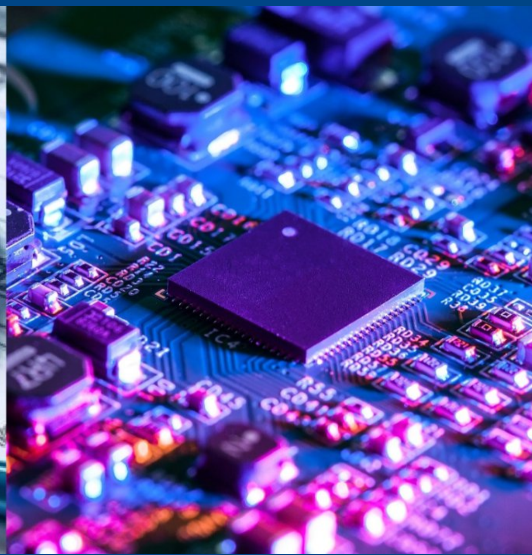
ISSN: 2582-7219

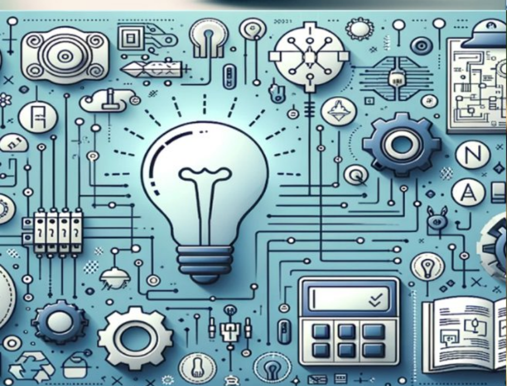




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904234

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

A Deep Learning–Driven Software Framework for Proactive Multi-Modal Surveillance

Mohamed Niyaz A[1], Vishwa B[2], Mohammad Hassan L[3], Pavithran P[4], and X. Birla[5]

Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, Chennai, Tamil Nadu, India[1]

Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, Chennai, Tamil Nadu, India[2]

Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, Chennai, Tamil Nadu, India[3]

Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, Chennai, Tamil Nadu, India[4]

Assistant Professor, Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, Chennai, Tamil Nadu, India[5]

ABSTRACT: Urban surveillance systems face critical limitations in real-time threat detection because camera performance decreases during low-light situations and cloud-based processing introduces delays. The paper introduces a hybrid edge-cloud framework that uses LiDAR–camera sensor fusion to create a secure and effective threat detection system for smart city environments. The framework uses vision models to perform semantic threat classification, while it uses LiDAR technology for tracking purposes in all light conditions. Edge devices conduct on-site multi-modal perception, while a geo-spatially aware, multi-tenant cloud layer handles emergency dispatch according to jurisdictional requirements. The proposed architecture has been designed to achieve effective detection performance during difficult lighting conditions while it maintains low latency for real-time system operations and can handle multiple simultaneous alerts. The document provides the system design together with the fusion methodology and an extensive evaluation plan, which includes performance targets that will be used for future experimental validation.

I. INTRODUCTION

The current urban surveillance systems operate as forensic tools because they document incidents for investigators to analyze after an event has occurred, which results in their ability to detect threats during active incidents. Two significant challenges exist for automated video stream detection via computer vision and deep learning technology, which stem from decreased performance under low-light and blocked visibility circumstances and from cloud- based AI systems needing time to process data.

The detection accuracy of camera-only systems at night and during foggy weather and in low-contrast situations [1, 4, 29], shows significant decline. The urban areas experience their highest criminal activity during these times, which makes vision-only surveillance systems ineffective for constant urban security monitoring. The AI systems that operate from cloud-based resources create two problems through their data transfer requirements and their processing times, which makes them unsuitable for urgent emergency operations that need immediate responses [14, 16].

The authors present an edge-cloud framework, which integrates LiDAR and camera sensors at edge locations with a multi-tenant cloud dispatch system that handles geo-spatial data. The research presents four main contributions through its development of a multi-modal sensor fusion architecture, which integrates YOLO-based 2D detection with LiDAR 3D point clouds through Kalman filtering; its development of a privacy-preserving edge processing strategy, which reduces biometric data transmission; its development of a multi-tenant cloud dispatch system, which operates through schema-per-tenant PostgreSQL and PostGIS routing; and its development of a comprehensive validation methodology, which defines detection accuracy and latency and scalability testing criteria.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. RELATED WORK

2.1 YOLO-Based Weapon Detection Systems

The YOLO (You Only Look Once) family of object detection algorithms has become the de facto standard for real-time weapon detection in surveillance systems due to their balance of accuracy and inference speed. The researchers Fernandez-Testa et al. [1], achieved 0.87 mAP at 4.43 FPS with custom dataset of 16,799 weapon images by deploying YOLOv5s that they optimized using TensorRT on edge devices for armed robbery detection. The distributed IoT architecture sends detected frames to a cloud-based 3DCNN classifier which achieves 0.88 accuracy to decrease false positive rates while demonstrating how edge-cloud partnerships create operational benefits.

The team of Ahmed et al. [29], created Scaled-YOLOv4 models which achieved 92.1% mAP at 85.7 FPS performance on RTX 2080Ti GPUs and they deployed optimized models successfully on Jetson Nano edge devices which operated with TensorRT framework. The researchers studied CPU and GPU platforms to determine which deployment methods functioned best during real-time operational conditions. Abins et al. [4], developed their YOLOv8 and PELSF-DCNN classifiers which used CSBO feature selection to achieve remarkable accuracy rates because they produced very few false-positive results without disclosing their precise measurement results.

The study by Choudhry et al. [31], conducted an extensive review of machine learning techniques used for detecting unusual activities in surveillance video footage. The researchers discovered that deep learning has improved detection capabilities yet single- modality video data causes detection systems to fail under changing environmental conditions which explains why multi-sensor systems are needed.

The researchers Singh et al. [6], used YOLO v4 to develop a weapon detection system for IoT devices which achieved 70% accuracy with low-quality video but improved to 95% with high-quality streams. The researchers Eberhardt et al. [23] used YOLOv8 on Raspberry Pi edge devices to create a gun and knife detection system which enhanced surveillance capabilities while the researchers Berardini et al. [24] analyzed YOLOv8 performance for edge AI video-surveillance systems through benchmark tests.

Critical Limitation: The performance of all reviewed YOLO-based systems shows excellent results in controlled conditions but they confront serious accuracy challenges when operating under low-light conditions. Camera-only methods face fundamental limitations since their detection performance drops below 45% during nighttime conditions [1, 4, 29]. The system functions as a critical security vulnerability which affects urban surveillance operations that run continuously throughout the day

2.2 LiDAR Crowd Monitoring Systems

The combination of LiDAR (Light Detection and Ranging) sensors with cameras enables users to obtain precise three-dimensional spatial data which operates successfully in all lighting situations. Torres et al. [2], deployed Velodyne VLP-16 LiDAR sensors with NVIDIA Jetson Nano and Xavier NX edge nodes for pedestrian detection in Aveiro, Portugal, achieving quasi-real-time processing with detection ranges up to 15 meters. The system employs SVM learning techniques for its "Neuron process" MEC application, which links edge publishers to smart-city platform consumers through the MQTT protocol.

