iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VI Month of publication: June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72772

Advanced CCTV Using Machine Learning and Internet of Things

Ms. Ambika A[1], Praveen S[2], Saravanan G[3], Surya R[4]
Dept.of Computer and Communication Engineering Sri Sairam Institute of Technology Chennai, India

Abstract: Current CCTV systems mainly act as surveillance tools, often just providing video evidence after an incident has occurred, without the ability to detect threats in real-time or respond automatically. This paper introduces an CCTV system that combines Machine Learning (ML) and Internet of Things (IoT) sensors, shifting from mere monitoring to proactive surveillance. The system employs YOLO V8 algorithms for real-time object detection, recognizing suspicious activities and analyzing behaviors, which helps in crime detection and minimizes response delays. Moreover, it integrates IoT-based environmental sensors—like motion, temperature, and acoustic sensors—to boost context-aware threat detection and automate alert systems.The proposed architecture allows for remote access, sends automated alerts to authorities, and supports intelligent decision-making, paving the way for smarter surveillance in public safety, critical infrastructure, and smart city initiatives. Performance evaluations show a notable increase in threat recognition accuracy and response times compared to traditional CCTV systems.

Keywords: CCTV Systems, Machine Learning, Suspicious Activities, Internet of Things, Object Detection, Smarter Surveillance, YOLO V8 Algorithms.

I. INTRODUCTION

Traditional CCTV Systems have long been used for Security and Surveillance in Public workplaces and Metropolitan cities.It is used to record live videos, images, and stored in DVR (Digital Video Recorder) and later used on for Evidence in Investigation. However, these systems typically rely on human operators to monitor footage, which is both time-consuming and prone to error. Moreover, Conventional CCTVSystems lack real-time intelligence, limiting their ability to detect suspicious behavior or respond promptly to incidents.

The rise of Internet of Things (IoT) and Machine Learning (ML) has opened new opportunities to transform surveillance systems into intelligent, automated solutions. IOT enables CCTV cameras and sensors to be connected to the internet, allowing remote access, centralized control, and data transmission to cloud or edge computing platforms. This can be integrated to perform a real-time detection.By integrating ML and IoT, CCTV systems can become aproactive rather than reactive—capable of sending real-time alerts, learning from new data, and reducingthe needfor the constant human oversight. This evolution in CCTVsurveillance technology enhances safety, responsivenessandoperational efficiencyin various sectors,includingsmart cities, transportation,industrial facilities, and varioussectors, including smart cities, transportation, industrial facilities, and residential environments.

II. LITERATURE SURVEY

Real-time computer surveillance for crime detection madewith smart surveillance designed and developed real timecomputer vision algorithms forvisual surveillance systems,[Abdul Jaleel et al., 2022], developed a real-time CCTV foredge-computing and deep learning to processing data at or nearthe data source (edge devices like cameras or real-timeservers), reducing the need to send raw video to centralized data servers.

IoT BasedTheft-Detection systems designed with sensors developed to environmental conditions like temperature, humidity etc. Thermostat-IoT based detection systems for leverages temperature variation as an indicator of human presence or intrusion, offering an innovative method fo theft-detection like detecting human temperature, heat, light, physical movements using Thermal Sensors. These sensors collectthe input through environment, analyze data andthrough actuators gives output in the form as alert or SMS.This System is connected to the Mobile Applications, orDatabase systems that provides information in the form of SMS or Signals. This was proposed by [Srinivas et al., 2022].

[Kumar et al., 2018] developed a smart surveillance system. The author created a surveillance system that utilizes,Raspberry Pi and IP cameras, allowing for remote monitoring via cloud and mobile applications.

This system offers live video streaming and basic motion detection features, but it doesn't include advanced video analytics or AI-driven decision-making capabilities. The information is shared to the mobile applications via internet or Bluetooth asSMS or alert signals.

Real-Time Object Detection with YoLoV3 introduced by [Redmon et al., 2016] uses bounding boxes and class probabilities for the object it contains. It uses predefined boxes to detect objects of varying shapes and size more accurately for object-detection. Darkest-53 as the backbone as CNN. This CNN system is more advanced and made up with multi-layer neural networks. This research explores with real-detection using bouncing boxes for three-scale dimensional boxes and sigmoid for class probabilities that collects data, analyzes probabilities with real-time circumstances.

