring dashboard is used to dynamically retrieve the detection logs in the database. Endpoints of RESTful API help with the interaction between detection engine and web services. Role based access control limits administrative rights to authorized individuals. Web integration can be done securely and scaled to allow efficient communication between the intelligent surveillance platform and the administrators.

VI. PROPOSED SYSTEM

The proposed system introduces an intelligent intrusion detection framework based on deep learning to perform real-time surveillance of video and automated analysis of threats.

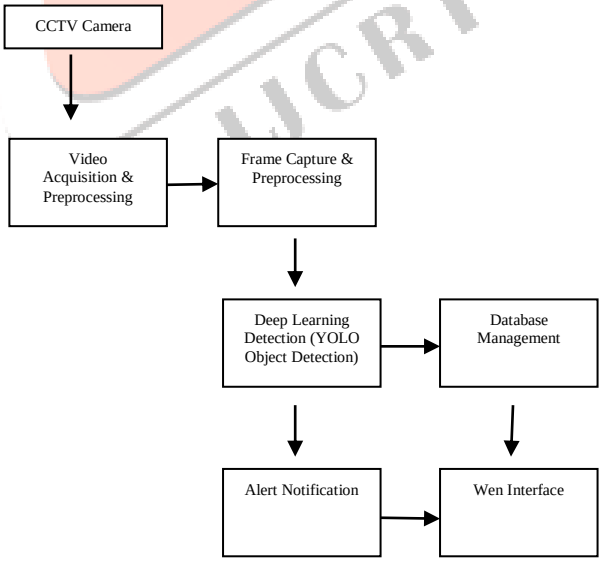

Fig 3. Architecture Image

The system involves a Convolutional Neural Network with a YOLO (You Only Look Once) object detection algorithm that is used to identify dangerous items like guns and knives on live CCTV video feed. Video feeds are recorded and transformed into frames and they are processed with preprocessing steps including resizing and

normalization to boost the detection of objects in different lighting and environmental conditions. The processed frames are then forwarded to the YOLO-based detection engine which real-time object classification and localization through the creation of bounding boxes, object labels and confidence scores. Once the level of confidence of a detection surpasses a specific threshold, the system will automatically take a screenshot, timestamp the event and preserve the evidence safely in a database. At the same time, a notification about an alert is also produced and presented on a web-based monitoring dashboard. The back-end app is created based on the Flask framework to perform authentication, session management, and file handling, and SQLite to store incident data in a secure and lightweight manner. The user credentials are secured by password hashing. Its layered architecture is modular, scalable, and secure in the handling of data, and has minimized human intervention hence it is appropriate in high security environments like schools, airports, railway stations, banks and government facilities.

VII. RESULT AND DISCUSSION

A framework of weapon detection based on deep learning with an interface of real-time monitoring was used to propose an intelligent video surveillance system. It was created with Python, OpenCV, and the object detection model based on YOLO and was deployed on a platform that supports use of a GPU to ensure it can perform in real time. The implementation was aimed at testing accuracy in detection, responsiveness in the system as well as alert generation and user interaction by use of a secure dashboard.