Miramá et al. [32], conducted a survey of machine learning in pedestrian localization systems, identifying that while ML improves accuracy, significant "open issues" remain regarding occlusion and environmental noise in single-sensor setups. The researchers discovered that successful urban area localization needs multiple sensor types because single sensors cannot provide enough information.

Jiang et al. [7], proposed an Ultra Large-Scale Crowd Monitoring (ULCM) system architecture which uses UAV agents together with airborne LiDAR sensors and ground sensory systems that include CCTV and infrared gas and ultrasonic sensors. The design uses software-defined networking (SDN) together with edge cloud technologies to achieve better performance in real-time data collection and analysis for multimodal and multiperspective crowding data although the research did not provide any experimental evaluation.Yu et al. [8], developed an Intelligent Visual-IoT system for real-time 3D visualization of outdoor scenes using fixed and airborne cameras, infrared cameras, and LiDAR. The system transmits multimodal data to cloud-based AI algorithms which use the data to calculate object locations and detect abnormal events and the results get sent to terminal devices for 3D animation and user interaction.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904234

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Chakraborty et al. [5], designed a LiDAR-assisted large-area video surveillance network which uses a LoRa-based IoT framework to conduct multi-sensor data fusion and video analysis through transfer learning. The architecture enables organizations to expand their operations across vast regions while maintaining cost-effective storage and processing capabilities, but the study did not include specific performance metrics.

Critical Limitation: The spatial localization capabilities of LiDAR remain intact because it functions without any lighting requirements, but the technology cannot differentiate between various threat categories which include weapons and phones through semantic classification. The system can detect and track objects through LiDAR-only technology, but it fails to identify threats because it lacks essential contextual knowledge [2, 5, 7, 8].

2.3 Edge Computing for Surveillance

The edge computing frameworks solve the latency and bandwidth issues which affect cloud- based surveillance systems by establishing inference points near where data is generated. The researchers in Dautov et al. [9], introduced the Metropolitan Intelligent Surveillance System (MISS) framework which combines IoT technology with cloud computing, edge computing, and big data systems to achieve facial recognition results that exceed standard cloud processing by eight times through the use of distributed edge device processing.

The process of using machine learning in networking operations creates multiple difficulties that must be addressed. The research by Ridwan et al. [33], examined machine learning applications in networking and found that efficient edge-cloud communication protocols serve as the vital component which enables real-time applications to meet their strict latency requirements. The researchers predict that organizations will face new challenges which require them to find ways to balance between edge processing requirements and delays caused by network data transmission.

The researchers Chen et al. [14, 16], established a distributed real-time object detection framework through edge-cloud collaborative efforts by implementing YOLOv3 on NVIDIA Jetson Xavier NX devices while using Intel Core i7-9700k with GeForce RTX 2080 Ti for cloud-based model training. The system achieved almost 94% bandwidth savings by conducting local inference on edge devices that sent only uncertain results to the cloud for validation while receiving updates through remote model delivery.

The researchers Ardiansyah et al. [3], developed EagleEYE which serves as an aerial edge- based disaster relief response system that operates at the OPTUNS edge data center with its 48-core CPU, 128GB RAM, and NVIDIA Tesla V100 GPU. The researchers used YOLOv3 with a scalable worker architecture which resulted in 90% decreased inference latency and 87% detection accuracy while each worker used 4.7% of the GPU memory (1.515 GB of 32 GB).

The researchers Rajavel et al. [20], designed the Cloud-based Object Tracking and Behavior Identification System (COTBIS) which implements edge computing at its gateway to detect abnormal falling activities through background subtraction and deep convolutional neural networks in smart healthcare surveillance systems. The system architecture reduces both network bandwidth and response time requirements for wireless camera connections to cloud servers.

The researchers Kumar et al. [11], enhanced smart city safety through their AI expert system which combines image-to-image stable diffusion with YOLO v7 for violent item detection (89.5% mAP) and MediaPipe with LSTM for action classification (88.33% accuracy), which runs on dash cameras deployed at edge devices for real-time testing.

Critical Limitation: The current edge surveillance systems depend on cameras as their only sensor type while they do not support any methods which combine multiple sensor data. The systems experience reduced latency because edge processing operates closer to data sources, but their security suffers from the inability to merge information from different sensor types which leads to system failures based on sensor-specific defects. [9, 11, 14, 16, 20].

2.4 Multi-Tenant Cloud Architectures for Emergency Response

The authors of Zhang et al. [10], designed a public safety alert system which uses IoT technology to provide real-time emergency response capabilities. The system achieved operational goals through an alert processing time of 450 milliseconds detection accuracy of 95 percent and capacity to handle 12000 simultaneous connections while maintaining a 99.8 percent system availability across four emergency situation The researchers Thota et al. [22], used IoT technology to detect emergency vehicles in smart cities but they did not share information about their system architecture. The researchers Cadet et al. [13], developed an AI-based threat detection system which combines machine learning and deep learning with stream processing and edge computing to solve network congestion problems but they did not conduct any tests to validate their system.

The researcher Cheedalla [17], studied Predictive Mobile AI as a tool to change emergency response from reactive methods to preventive strategies while he examined neural network models for early warning detection which use multimodal data streams from satellites and sensors and social media. The research did not provide information about its practical execution methods.

Critical Limitation: While multi-tenant cloud architectures demonstrate scalability and reliability, existing emergency response systems lack geo-spatial routing mechanisms for jurisdiction-aware alert dispatch. Most systems broadcast alerts uniformly rather than intelligently routing to appropriate agencies based on incident location [10, 13, 17, 22].

2.5 Positioning Our Solution

The proposed system intends making four key innovations to address the noted gaps in operational effectiveness and efficiency.

Sensor Fusion at the Edge: The proposed framework achieves early fusion by integrating data from both camera systems [1, 4, 6, 23, 29], and LiDAR systems [2, 5, 7]. Kalman filtering is used in this process to combine YOLO-based 2D bounding boxes with LiDAR-derived 3D point clusters. This fusion method was developed to enhance detection performance across different lighting conditions while preserving system ability to detect semantic threats.

Privacy-Preserving Architecture: The framework uses LiDAR technology for spatial tracking without biometric data while cameras perform semantic classification according to edge-computing principles that were established in previous research [9, 11, 14, 16, 20]. The architecture conducts local perception while sending out only structured alert data to protect personally identifiable information (PII) which public surveillance systems must safeguard at all times.

Geo-Spatial Cloud Routing: The proposed system uses PostGIS Point-in-Polygon queries to enable jurisdiction-aware routing to respond to emergency situations through its design based on earlier multi-tenant emergency response frameworks [10, 13, 17, 22]. Agencies should receive incidents based on their geographic location to improve emergency response efficiency and operational relevance instead of sending out alerts to everyone simultaneously.