III. IDEAL STATE

An Ideal state for CCTV surveillance system that seamlessly blends Machine Learning (ML) and the Internet of Things (IoT). This creates a smart, adaptable, and real-time monitoring environment. Unlike the old-CCTV surveillance systems that just record video footage and images in Digital-Video Recorder and does not have the capability to detect the real- time intrusions or unusual behaviours. This ideal system is proactive. It actively analyzes video feeds, spots unusual behavior, and can even take action on its own when it senses a security threat. This makes a setup where edge devices like IP cameras and microcontrollers. For IP Cameras , Raspberry Pi are fitted with lightweight machine learning models that can handle video analytics right on the device. These smart models can spot human presence, recognize faces, classify objects like weapons and even pick up on motion patterns that might signal an intrusion or violence.

This integrated module with IoT connectivity devicescan easily connected and share data with IP cameras for centralized management, real-time alerts, and remote access.

IoT devices like smart sensors like Acoustic Sensors/Microphones and integrated with IP cameras used to detect abnormal audio patterns such as glass breaking, screaming, or gunshots. These abnormal sounds is shared as to the ML models as a sound-models and these models classify sound types for contextual anomaly detection. These data analyzed by ML models using YOLOv8 algorithms to detect the trained models, compare and then detect whether it is a usual or unusual activity. After further classification of data models and detection , information is shared to IoT connectivity devices for making alerts or sending SMS notifications to the nearby authorities via internet or mobile applications.

Cloud storage plays a crucial role in deep analytics, storing historical data, and updating models. Perfect system would send real-time notifications to authorities or users through mobile apps, automate access control, and seamlessly integrate with smart city infrastructure.

IV. METHODOLOGY

The System methodology for implementing this intelligent CCTV system integrates Machine Learning (ML) algorithms and Internet of Things (IoT) devices that employs a system design and architecture development that defines the overall system architecture, including CCTV cameras, IoT sensors, cloud integration, and processing units.

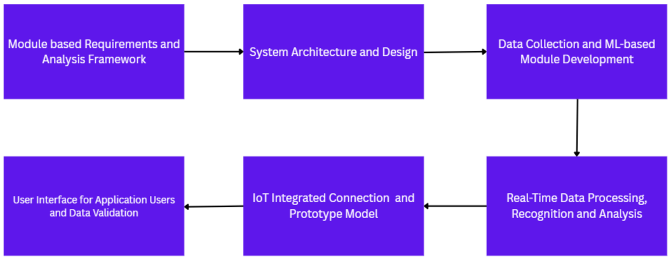

Module based Requirements and Analysis Framework
System Architecture and Design
Data Collection and ML-based Module Development
Real-Time Data Processing, Recognition and Analysis
IoT Integrated Connection and Prototype Model
User Interface for Application Users and Data Validation

Fig.1 Flow Diagram

A. *Module based Requirements and Analysis Framework*

The module development focuses on proposed advanced CCTV system withthecombination of Machine Learning (ML) and the Internet of Things (IoT) to boost real-time surveillance, spot anomalies, and automate alert systems. Its architecture is built to function across several layers, starting with the perception layer. Here, ML-enabled CCTV cameras and with acoustic sensors are set up. These devices gather raw data, including live video feeds and environmental signals, forming the essential input for the entire system.

Initially, this data is collected at edge devices, which are often embedded systems like Raspberry Pi connected IP cameras acting as local processing hubs. At this point, some preliminary processing takes place, including frame extraction, image resizing, background subtraction, and noise filtering, all aimed at getting the data ready for machine learning inference.

Once the data is gathered and assembled, it gets sent off to the various functional modules that are built to trackdifferent tasks within the surveillance system. The first of these is the video analytics module, which uses trained machine learning models like YOLOv8 to spot and categorize objects in the video framesthink people, vehicles, or even weapons. This module also features facial recognition and behavior analysis, employing techniques like CNN-RNN to pick up on suspicious activities such as loitering, running, or large groups forming. Next up is the sensor fusion module, which brings together and correlates data from various sensors to boost detection accuracy. The anomaly detection module, uses pattern recognition and machine learning algorithms to identify any unusual behavior, like unauthorized access, intrusions, or environmental threats. These alerts arein real time and sent to the alert generation module, which organizes the important details and sends out notifications via SMS, email, or mobile apps, often including video clips or images. These alert modules data is shared to the IoT integrated connectivity devices like sensors for alert sounds. These alerts used to alert the local authorities to take further steps or actions for the public threats.