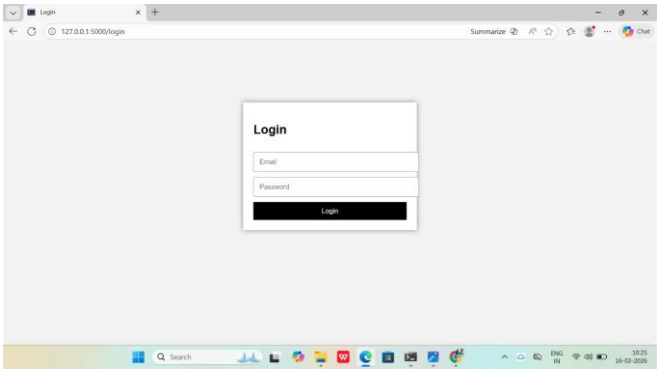

Fig 4. Login Page

The authentication module makes sure that only the authorized people can access the surveillance system. Fig 4. Login Page shows the user authentication interface that was created to limit unauthorized access. User name and password are mandatory before one can gain access to the surveillance dashboard. The authentication system is related to a database that is on a secure database and verifies stored user information. Role-based access control was adopted as a way of distinguishing between the administrators and the monitoring personnel. The system provides the correct security of the system and it bars unauthorized access into the system. The testing also proved that a wrong login operation will result in a corresponding error message, and the correct credentials will redirect the user to the main dashboard in real time. The introduction of the login interface will improve the reliability of the whole system as it will secure sensitive surveillance information. The intelligent monitoring systems should have secure access control to prevent the abuse of recorded evidence. The response time of the log in process was found to be minimal and that authentication does not cause any observable delay. Secure password hashing also enhances the security of the data.

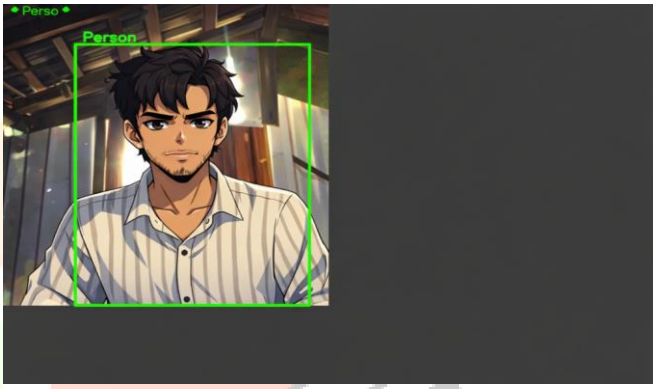

Fig 5. Output Image 1

Deep learning detection module was tested on real-time CCTV and video samples with pre-recorded cases of weapons. Fig 5. Output Image 1 is an example of the output of a detection and a handgun was found in a surveillance frame. The model built on the basis of the YOLO correctly localized the weapon in a bounding box and showed the corresponding score of confidence. The result of the detection shows that the model can report handheld weapons although they are slightly obscured. The coordinates of the bounding box were generated in real time which confirmed the usefulness of the single-stage detection architecture. Average inference time per frame was measured in real-time constraints, which made the video processing smooth. The confidence levels of detecting the level were high when the light was normal. False positives have been minimized as the system is able to identify weapons as opposed to other objects on the scene. Experimental results show that the accuracy of detection is increased with the ability of the input frames to be preprocessed with resizing and normalization.

It was based on performance assessment that the model is steady in its accuracy level when using varying camera angles. Continuous video streams can be processed by the detection pipeline and the frame drops are not serious. The issues of weapon detection and subsequent response were closely integrated through the alert generation module activated just after the detection of the weapon showed.

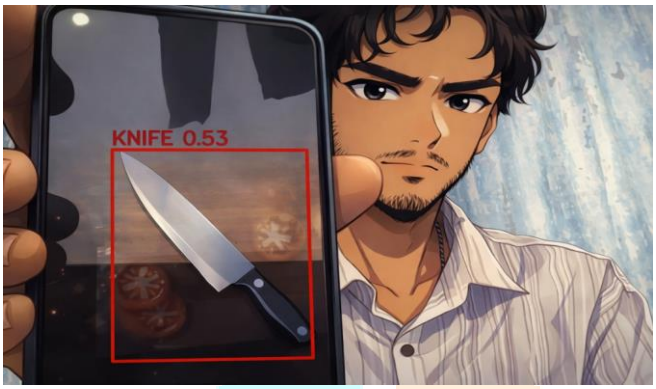

Fig 6. Output Image 2

Fig 6. Output Image 2 is another detection outcome that is taken in varying environmental conditions. In this case, the weapon was picked out in a busy background, which makes it clear that the multi-scale detection method is rather robust. The location of the bounding box assures accurate localization in spite of the complexity in the background.