Comprehensive Validation Methodology: The research introduces an organized testing framework which scientists will use to evaluate detection accuracy in different light situations while they also measure complete system delay times and test system performance with numerous edge devices operating together. The validation framework for this study enables researchers to test multiple performance metrics simultaneously which previous research studies did not examine together [4, 5, 7, 13, 17, 18, 27Amendola et al. [19], presented an edge-distributed camera–LiDAR fusion approach for accurate moving-object localization, reporting mean errors of 5 cm in indoor settings and 30 cm on the outdoor KITTI 360 dataset, while maintaining low-latency performance. Their results highlight the strengths of sensor fusion at the edge, but the work is primarily oriented toward autonomous navigation. It does not address threat detection scenarios or incorporate multi-tenant cloud-based dispatch mechanisms for emergency response. In contrast, the presented functionality does further apply these fusion principles of urban threat detection and integrates them with a geospatially aware emergency dispatch layer.

III. SYSTEM ARCHITECTURE

The system design uses a dual-layer structure which includes (1) an Edge Perception Layer dedicated to immediate multi-modal threat detection through privacy-preserving methods, and (2) a Cloud Dispatch Layer which enables geo-spatial routing and notification services for multiple tenants. The total system architecture is demonstrated through the conceptual design presented in Figure 1.

Edge Perception Layer

3.1.1 Hardware Setup

The edge perception node operates as an embedded multi-sensor platform which includes three essential components.

The **NVIDIA Jetson Nano** serves as the target edge inference device. The device contains a 128-core Maxwell GPU

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

which can achieve 472 GFLOPS of processing power while using minimal energy. The platform utilizes Ubuntu as its operating system together with NVIDIA JetPack and CUDA and TensorRT acceleration libraries which enable efficient deep neural network execution for real-time perception tasks on devices with limited resources [1, 15, 29].
The Logitech C922 Pro Stream webcam **RGB camera** captures visual semantic information. The device operates at a resolution of 1920×1080 pixels at 30 frames per second and incorporates autofocus and low-light correction features. The camera provides the necessary visual detail for semantic threat classification tasks [12].

To enhance vision performance during dim lighting conditions, the system integrates a **Velodyne VLP-16 LiDAR sensor**, which delivers complete 3D spatial data collection that does not depend on ambient light levels. The VLP-16 offers a 360-degree horizontal field of view and a vertical field of view of ±15 degrees across 16 laser channels, enabling reliable depth perception in both bright and near-dark environments [2, 19].

The sensors undergo calibration procedures through checkerboard-based camera– LiDAR calibration methods. The extrinsic parameters which result from this process enable accurate geometric projection from 3D LiDAR coordinates to 2D image coordinates through a rotation matrix (R) and a translation vector (t) [19, 30]

3.1.2 Sensor Fusion Algorithm: Early Fusion with Kalman Filtering
The perception framework uses early feature-level fusion to combine 2D semantic detections from the camera with 3D spatial information from LiDAR point clouds. The fusion pipeline consists of four distinct processing stages.

Stage 1: Camera-Based Threat Detection
The object detection models that use YOLO technology will undergo optimization through TensorRT to enable real-time inference capabilities on Jetson Nano devices. The model is trained on a composite dataset including weapon imagery and augmented low-light samples. The detector generates 2D bounding boxes for each frame at time (t):

$$B_{cam}(t) = \{b_i = (x_i, y_i, w_i, h_i, c_i, s_i)\} \quad (1)$$

The box center is represented by (x_i, y_i) while (w_i, h_i) denotes width and height. The predicted class is indicated by c_i and s_i stands for the confidence score. Detections with low confidence may be filtered to reduce false positives [1].

Stage 2: LiDAR Point Cloud Processing
The VLP-16 sensor produces a 3D point cloud at each time step:

$$P_{lidar}(t) = \{p_j = (x_j, y_j, z_j, r_j,)\} \quad (2)$$

RANSAC may be used to remove the ground plane from the point cloud data before Euclidean clustering divides the remaining data into different object groups which results in clusters:

$$C_{lidar}(t) = \{C_k\} \quad (3)$$

This process enables spatial grouping of points likely belonging to individual objects [19].

Stage 3: 2D–3D Association via Projection
The projection of LiDAR points into the image plane is done through the calibrated intrinsic matrix (K) and extrinsic parameters (R, t) which correspond to each camera bounding box b_i:

$$p_{img}(t) = K[R|t]\, P_{lidar} \quad (4)$$

The bounding box b_i identifies LiDAR clusters from projected points which meet the requirement of appearing within the bounding box area on the image. This process produces fused detections:

$$F(t) = \{(b_i, C_k)\} \quad (5)$$

Stage 4: Kalman Filter Tracking and Velocity Estimation
The fused detection undergoes tracking through a Kalman filter to establish continuous time links. The state vector is defined as:

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

$$x = [x, y, z, v_x, v_y, v_z, w, h, d]^T \tag{6}$$

The 3D centroid of the LiDAR cluster is represented by (x, y, z) while the velocity components are described through (v_x, v_y, v_z) and the bounding box dimensions are depictedthrough (w, h, d). The system achieves velocity estimation and continuous tracking through the interaction of prediction and update steps [19].

This fusion strategy is designed to allow LiDAR to provide robust 3D localization under poor lighting, while the camera contributes semantic classification capabilities that LiDAR alone cannot provide.

3.1.3 Privacy Engine: LiDAR for Non-Biometric Tracking

The system design approach establishes privacy protection as its fundamental design requirement. LiDAR point clouds enable spatial tracking without capturing biometric identifiers such as facial features or clothing details [2], [26]. The privacy-oriented processing pipeline is designed as follows.

The system needs to use LiDAR clusters for its primary tracking and motion estimation tasks instead of relying on continuous video footage. The RGB camera may be activated selectively when high-confidence threat detections occur to record short verification clips rather than continuous footage. All sensor fusion and inference processes are designed to execute locally on the Jetson Nano. The system will transmit to the cloud only structured alert metadata which includes timestamp and location and threat type and confidence information while preventing the transmission of raw video and point cloud data [14] [16] [26].

3.1.4 Technical Details: Mapping Bcam to Plidar for Volumetric Velocity Derivation

The fusion process requires exact geometric mapping which connects camera detections with LiDAR points. The image region which corresponds to bounding box ($b_i = (x_i, y_i, w_i, h_i)$) can be defined through the following equation:

$$R_i = \{(u,v) \mid x_i - w_i 2 \le u \le x_i + wi/2,\ y_i - h_i/2 \le v \le y_i + h_i/2\} \tag{7}$$

The image coordinates of LiDAR points are determined through the following transformation:

$$[uj, vj, 1]^T = K[R|t][xj, yj, zj, 1]^T \tag{8}$$

A cluster C_k is associated with bounding box b_i if the fraction of its projected points within R_i exceeds a predefined threshold [19].