Fig.2 Sample datasets of threatening weapons for Module based Requirements and Analysis Framework

B. System Architecture Design and Module Layering

This proposed advanced CCTV solution features an interconnected system architecture design and module that brings together IoT-enabled hardware sensors, ML-trained edge camera, cloud storage, and smart software layers to facilitate real-time suspicious surveillance and threat detection. This architecture is built on different functionality layers integrated to perform a real-time surveillance. The IoT hardware devices,ML IP camera layer, network layer, cloud layer, and application layer. The IoT hardwarelayer with high-definition CCTV cameras alongsidesensors, including motion detectors, infrared acoustic sensors, all linked through embedded systems. These devices serve as data collection units, constantly capturing video feeds and contextual sensor information. The edge devices take charge of local processing, utilizing devices like Raspberry Pi, or ARM-based systems that come equipped with pre-trained machine learning models. This layer handles specific tasks such as object detection, facial recognition, and identifying suspicious activities, which lightens the load on the cloud and speeds up decision-making. The network layer devices that guarantee secure and dependable data transmission Wi-Fi, Internet or Mobile Applications. Meanwhile, the cloud layer is responsible for high-level processing, centralized data storage, and extensive analytics. Lastly, the application layer offers an interactive interface for users through mobile or web applications, featuring real-time video monitoring, event-based alert notifications, access to data login and system dashboards.

This modular and layered architecture ensures scalability, robustness, and real-time responsiveness, making it ideal for smart surveillance in public spaces, restricted authoritarian, and critical infrastructure environments.

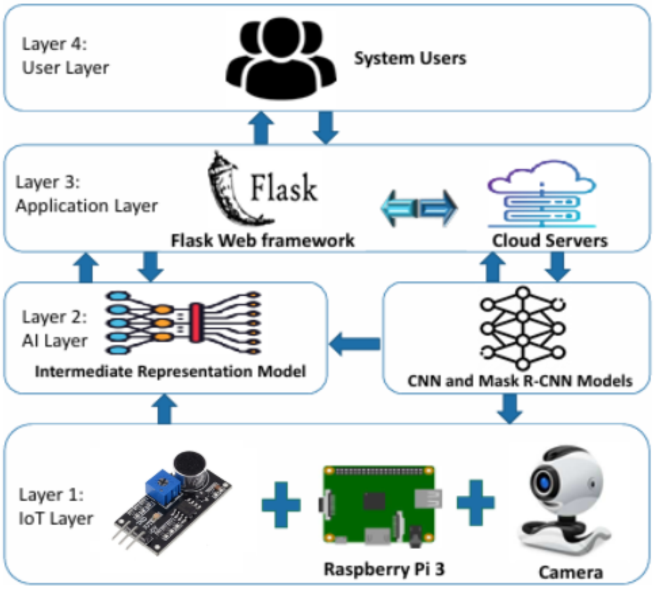


Fig.3 System Architecture and Design

C. Data Collection and ML-basedModuleDevelopment

The real-time detection using machine learning modulesis the main analytical core that begins working with the powerful edge devices like IP cameras.The data is first pre-defined or trained where videos or imagedata captured, recognized, trained with pre-defined datasetswith the help of ML algorithm models. The object detection module is then activated, where deep learning algorithms such as YOLOv8 scan each frame to identify and label suspicious activities, anomaly detection and weapons detection. These models use convolutional neural networks (CNNs) to detect spatial patterns and generate bounding boxes around detected objects, along with class probabilities.

These devices break down the data into individual frames, adjust their size and format to the data consistent and then feed them into the YOLOv8 algorithms. Here the system starts to recognize about the intrusions. Using advanced deep learning models like YOLOv8algorithm, the object detection module scans each frame to identify things like object-detection, facial-analysis, suspicious detection These models use techniques similar to human-based recognition by detecting patterns. Once the data patterns identified, the system uses bounding boxes, class patterns around it and estimates the activity detection allowing it to take further decisions and information.

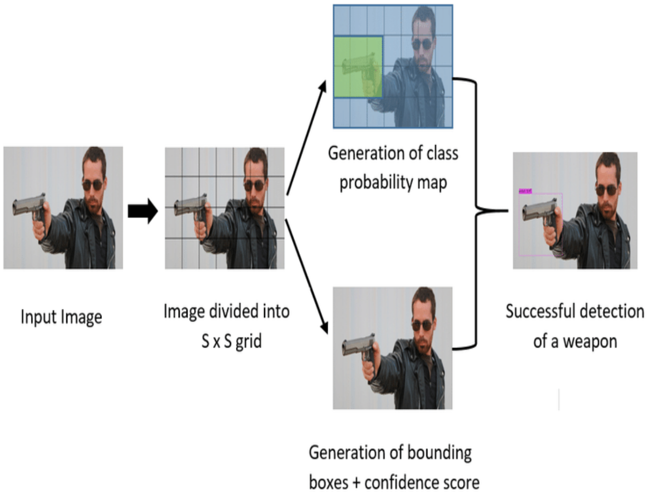


Fig.4 ML-based dataset training modules using bounding boxes and class probabilities