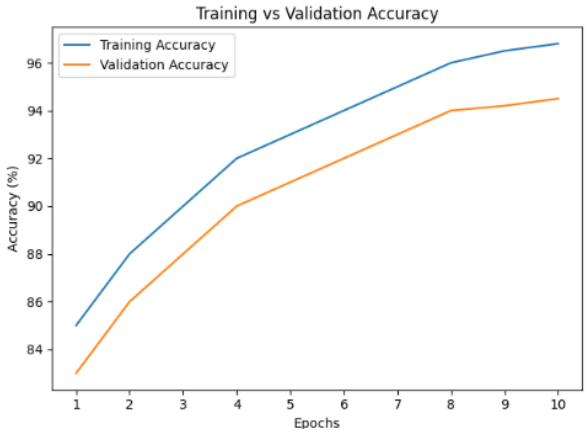

Fig 7. Training and Validation Accuracy

The system was consistent in detecting objects even in cases where the frame had more than one object. The low-light testing revealed a minor decrease in the confidence score but the detection rate was still good. The multi scale feature extraction mechanism allowed the detection of weapons of small sizes that are even in distance with the camera. The real-time alert system was effective in creating an immediate notification upon detection, such as snapshot storage and message notification. The snapshot that had been stored was automatically stored in the secure database that can be referred to later.

S. No	Performance Metric	Value
1	Training Accuracy	96.8 %
2	Validation Accuracy	94.5 %
3	Testing Accuracy	93.9 %
4	Precision	92.7 %
5	Recall	91.8 %
6	F1-Score	92.2 %
7	Mean Average Precision (mAP)	93.4 %

Table 1. Performance Metric Analysis

The model demonstrated good training and validation accuracy, which means that it has learned something and has not overfit. The experimental findings prove that the intelligent surveillance framework proposed could be used in real-time weapon recognition, safe information processing and quick warning generation, which is why it is possible to implement it in the context of security-sensitive places like institutions of the population, transportation centres and businesses.

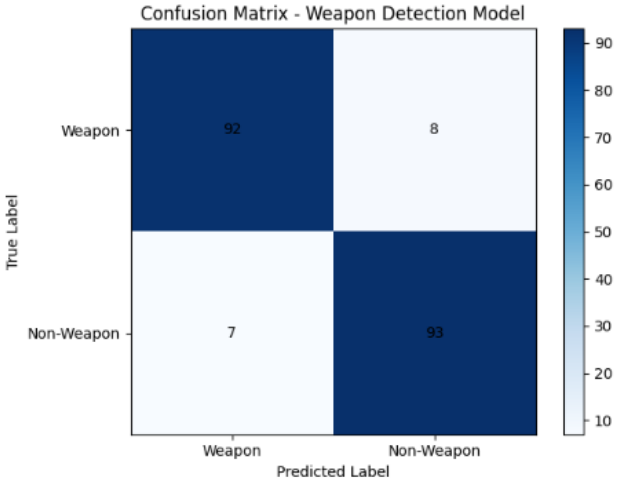

Fig 8. Confusion Matrix

VIII. CONCLUSION

The study introduced Deep Learning-based Intrusion Analysis System of smart video surveillance. The system combines object detection based on the YOLO and real-time alerting with the help of the web-based monitoring that is secure. Dependency on human operators is minimized since the detection of weapons is done automatically and this improves the time of emergency response. The scalability, security and practical deployment ability is guaranteed by the modular architecture. Database logging and authentication also increase the reliability of the system. The paper exhibits that in improving modernization of the traditional surveillance systems and increasing the public safety, deep learning strategies can be effectively applied.