The centroid of cluster C_k is computed as:

$$c_k(t) = \frac{1}{|C_k|} \sum_{p_j \in C_k} p_j \tag{9}$$

Temporal differentiation enables the calculation of velocity:

$$v_k(t) = \frac{c_k(t) - c_k(t-1)}{\Delta t} \tag{10}$$

The term (\Delta t) represents the period which separates two frames. The bounding box dimensions function as volumetric size estimators through the equation:

$$V_k = (x_{max} - x_{min})(y_{max} - y_{min})(z_{max} - z_{min}) \tag{11}$$

The volumetric velocity representation enables threat motion assessment and trajectory analysis [19].

Cloud Dispatch Layer

The cloud dispatch layer functions as a scalable multi-tenant system which enables both jurisdiction-based routing and notification delivery. The system supports deployment on cloud platforms which include AWS EC2 instances while PostgreSQL and PostGIS enable geo-spatial database operations.

3.1.5 Multi-Tenancy Implementation: Schema-Per-Tenant in PostgreSQL

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The architectural design uses a schema-per-tenant approach to create separate data access pathways which protect agency data from unauthorized access. The system uses dynamic model binding as shown in **Listing 3.1** to enable users to choose the correct tenant schema during runtime while the application functions with its common operational framework.

Listing 3.1. Dynamic tenant model loader function

```
const getTenantModel = (tenantSchema, modelName) => { const modelDef = require(`../models/${modelName}`);
return modelDef(sequelize, DataTypes, tenantSchema);
};
```

The design permits authenticated users to access their specific tenant information while keeping the public schema accessible for all users to view shared reference data.

Role-Based Access Control and Emergency Identity Assistance

The proposed system uses Role-Based Access Control (RBAC) because it provides a secure method for users to interact with the platform. The system provides three main user roles which include **Administrator**, **User**, and **Personnel (Responder)**. Administrators handle all system tasks which involve device control and user monitoring. Users who do not have administrative rights can use the mobile and web applications to receive alerts and provide information and create personal accounts. Emergency response and security personnel use operational dashboards to view current incidents and their corresponding response resources.

The system enforces access controls by establishing specific permission levels which depend on user roles. Administrators control all system settings while maintaining their right to access necessary field data. Users are restricted to viewing and handling their personal information only. Responders have permission to see emergency-related information which is needed for their response work but they cannot change any system settings. The system separation reduces risks of misuse while enabling organizations to track activities and achieve secure multi-user system operations which follow industry standards The system provides two security functions which include alert-based threat detection and emergency identity assistance through facial recognition technology. The system allows users to store their facial embeddings and essential identification details which include their name and emergency contacts and medical notes after they give permission during the registration process. Emergency responders who need to identify an unconscious person can use a camera device to scan for facial features. The system utilizes DeepFace to conduct facial embedding matching with registered profiles and obtain corresponding emergency information.

The emergency response support feature operates through controlled access which restricts its use to authorized personnel who need to respond to emergency situations. The design establishes privacy rights through facial data enrollment which users voluntarily give and secure storage of their data and emergency identification access rights to their data. The system does not perform continuous public facial surveillance; instead, recognition operates through event-based triggers which restrict access to specific user roles. The approach establishes a framework which delivers fast emergency assistance while protecting ethical standards and privacy rights.

3.2.2 Geo-Spatial Routing: PostGIS Point-in-Polygon Queries

PostGIS stores jurisdiction boundary data in geometry columns while GiST indexing allows fast execution of spatial search operations. The spatial query used to identify the responsible agency is shown in **Listing 3.2**.

Listing 3.2. Spatial containment query for jurisdiction lookup

```
const responsibleAgency = await Jurisdiction.findOne({
where: sequelize.fn('ST_Contains', sequelize.col('boundary'), incidentLocation),
order: [['priority', 'DESC']], limit: 1
});
```

The routing system establishes a secure method for assigning incidents to the appropriate agency while protecting the system from potential SQL injection attacks [10].

3.2.3 Real-Time Notification System Architecture and Visualization

The proposed framework uses RTSP streams to obtain video input from surveillance cameras, which is processed at the edge device for local analysis. The AI model performs real-time inference on these streams, and structured alerts are generated only when potential threats are detected. The alert system transmits all essential content through different fields, which include timestamp data, device identifier information, threat category details, confidence score data, and estimated location, while preventing continuous cloud transmission of raw video content.

The edge system sends detected threat events to the backend server through protected API connections. The backend system handles all tasks related to alert record storage and user access control, while it also oversees the function of notification systems which operate between different services. The system uses Firebase Cloud Messaging (FCM) to distribute real-time alerts by sending push notifications to users who have permission to access mobile applications and web clients. The method enables notification delivery with low latency because it does not need ongoing streaming connections.

The centralized dashboard handles visualization for monitoring and situational awareness by retrieving processed alert information from the backend database. The interface system presents threat locations, timestamps, and classification results through an organized format, which enables operators to quickly assess and respond to incidents. Human operators can use visualization strategies which originate from machine learning–assisted visualization research principles, which focus on showing vital events while reducing the amount of information that operators have to process.

The backend system includes an escalation mechanism, which triggers automatic alerts to additional responders or supervisory personnel when alerts remain unacknowledged for a defined interval. The system guarantees that response activities will continue despite initial notification failures.

3.2.4 Database Schema Design for Logical Isolation

The Sequelize ORM framework is used for database manipulation, which establishes database models for Alert and Personnel and Incident Log data which utilize spatial and temporal indexing to enable both real-time and post-incident investigations. An example schema field definition is shown in **Listing 3.3**.

Listing 3. Geometry column definition for incident location

```
location: DataTypes.GEOMETRY('POINT', 4326)
```

The schema design enables multiple agencies to manage their data through a system that provides scalable and isolated and efficient query performance.

IV. METHODOLOGY

4 Proposed Evaluation Methodology and Expected Outcomes

This section describes the planned experimental methodology for validating the proposed multi-modal edge–cloud surveillance framework. The section establishes upcoming research activities by presenting the datasets and training methods together with the evaluation standards and performance objectives which will be used in future empirical testing.

Datasets and Experimental Setup

The multi-modal perception framework requires both visual and spatial datasets which represent actual urban safety situations for its training and evaluation processes.

Visual Dataset (YOLOv8 Training)

The object detection model will be trained using a composite dataset which consists of multiple publicly available sources including Roboflow datasets [36–40], and Kaggledatasets [41], and open-source repositories on GitHub [42]. The dataset contains more than 15,000 annotated images which span four essential categories for urban threat monitoring:

- **Weapon:** Handguns and rifles and knives
- **Fire:** Flames and smoke signatures
- **Anomaly:** Suspicious or unusual crowd behavior
- **Crowd Count:** Dense crowd presence estimation

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The research will implement data augmentation methods which include brightness changes, blurring effects, noise additions, and contrast modifications to create simulations of low-light and difficult visual environments. The training process for the project will use Kaggle GPU instances which operate in a cloud environment while the optimized model will be deployed to the Jetson Nano system for edge inference operations.

LiDAR Dataset:
The development of spatial perception components will utilize Point cloud data sorced from **IEEE DataPort** [43], was collected using frontal and rear 3D LIDAR (Velodyne VLP-16) at 20 Hz, and a frontal facing RGBD camera (Real Sense D435). The researchers employed fusion techniques to convert LiDAR 3D coordinates into camera 2D bounding boxes for accurate measurements according to the 3D point cloud and RGBD of pedestrians in robot crowd navigation: detection and tracking dataset. The dataset includes 250k frames of data which consist of over 200 minutes of recordings.