D. Real-Time Data Processing, Recognition and Analysis

Real-time data processing, recognition, and analysis uses surveillance accuracy and response times. In this module, CCTV cameras and IoT-enabled sensors are constantly recording videos and image data and transmits the data to cloudservers forprocessing.Machine learning algorithms then divide into data streams in real-time, spotting patterns, recognizing faces, identifying objects, and detecting any unusual activities like weapon detection, abnormal activities and recognizing the person behind these activities. In real-time analytics, threats can be recognized immediately, and alerts can be automatically sent to security teams or emergency services. This processing reduces human dependency, reduces false alarms, and speeds up decision-making. The data gathered from various points in the network can be compiled for trend analysis, predictive insights, and adaptive learning, helping the system to grow and improve over time.These Real-time data processing gets data from the cameras, extracts the data and identifies whether a person involved in usual or unusual activity and related to that, the data is finally processed and alert is transmitted to mobile applications.

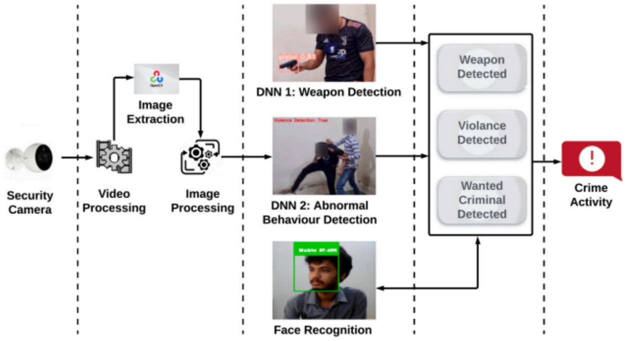
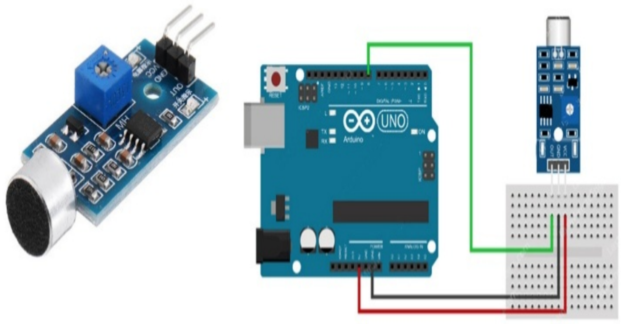


Fig.5 Data Processing and Recognition Analysis

E. IoT Integrated Sensor Module

The IoT Integrated Data Transmission and Alert Notification Module is a crucial part of cutting-edge CCTV systems that leverage machine learning (ML) and Internet of Things (IoT) technologies. This module facilitates real-time communication among edge devices, sensors, cloud storage, and end-users, ensuring real-time detection and response to potential threats. It uses gathering real-time data from CCTV cameras and IoT sensors like motion detectors, and acoustic sensors. These data streams are processed right at the edge, where lightweight ML models analyze video and sensor inputs to spot anomalies such as intrusions, unusual sounds, or suspicious activities. When an event is identified, the relevant metadata or processed data is sent out using efficient IoT communication protocols. This guarantees low-latency and bandwidth-optimized data exchange. At the same time, the system sends out real-time alerts through various channels, including mobile notifications, SMS, emails, and web-based dashboards, ensuring that authorized personnel are immediately informed. The notification system is integrated with cloud services like Firebase forreliable delivery. All event logs, media clips, and sensor data are encrypted and securely stored in the cloud, allowing for remote access, historical analysis, and adherence to surveillance policies. A feedback mechanism is also included, enabling users to classify alerts as accurate or false, which aids in retraining and enhancing ML model performance over time. In summary, this module guarantees intelligent surveillance with real-time decision-making capabilities, efficient data management, and robust alert systems, significantly boosting the responsiveness and reliability of next-generation CCTV infrastructures.