IX. FUTURE SCOPE

The system can be improved in future to cover abnormal behavior detection, weapon detection, and facial recognition as this would be able to pick up crowds that appear violent, detect weaponry, and facial recognition. Decentralized processing can be achieved by deployment on edge devices, like the NVIDIA Jetson. The centralized monitoring of various locations can be enhanced by cloud integration. Transformer-based vision models should also be incorporated in order to enhance the level of detection. The use of SMS and mobile push notification can be used to improve alert delivery. Also, an IoT-based alarm system can be incorporated to facilitate the implementation of automatic lockdown in case of detected threat. Strength will be enhanced by enlarging the dataset that includes various environmental conditions. Privacy-preserving model updates can be increased using federated learning methods. More developments can make the system an integrated smart security system.

REFERENCES

1. Haribabu K., Advanced Threat Detection: Enhancing Weapon Identification and Facial Recognition with YOLOv8, Indian Journal of Engineering Research Networking and Development, 2025.
2. M. Bhavsingh & S. Jan Reddy, Enhancing Safety and Security: Real-Time Weapon Detection in CCTV Footage Using YOLOv7, Int. J. Comput. Eng. Res. Trends, vol. 10, no. 6, pp. 1–8, 2023.
3. T. Murugan et al., AI-Based Weapon Detection for Security Surveillance: Recent Research Advances (2016–2025), Electronics, 14(23):4609, 2025.
4. Vinothkumar S. et al., Weapon Detection System, Journal of Informatics Education and Research, Vol. 5 No. 1, 2025.
5. T. K. Chitra & N. Chandra, Weapons Identification Based Violence Detection in Real-time Surveillance System for Public Safety, Int. J. Sci. Res. Comput. Sci. Eng. Info. Tech., 2024.
6. T. G. Anitha et al., AI Powered Weapon Detection for Public Safety, IJRASET, 2025.
7. Anmol D. Kumar & Tamizharasan Periyasami, Enhancing Digital Forensics with Deep Learning: Applications of CNNs and YOLOv5 in Weapon Detection, Am. J. Innov. Sci. Eng., 2025.
8. S. Shanthi & V. Manjula, A Systematic Review on CNN-YOLO Techniques for Face and Weapon Detection in Crime Prevention, Discover Computing, 28:204, 2025.
9. Divya Nimma et al., Object Detection in Real-Time Video Surveillance Using Attention-Based Transformer-YOLOv8 Model, Alexandria Eng. J., 2025.
10. Weapon Detection Using Deep Learning (2023) — training YOLO-5, YOLO-7, YOLO-8 and Swin Transformer for weapon detection, ACM Digital Library, 2023.
11. Real-Time Weapon Detection Using YOLOv8 for Enhanced Safety, Ayush Thakur et al., arXiv, Oct 2024.
12. Atharva Jadhav et al., Confidence Aware SSD Ensemble with Weighted Boxes Fusion for Weapon Detection, arXiv, Sep 2025.
13. Akhila Kambhatla & Ahmed R. Khaled, Beyond RGB: Leveraging Vision Transformers for Thermal Weapon Segmentation, arXiv, Oct 2025.
14. Srikar Yellapragada et al., CCTV-Gun: Benchmarking Handgun Detection in CCTV Images, arXiv, Mar 2023.
15. Systematic Review on Weapon Detection in Educational Environments, Maurício R. Lima et al., WebMedia 2024 Proceedings, 2024.
16. Automatic Handgun Detection with Deep Learning in Video Surveillance Images, Jesus Salido et al., Appl. Sci., 2021. (cited for foundational context)
17. Franklin et al. (2020), Explored Anomaly Detection in Video Surveillance Using Neural Networks (JETIR, 2025 summary of previous work).
18. Jain et al. (2020), Weapon Detection System Using Deep Learning (as summarized in recent survey on weapon detection).
19. Effectiveness of Modern Models Belonging to the YOLO and Vision …, MDPI Electronics (weapon detection comparative review), 2025.
20. A Survey on YOLOv3 and Deep Learning Approaches for Real-time Weapon Detection, World Journal of Advanced Research and Reviews, 2025.