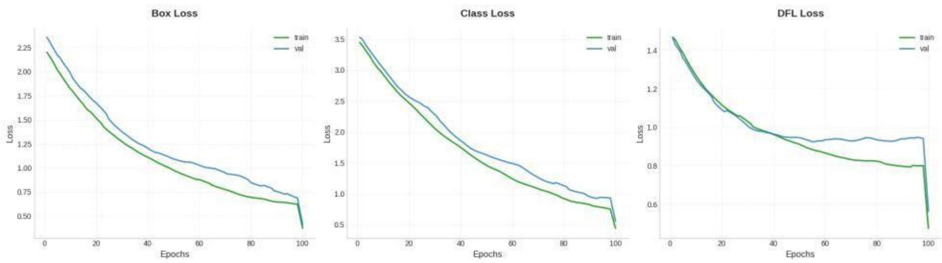
Experimental Environment:
The evaluation process will test two separate scenarios which include controlled indoor environments that allow researchers to alter light conditions and outdoor testing environments that replicate different movement patterns at various distances. The lighting conditions will test system performance at well-lit conditions which exceed 800 lux and at near-dark conditions which fall below 5 lux to create difficult visual environments.

YOLOv8 Training Analysis
The YOLOv8 object detection model will start its training process by using transfer learning from pretrained weights to achieve faster convergence and better generalization results. The training process will take place on cloud-based systems that use Kaggle GPU instances before the optimized model is moved to Jetson Nano for edge computing. The training process is expected to run for up to 300 epochs, with early stopping criteria considered if validation performance stabilizes.

The training pipeline will use extensive data augmentation methods to achieve strong learning results in all environmental conditions. Planned augmentation strategies include random brightness and contrast adjustments and Gaussian noise injection and motion blur simulation and horizontal flipping and scaling and random cropping. These augmentations create urban surveillance scenarios through which low-light conditions and motion and real- world camera variability can be simulated.

The training process will use YOLOv8 to monitor multiple loss components which will help determine when the system achieves convergence. The system uses three loss functionswhich include bounding box regression loss and classification loss and objectness loss. The analysis of training and validation loss curves will reveal divergence patterns and underfitting patterns and overfitting patterns. The optimization process will remain stable through the adjustment of learning rate scheduling and weight decay parameters. The Fig. 4.2.1 shows the graph on loss functions for YOLOv8 model we have trained for 50 epochs.

Fig. 4.2.1 . Crowd Density Around Platform



The model will undergo evaluation through standard object detection metrics which assess its performance. Detection precision and recall measurement will determine how reliable detection works and what percentage of detections were missed. The F1-Score will provide a balanced measure of detection performance while mean Average Precision works to evaluate localization and classification accuracy throughout various IoU thresholds.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The inter-class confusion analysis will create a normalized confusion matrix which focuses on how well the "Weapon" class distinguishes itself from mobile phones and tools and other similar-looking non-threatening items. The analysis process aims to uncover specific weaknesses in each class which require additional dataset balancing treatment.The analysis will focus on a validation data subset which has undergone synthetic low- light augmentation to test system robustness against visual degradation. The analysis will show how reduced illumination affects performance which proves vital for nighttime surveillance operations. The results obtained from these analyses will help in refining the training strategies and augmentation pipelines and class weighting schemes.

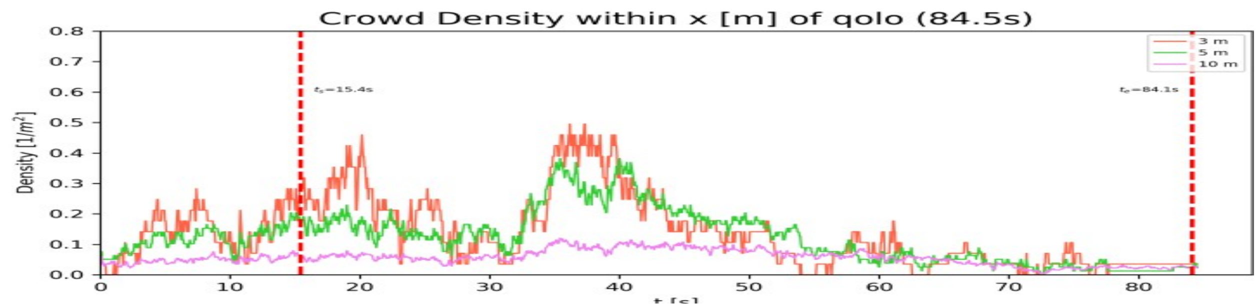
LiDAR Spatial Accuracy Evaluation

The LiDAR subsystem provides a system that can locate and track movement in space without needing light to function. The system uses its estimated object centroids from point cloud clusters to measure performance against established ground truth measurements from reference points The Root Mean Square Error (RMSE) in meters serves as the primary method for measuring spatial accuracy. The measurement process will assess how well sensors localize objects at three different distance categories, which include short, medium, and long. The evaluation will assess both static accuracy and tracking stability through time by studying how centroid positions and velocity estimates from the Kalman filter change over different time periods.LiDAR processing algorithms require development and validation support through public pedestrian tracking datasets which contain LiDAR point clouds and matching motion states and environmental information. The following figures are reproduced from the referenced dataset repository to illustrate the nature of the LiDAR-based motion and environment data. These images are included strictly as dataset examples and **do not represent results generated by the proposed system.**

Dataset Reference Visualizations from [43]

The dataset visualizations show the LiDAR dataset which offers information about dynamic motion and crowd interaction and spatial measurements as shown in Fig. 4.1 and Fig. 4.2 respectively. The data enables researchers to evaluate how well a system maintains tracking stability while estimating velocity and driving performance in crowded areas.The platform area experiences pedestrian density changes throughout different time intervals at various radial distances.The time series data displays pedestrian distances which reach their closest point during situations of crowd movement.

Fig. 4.1. Crowd Density Around Platform.



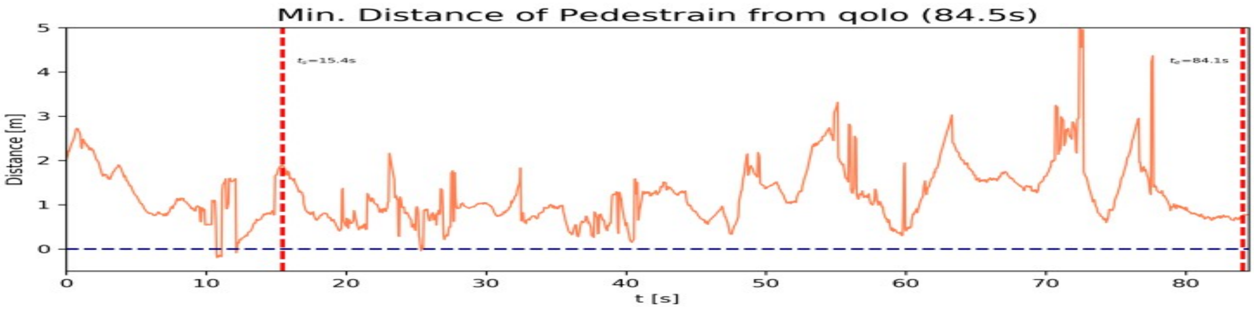


Fig. 4.2. Minimum Distance to Pedestrians.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904234

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The evaluation aims to confirm that
LiDAR-based tracking achieves accuratespatial measurements which remain dependable during times whencamera visual data becomes untrustworthy. The result proves that LiDAR works as a complementary technology to the fusion architecture'soperationalprocess.