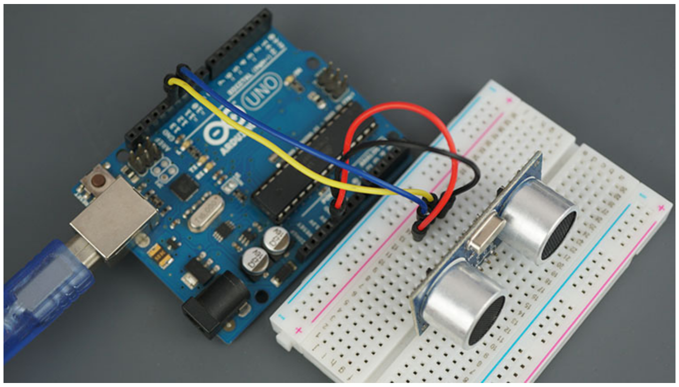
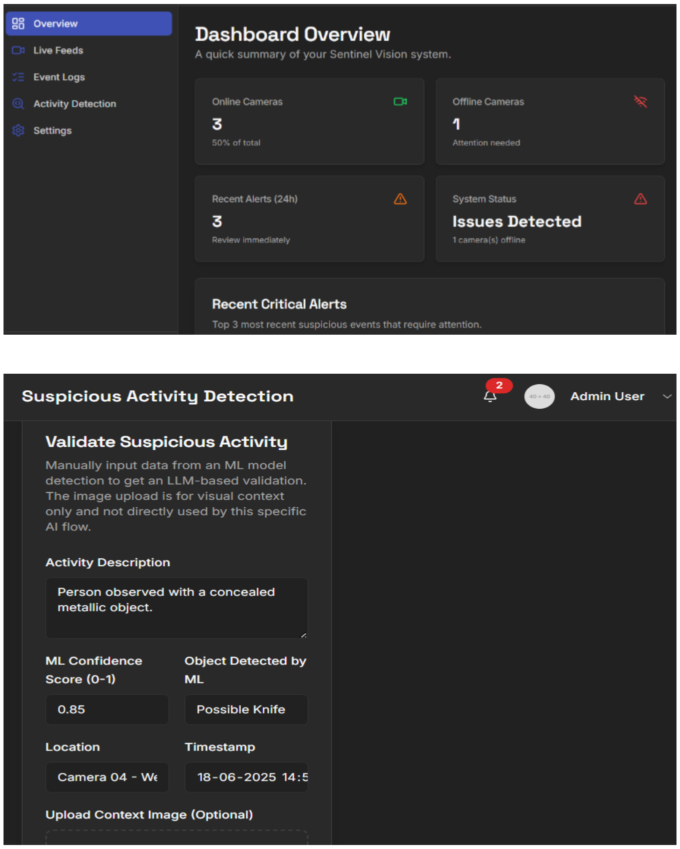


Fig. 6 IoT Acoustic Sensor Connection Diagram and Prototype Model

F. User Interface for Application users and Data Validation

The user interface (UI) is developed to provide an interaction between users and the surveillance setup. This UI is designed to be data validating and responsive, giving real-time access to live camera feeds, alerts, and analytical insights through dashboards and validating suspicious activity. It boasts features like personalized camera views, event-triggered video playback, and remote device management, allowing users to keep an eye on their surroundings efficiently from both desktop and mobile devices. Additionally, strong data validation processes are in place to guarantee the accuracy and integrity of input data, including sensor readings, ML-generated alerts, and user inputs. These validation measures help to avoid false positives or duplicate data from impacting system performance, ensuring that only verified, meaningful information feeds into analytics and decision-making. Dashboards used to monitor the recent alerts, cameras that in online or offline, active camera feeds monitoring different locational cameras like in parking slot, main entrances and through validating suspicious activity, data can be feed in file format with location enabled for real-time visualization.



Dashboard Overview

A quick summary of your Sentinel Vision system.

Online Cameras
3
50% of total

Offline Cameras
1
Attention needed

Recent Alerts (24h)
3
Review immediately

System Status
Issues Detected
1 camera(s) offline

Recent Critical Alerts

Top 3 most recent suspicious events that require attention.

Suspicious Activity Detection

Validate Suspicious Activity

Manually input data from an ML model detection to get an LLM-based validation. The image upload is for visual context only and not directly used by this specific AI flow.

Activity Description: Person observed with a concealed metallic object.

ML Confidence Score (0-1): 0.85

Object Detected by ML: Possible Knife

Location: Camera 04 - We

Timestamp: 18-06-2025 14:5

Upload Context Image (Optional)

Fig.7 User Interface Dashboard and Validation Activity for Application users

V. SYSTEM REQUIREMENTS

The suggested framework for system requirements involves an advanced CCTV system that integrates Machine Learning (ML) and the Internet of Things (IoT), with the combination of ML-based modules, IoT integrated sensors, and mobile applications that work together for real-time surveillance, smart analysis, and effective data management. On the hardware systems, IP cameras equipped withedge computing capabilities, along with IoT sensors like acoustic, motion, and thermal detectors. Processing units such as Raspberry Pi or NVIDIAdevice modules must be integrated as an edge device. Reliable 5V control power supply with backup options to ensure continuous monitoring. For software systems, YOLOv8 ML algorithms for tasks like object detection, facial recognition, anomaly detection, and behavior analysis, that are usually trained and need to be deployed using popular frameworks like TensorFlow, Pytorch, or OpenCV. Cloud module for scalable storage, remote access, and keeping the models up to date, while secure APIs maintains the edge devices communicate with cloud platforms. The networking makes the overall system reliable with a high-speed internet connection or LAN with enough bandwidth for streaming and data transfer is essential, supported by protocols like MQTT or HTTP. Cybersecurity protocols like encryption, firewalls, and access controls made vital for safeguarding sensitive video and sensor data to prevent the module from any physical or network threats.A dashboard or mobile interface app uses dashboard and validate suspicious activity for real-time alerts, validating usual and unusual activities, fetching data with time and location, remote monitoring, and managing the system, making the whole systems efficient, smart, secure and functional.

VI. EXPERIMENTAL RESULTS

A. *Confusion Matrix*

This confusion matrix illustrates model’s classification performance with true and false positives and true and false negatives. Binary classification with two classes “No Weapon” and “Weapon” and labels with predicted labels.

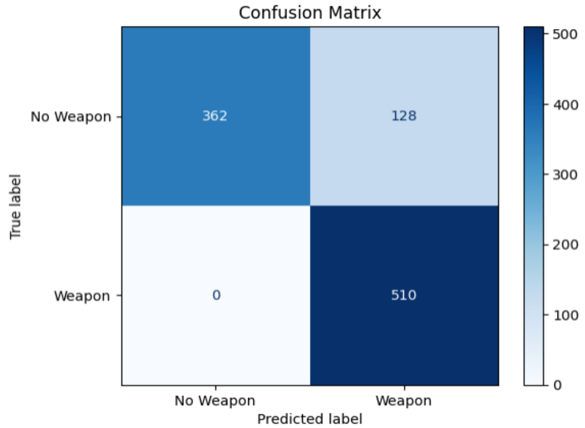
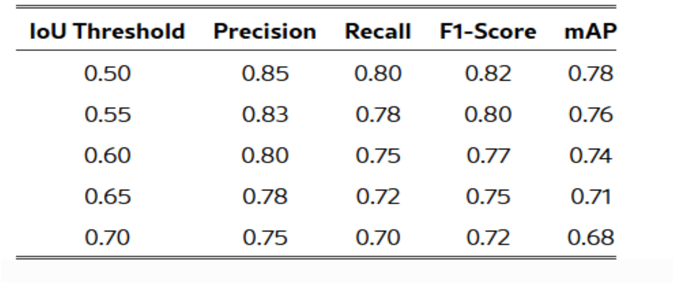


Fig 8.1 Confusion Matrix

B. *Performance Metrics of YOLOv8 Model*

This table illustrates that the YOLOv8 model shines at lower IOU thresholds (0.50–0.55), which makes sense for real-world object detection tasks where a bit of localization is often acceptable. However, as the evaluation criteria tighten up, the model's knack for accurately pinpointing objects takes a slight hit, which is pretty standard for most object detectors.



IoU Threshold	Precision	Recall	F1-Score	mAP
0.50	0.85	0.80	0.82	0.78
0.55	0.83	0.78	0.80	0.76
0.60	0.80	0.75	0.77	0.74
0.65	0.78	0.72	0.75	0.71
0.70	0.75	0.70	0.72	0.68

Table shows Performance Metrics of YOLOv8 Model

C. Training and Validation Loss

This graph shows training loss and validation loss over epochs during the training of a machine learning model. The x-axis consists of epochs with the number of times the learning algorithms worked through training entire dataset and y-axis denotes the loss value, a measure of error. The provided plot lines show the training loss and validation loss of the YOLOv8 model over epochs. Losses steadily decreasing and stabilizing shows the effective model learning and good training performance.