Multi-Modal Performance in Low-Light

The proposed framework exists to support reliable threat detection under low visibility conditions since camera-only systems experience severe performance drops in these situations according to research [1] [4] [29]. The researchers will assess LiDAR–vision fusion capability through experiments which will occur in low-light environments that reach 5 lux or lower to simulate urban nighttime conditions.

The evaluation will compare three system configurations to isolate the contribution of each sensing modality:

- **Camera-Only Detection:** The YOLOv8 model processes RGB frames without LiDAR input. This configuration represents conventional vision-based surveillance systems.
- **LiDAR-Only Tracking:** Object clusters are extracted from LiDAR point clouds without semantic classification. This setup evaluates geometric tracking reliability independent of lighting.
- F**usion-Based Detection and Tracking:** The proposed early fusion pipeline combines camera-based semantic detections with LiDAR spatial clustering and Kalman filter tracking.

Experimental Setup

The facility will create low-light environments through its design which permits staff to control different lighting conditions. The lux meter will be used to measure lighting intensity in order to maintain consistent lighting conditions throughout the different testing periods. The testing process will involve:

- Stationary and moving human subjects
- Subjects carrying objects that may resemble weapons
- Occlusion scenarios where objects are partially blocked
- Varying distances between subjects and sensors

The researchers will conduct multiple test runs of each scenario to establish statistical reliability. The team will create ground truth annotations for object presence and class labels and spatial positions through manual preparation or reference marker usage.

Detection Performance Evaluation

The detection performance assessment will examine each configuration through three evaluation methods:

- **Precision:** Ratio of true threat detections to all detections
- **Recall:** Ratio of detected threats to total ground-truth threats
- **F1-Score:** Harmonic mean of precision and recall
- **mAP:** Mean Average Precision across IoU thresholds

The researchers will study performance decline patterns which happen at lower light levels to establish the point when camera-only detection becomes less dependable than fused detection methods. Tracking and Spatial Stability Metrics The evaluation of tracking performance will focus primarily on both LiDAR-only and fusion system configurations through the following methods:

- **Spatial RMSE:** Deviation between estimated object centroids and ground-truth positions
- **Track Continuity:** Percentage of frames where an object track remains consistently maintained
- **Velocity Stability:** Variance of estimated velocity over time for the same object

The metrics will determine whether LiDAR enables accurate motion estimation during periods when camera detection suffers from low light conditions.

Failure Case Analysis

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904234

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The analysis will focus on two main failure modes which include:

- Missed detections due to poor contrast
- False positives from visual noise
- Loss of track continuity during occlusion
- Misalignment between 2D detections and 3D clusters

The research team will analyze how sensor fusion technology operates across all three configurations to identify system detection failures.

Expected Outcomes

The research team expects that the evaluation results will demonstrate how LiDAR tracking performance remains constant while camera detection ability decreases under dim lighting conditions. The fusion-based system is therefore anticipated to maintain higher overall detection reliability and consistent spatial tracking by leveraging complementary sensing strengths. The results will help create a resilient perception system which operates effectively in low-light urban surveillance environments.

System Latency

The system will assess its operational delay by tracking the duration from edge device sensors when they capture data until the cloud dashboard displays alerts. The measurement of latency will be divided into three components which include:

- Time required for edge devices to perform inference and fusion operations
- Time needed for data to move across the network
- Time required for cloud systems to process data and perform geo-spatial routing functions

The testing process will use different network conditions to create scenarios that represent actual variations in cellular bandwidth throughout the testing period. The system supports emergency response operations by maintaining a maximum end-to-end latency of 200 milliseconds for its entire operational range

Scalability and Load Testing

The process of evaluating scalability will use simulated edge-device workloads which create alert traffic at progressively higher rates. The load testing tools will create a situation where they simulate hundreds to thousands of devices that send alerts to the cloud platform. The following metrics will be monitored during testing:

- Alert processing throughput (alerts per second)
- Packet drop rate (%)
- Cloud processing latency time
- Geo-spatial query response time

The architectural design enables city-wide deployments with more than 1000 simultaneous devices. It achieves this goal through its ability to deliver stable network performance and keep packet loss rates below acceptable limits. Assessing horizontal scaling methods will focus on how load-balanced cloud services and database connection pooling enable systems to sustain their performance levels when system workloads increase.

Expected Performance Targets

Table 4.1 shows the complete system performance targets together with the explanations which were used to establish these performance boundaries.

Table 4.1. Expected system performance targets.

Metric	Target Performance	Rationale
Detection Accuracy (Low-Light)	> 85%	Reliable 24/7 threat monitoring
End-to-End Latency	< 200 ms	Real-time emergency response
LiDAR Spatial RMSE	< 0.2 m	Accurate motion tracking
Concurrent Devices Supported	> 1000	City-scale deployment
Packet Drop Rate	< 1%	Mission-critical reliability
Bandwidth per Device	< 0.5 Mbps	Feasible over 4G/5G networks

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904234

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V DISCUSSION

Advantages of the Hybrid Edge–Cloud Approach

The proposed hybrid edge–cloud architecture is designed to combine the strengths of both distributed edge intelligence and centralized coordination. The integrated system provides multiple benefits at the system level compared to systems which use either edge-based or cloud-based surveillance systems.

The fusion of LiDAR and camera sensing is intended to provide **resilient detection across varying lighting conditions**. Vision-based detectors become less effective when operating in environments with low light, glare, or visual obstructions [1], [4], [29]. LiDAR technology provides spatial measurements which remain unaffected by lighting conditions yet it fails to identify semantic content [2], [19]. The framework integrates thesecomplementary sensing technologies to achieve continuous detection and tracking performance throughout day and night for urban monitoring purposes.

The architecture achieves **real-time responsiveness** through local edge inference which processes data at the edge element. Edge perception tasks handle their data processing requirements by operating without sending unprocessed sensor data to the cloud [14], [16]. The system meets emergency dispatch timing requirements because it transmits only essential alert metadata.

The system design promotes **bandwidth efficiency** because. The system needs to reduce network traffic by sending compact alert metadata instead of transmitting high- resolution video or unprocessed LiDAR point cloud data. The system enables better operational cost efficiency because it works with existing cellular network systems (4G/5G) to deploy city-wide systems.