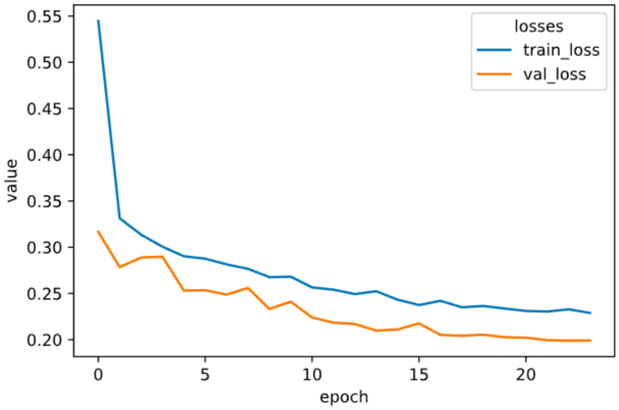


Fig 8.2 Training and Validation Loss

D. Precision Recall Curves

The precision-recall curve is a crucial metric for understanding the trade-offs between precisionand recall at various thresholds. Figure 3 illustrates the precision-recall curve for the YOLOv8 model, showcasing its performance across different confidence thresholds.

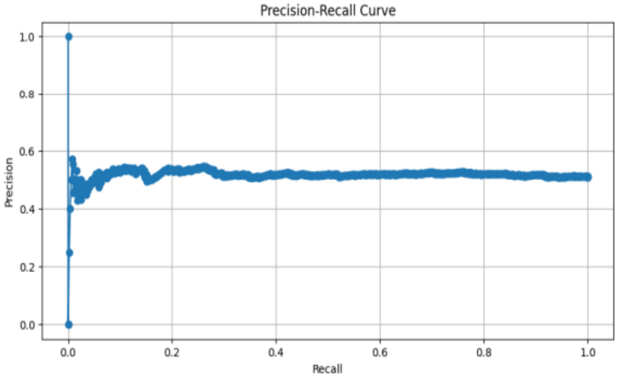
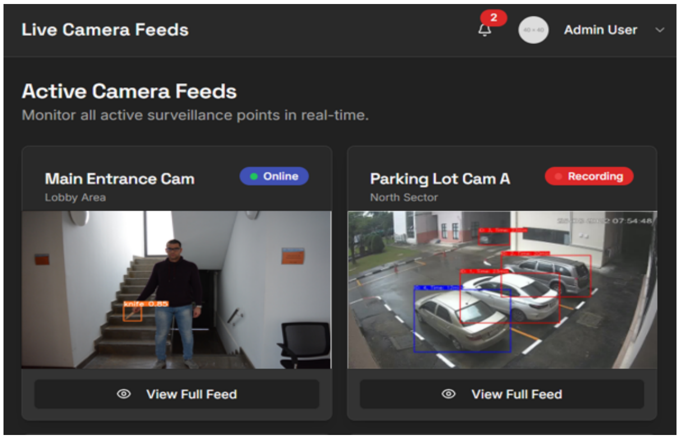


Fig. 8.3 Precision-Recall Curve



Live Camera Feeds

Active Camera Feeds

Monitor all active surveillance points in real-time.

Main Entrance Cam — Online — Lobby Area

Parking Lot Cam A — Recording — North Sector

Fig 8.4Live Action Camera Feeds

VII.FUTURE SCOPE

The Future of Advanced CCTV using Machine Learning (ML) and the Internet of Things (IoT) in cutting-edge CCTV systems is paving the way for the future of surveillance, public safety, and smart infrastructure. For instance, edge computing will allow for real-time analytics right on the devices. By using federated and continual learning models, these systems will be able to adapt over time while keeping data private. Multi-modal sensor fusionmerging video with sound, heat, and environmental datawill improve context-aware detection. Predictive surveillance techniques could even anticipate unusual behavior before incidents happen, providing proactive security measures. Cloud-IoT architectures will facilitate scalable and efficient data management, especially in smart cities. Privacy and cybersecurity will become even more crucial preventing device from cyber threats. Future systems will also prioritize energy-efficient designs and better integration with other smart city elements, all while fostering effective collaboration between humans and AI to support, rather than replace, security personnel. All these advancements point to a bright and impactful future for intelligent CCTV surveillance.

A. *Edge Computing and Real-Time Analytics*

- With advanced edge computingfor quicker data processing right at the camera.
- Ability to perform real-time object recognition, behaviour analysis, and anomaly detection without having to rely entirely on cloud computing.

B. *Federated and Continual Learning*

- Enhanced federated learning models, CCTV systems can enhance their accuracy over time.
- Techniques for continual learning will enable these systems to adjust to new situations, like changes in lighting, different human behaviour, or emerging threats, all without needing to completely retrain the model.

C. *Predictive Surveillance and Behaviour Forecasting*

- Using predictive analytics powered by machine learning, CCTV systems can anticipate potential incidents before.
- This includes spotting unusual movement patterns, recognizing suspicious behaviours, and alerting authorities in advance to boost public safety.