The use of PostGIS-based Point-in-Polygon queries enables **intelligent geo-spatial dispatch**. The system directs alerts to specific locations instead of sending alerts to all areas [10]. The system enables multiple agencies to respond together while operators need less cognitive effort because the system delivers alerts which match their operational needs.

Privacy Preservation Through Selective Processing

The proposed system has privacy preservation as its main design objective. The framework uses multiple protection mechanisms which aim to minimize the collection and sharing of personal identifiable information.

LiDAR-based tracking provides **non-biometric spatial monitoring**, as point clouds do not contain facial features, skin tone, or clothing detail [2], [26]. The system enables object tracking and motion monitoring without using any biometric information.

The system allows **selective camera activation** because it enables users to keep cameras off except when dangerous situations occur. The system prevents excessive visual monitoring through its camera system while maintaining essential video recording capabilities.

All primary perception and fusion operations are designed to occur locally on the edge device, supporting **local processing of raw sensor data** [14], [16], [26]. The system sends only alert metadata to the cloud which helps protect sensitive information. The multi-tenant cloud architecture allows agencies to set up **data retention policies** according to their local regulatory requirements [10].

Scalability for City-Wide Deployment

The architecture is designed with scalability as a primary consideration. The system uses distributed edge nodes together with horizontally scalable cloud services to enable gradual growth which starts from pilot projects and extends to citywide systems. The multi-tenant infrastructure enables shared resource use among multiple agencies while each agency keeps its data safety through logical separation The big cities need to create dense sensor networks while they choose their hardware based on their budget. The system currently relies on high-performance LiDAR sensors which account for most of its operational costs but upcoming solid-state LiDAR innovations will lower this economic barrier [2].

Limitations and Future Work

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The framework has multiple limitations which need to be addressed through upcoming research efforts.

The performance of both camera and LiDAR systems suffers when urban areas experience occlusion due to dense construction. Future research will focus on developing **multi-view fusion across neighboring edge** nodes which will deliver overlapping coverage while eliminating blind spots [7, 8, 30].

Deep learning–based detectors struggle with **adversarial robustness challenges** because their systems can be disabled through adversarial patches or camouflage methods. The researchers should implement adversarial training strategies along with ensemble modeling because these techniques will strengthen resilience against attacks [11, 29].

LiDAR signal quality decreases under environmental conditions which include extreme rainfall and fog and snowy weather. The use of radar together with three different sensing technologies will enhance operational performance during all weather conditions [30].

The system needs further research into **adaptive confidence thresholding and human-in-the-loop verification** to increase operational reliability because current fusion methods only reduce false positives.

The current design assumes that model training needs to happen at a central location. The implementation of **federated learning** enables multiple parties to create shared models while they keep their data on edge devices [10, 17].

The current framework works to detect threats as they occur. The system needs **predictive analytics and trajectory modeling** to conduct risk assessment before any unsafe situation arises. Prior surveys highlight the potential of machine learning for forecasting high-risk zones and movement patterns [17], [35], suggesting a pathway toward anticipatory urban safety systems.

VI. CONCLUSION

The study proposes a hybrid edge-cloud system which overcomes three main weaknesses of urban surveillance systems, which include their inability to operate in low-light environments and their extended cloud processing delays and their protection of private information. The framework establishes resilient threat detection abilities which function in real-time for smart city environments through its combination of LiDAR spatial sensing with vision-based detection and its geographical based cloud dispatch system. The research solution aims to create a detection system which operates reliably under different environmental conditions while it processes private data and maintains operational capacity.The work demonstrates how urban safety infrastructure will evolve because of multi-modal sensor fusion and distributed intelligence systems which enable new designs for urban safety systems. The system-level framework establishes an approach which unifies perception of robustness with bandwidth efficiency and emergency coordination between different jurisdictions. The system design enables organizations to operate their existing communication networks through its fundamental design, which eliminates the need for central video processing while creating defined alert transmission channels. The system demonstrates how surveillance technologies can support operational requirements and ethical standards, which will influence the creation of public safety technologies that protect individual privacy through its implementation of role-based access control with emergency identity verification systems.

The proposed framework shows great potential but contains multiple restrictions which require additional study. Performance under severe weather conditions, dense occlusion, and adversarial visual manipulation remains an open challenge. The high costs and power requirements of LiDAR equipment will limit large-scale deployments, which will drive organizations to investigate cheaper sensing solutions. The research should concentrate on three areas: distributed node multi-view sensor fusion, all-weather operation through tri- modal sensing, and federated learning for secure model enhancement through predictive analytics, which will establish a risk assessment framework. The framework will develop from threat detection operations into urban safety intelligence through these advanced directions.

REFERENCES

[1] C. Fernandez-Testa, et al., "Distributed Intelligent Video Surveillance for Early Armed Robbery Detection Based on Deep Learning," in Proc. Conference on Graphics, Patterns and Images (SIBGRAPI) (2024). https://doi.org/10.1109/sibgrapi62404.2024.10716299