D. *Scalable Cloud-IoT Architecture*

- A flexible architecture that merges cloudand IoT technologies will facilitate high-volume video streaming and data storage across extensive deployments.
- This setup will be crucial for managing urban surveillance in smart cities, transportation hubs, and large-scale events.

E. *Enhanced Data Privacy and Cybersecurity*

- Future systems must address growing concerns about data security and privacy.
- End-to-End encryption, and secure IoT protocols will be crucial in building trust and regulatory compliance in surveillance.

VIII. CONCLUSION

The Advanced CCTV System that leverages Machine Learning and the Internet of Things brings a change in world of security surveillance. By combining ML-driven video analysis, IoT-enabled sound detection, and automated threat response, this system takes real-time monitoring and security efficiency to a whole new level, all while proactively preventing potential threats. With the power of cloud computing, edge processing, and predictive analytics, it ensures quicker response times, fewer false alarms, and greater accuracy in surveillance. Even though there are hurdles like data privacy issues, hardware constraints, and the costs of initial setup, ongoing improvements in AI model training, network reliability, and cybersecurity measures will keep enhancing its performance. As security threats continue to evolve, the system's ability to adapt, learn, and incorporate new technologies guarantees that these next-gen surveillance systems will stay smart, effective, and absolutely essential.

REFERENCES

[1] Gun Violence Archive. Accessed: Apr. 15, 2021. [Online].https://www.gunviolencearchive.org/

[2] G. F. Shidik, E. Noersasongko, A. Nugraha, P. N. Andono, J.Jumanto, and E. J. and Kusuma, ‘‘A systematic review of intelligence video surveillance: Trends, techniques, frameworks, and datasets,’’ IEEE Access, vol. 7, pp. 457–473, 2019.

[3] Abdelmoamen, ‘‘A modular approach to programming multi- modal sensing applications,’’ in Proc. IEEE Int. Conf. Cogn. Comput. (ICCC), 2018, pp. 91–98, doi: 10.1109/ICCC.2018.00021.

[4] K. He, G. Gkioxari, P. Dollár, and R. Girshick, ‘‘Mask R-CNN,’’ in Proc. IEEE Int. Conf. Comput. Vis., Apr. 2017, pp. 2980–2988.

[5] S.-C. Huang, ‘‘An advanced motion detection algorithm with video quality analysis for video surveillance systems,’’ IEEE Trans. Circuits Syst. Video Technol., vol. 21, no. 1, pp. 1–14, Jan. 2011.

[6] A. A. Moamen and N. Jamali, ‘‘Opportunistic sharing of continuous mobile sensing data for energy and power conservation,’’ IEEE Trans. Services Comput., vol. 13, no. 3, pp. 503– 514, May/Jun. 2020, doi: 10.1109/TSC.2017.2705685.

[7] Indrahayu, R. Y. Bakti, 1. S. Areni, and A. A. Prayogi, "Vehicle detection and tracking using gaussian mixture model and kalman filter," in Pro ceedings of the International Conference on Computational Intelligence and Cybernetics. 2016. pp. 115-119.

[8] "Kaggle. Machine learning and data science community, accessed Apni 15, 2021. [Online]. Available: https://www.kaggle.com/

[9] A. Dutta and A. Zisserman, The via annotation software for images. audio and video." in Proceedings of the 27th ACM International Con ference on Multimedia, ser. MM 19, 2019, pp. 76-79.

[10] "Keras: A python deep learning api, accessed 2021. [Online]. Available: https://keras.io/

[11] Opencv: A python library for real-time computer vision accessed April 15. 2021. [Online]. Available: https://pypi.org/project/opency-python/

[12] Abdelmoamen and N. Jamalı, "A model for representing mobile Jistributed sensing-based services." in Proceedings of the IEEE Interna nonal Conference on Services Computing, set. SCC 18, San Francisco, USA, 2018, p. 282-286.

[13] A. Moamen and N. Jamali, "ShareSens: An approach to optimizing energy consumption of continuous mobile sensing workloads," in Pro ceedings of the 2015 IEEE International Conference on Mobile Services (MS 15), New York, USA, 2015, pp. 89-96

[14] Abdelmoamen. D. Wang, and N. Jamali. "Approaching actor-level resource control for akka in Proceedings of the IEEE Workshop on Job Schedding Strategies for Parallel Processing, ser. JSSPP 18 Vancouver, Canada. 2018. pp. 1-15.

[15] A. A. Moamen and N. Jamali, "CSSWare: A middleware for scalable mobile crowd-sourced services," in Proceedings of MobiCASE, Berlin. Germany, 2015, pp. 181-199.