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)
[2] D. Torres, et al., "Pedestrian Detection with LiDAR Technology in Smart-City Deployments-Challenges and Recommendations," Computers, 12 (2023). https://doi.org/10.3390/computers12030065
[3] A. Ardiansyah, et al., "EagleEYE: Aerial Edge-enabled Disaster Relief Response System," in Proc. Europea n Conference on Networks and Communications (EuCNC) (2020). https://doi.org/10.1109/EUCNC48522.2020.9200963
[4] A. Abins, et al., "Weapon Recognition in CCTV Videos: Deep Learning Solutions for Rapid Threat Identification," in Proc. International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (2024). https://doi.org/10.1109/ic- etite58242.2024.10493569
[5] S. Chakraborty, et al., "Design and IoT Implementation of a LiDAR Assisted Large- Area Video Surveillance Network," in Proc. International Conference on Power, Instrumentation, Energy and Control (PIECON) (2023). https://doi.org/10.1109/punecon58714.2023.10449997
[6] A. Singh, et al., "IoT based weapons detection system for surveillance and security using YOLOV4," in Proc. International Conference on Communication and Electronics Systems (ICCES) (2021).
https://doi.org/10.1109/ICCES51350.2021.9489224X. Jiang, et al., "Ultra Large-Scale Crowd Monitoring System Architecture and Design Issues," IEEE Internet of Things Journal, 8, 10356-10366 (2021). https://doi.org/10.1109/JIOT.2021.3076257
[7] J. Yu, et al., "Intelligent Visual-IoT-Enabled Real-Time 3D Visualization for Autonomous Crowd Management," IEEE Wireless Communications, 28, 138-145 (2021). https://doi.org/10.1109/MWC.021.2000497
[8] R. Dautov, et al., "Metropolitan intelligent surveillance systems for urban areas by harnessing IoT and edge computing paradigms," Software - Practice and Experience, 48, 1475-1491 (2018). https://doi.org/10.1002/SPE.2586
[9] L. Zhang, et al., "Developing real-time IoT-based public safety alert and emergency response systems," Scientific Reports, 15 (2025). https://doi.org/10.1038/s41598-025- 13465-7
[10] A. Kumar, et al., "Enhancing Smart City Safety and Utilizing AI Expert Systems for Violence Detection," Future Internet, 16 (2024). https://doi.org/10.3390/fi16020050
[11] S.O. Ajakwe, et al., "ALIEN: Assisted Learning Invasive Encroachment Neutralization for Secured Drone Transportation System," Sensors, 23 (2023). https://doi.org/10.3390/s23031233
[12] J. Cadet, et al., "AI-powered threat detection in surveillance systems: A real-time data processing framework," Open Access Research Journal of Engineering and Technology, 7 (2024). https://doi.org/10.53022/oarjet.2024.7.2.0057
[13] Y. Chen, et al., "Distributed real-time object detection based on edge-cloud collaboration for smart video surveillance applications," IEEE Access, 10, 93745-93757 (2022). https://doi.org/10.1109/ACCESS.2022.3203053
[14] "Real-time fire detection and alarm system using edge computing and cloud IoT platform," International Journal of Wireless and Mobile Computing, 23, 1-11 (2022). https://doi.org/10.1504/ijwmc.2022.124822
[15] "Distributed Real-Time Object Detection Based on Edge-Cloud Collaboration for Smart Video Surveillance Applications," IEEE Access, 10, 93745-93757 (2022). https://doi.org/10.1109/access.2022.3203053
[16] S. Cheedalla, "Predictive mobile AI: Transforming emergency response from reactive to preventative," World Journal of Advanced Engineering Technology and Sciences, 15 (2025). https://doi.org/10.30574/wjaets.2025.15.3.1108
[17] X. Chen, "Smart Camera Video Processing and Recognition System Based on Edge Computing," Applied and Computational Engineering, 97 (2025). https://doi.org/10.54254/2755-2721/2025.bj26597
[18] C. Amendola, et al., "Edge-Distributed Fusion of Camera-LiDAR for Robust Moving Object Localization," IEEE Access, 11, 76890-76903 (2023). https://doi.org/10.1109/access.2023.3295212 R. Rajavel, et al., "IoT-based smart healthcare video surveillance system using edge computing," Journal of Ambient Intelligence and Humanized Computing, 13, 3195-3207 (2021). https://doi.org/10.1007/S12652-021-03157-1
[19] Y. Zhang, et al., "Autonomous dam surveillance robot system based on multi-sensor fusion," Sensors, 20 (2020). https://doi.org/10.3390/S20041097
[20] S. Thota, et al., "Emergency Vehicle Detection Using IoT in Smart Cities," in Proc. International Conference on Smart Electronics and Communication (ICOSEC) (2023). https://doi.org/10.1109/icseiet58677.2023.10303407
[21] M. Eberhardt, et al., "An Enhanced Surveillance System for Gun and Knife Detection using YOLOv8 and Raspberry Pi," in Proc. International Conference on Computing, Communication, and Intelligent Systems (IC3IOT) (2024). https://doi.org/10.1109/ic3iot60841.2024.10550256
[22] D. Berardini, et al., "Benchmark Analysis of YOLOv8 for Edge AI Video-Surveillance Applications," in Proc. Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (2024). https://doi.org/10.1109/mesa61532.2024.10704889
[23] J. Wu, et al., "A Survey of Joint Security-Safety for Function, Information and Human in Industry 5.0," Security and Safety, 3 (2024). https://doi.org/10.1051/sands/2024014
[24] A. Fitwi, et al., "A Lightweight Blockchain-based Privacy Protection for Smart Surveillance at the Edge," (2019).
[25] "Real-Time Surveillance Video Analytics: A Survey on the Computing Infrastructures," Lecture Notes in

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Networks and Systems, 490, 267-278 (2022). https://doi.org/10.1007/978-981-19-5292-0_23
[26] A. Fares, et al., "Edge/Fog-based Architecture Design for Intelligent Surveillance Systems in Smart Cities: A Software Perspective," in Proc. International Conference on Mobile and Ubiquitous Computing (MIUCC) (2021). https://doi.org/10.1109/MIUCC52538.2021.9447632
[27] I. Ahmed, et al., "Development and optimization of deep learning models for weapon detection in surveillance videos," Applied Sciences, 12 (2022). https://doi.org/10.3390/app12125772
[28] Z. Bai, et al., "A Survey and Framework of Cooperative Perception: From Heterogeneous Singleton to Hierarchical Cooperation," arXiv preprint arXiv:2208.10590 (2022). https://doi.org/10.48550/arXiv.2208.10590
[29] N. Choudhry, et al., "A Comprehensive Survey of Machine Learning Methods for Surveillance Videos Anomaly Detection," IEEE Access, 11, 114680-114713 (2023). https://doi.org/10.1109/ACCESS.2023.3321800
[30] V.F. Miramá, et al., "A Survey of Machine Learning in Pedestrian Localization Systems: Applications, Open Issues and Challenges," IEEE Access, 9, 120138-120157 (2021). https://doi.org/10.1109/ACCESS.2021.3108073M.A. Ridwan, et al., "Applications of Machine Learning in Networking: A Survey of Current Issues and Future Challenges," IEEE Access, 9, 52523-52556 (2021). https://doi.org/10.1109/ACCESS.2021.3069210
[31] Q. Wang, et al., "A Survey on ML4VIS: Applying Machine Learning Advances to Data Visualization," IEEE Transactions on Visualization and Computer Graphics, 28, 5134-5153 (2022) . https://doi.org/10.1109/TVCG.2021.3106142
[32] V. Mandalapu, et al., "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," IEEE Access, 11, 60153-60170 (2023). https://doi.org/10.1109/ACCESS.2023.3286344
[33] Roboflow Universe, "crowd counting dataset Computer Vision Model" Available : https://universe.roboflow.com/crowd-dataset/crowd-counting-dataset-w3o7w/dataset/2 [Accessed: Jan 2026].
[34] Roboflow Universe, "Crowd anomaly detection Computer Vision Model" Available: https://universe.roboflow.com/crowd-detection-cnw22/crowd-anomaly-detection- r9c66/dataset/3 [Accessed: Jan 2026].
[35] Roboflow Universe, "weapon-detection Computer Vision Model" Available: https://universe.roboflow.com/yolov7test-u13vc/weapon-detection-m7qso/dataset/4 [Accessed: Jan 2026].
[36] Roboflow Universe, "weapon-detection Computer Vision Model" Available: https://universe.roboflow.com/yolov7test-u13vc/weapon-detection-m7qso/dataset/1 [Accessed: Jan 2026].
[37] Roboflow Universe, "fire Computer Vision Dataset" Available: https://universe.roboflow.com/anomaly-detention/fire-z3djy [Accessed: Jan 2026].
[38] Kaggle, "Abnormal High-density Crowds" Available: https://www.kaggle.com/datasets/angelchi56/abnormal-highdensity-crowds [Accessed: Jan 2026].
[39] GitHub, "awesome-crowd-datasets" Available: https://github.com/klchang/awesome- crowd-datasets?tab=readme-ov-file [Accessed: Dec 2025].
[40] D. Paez-Granados, Y. He, D. Gonon, L. Huber, A. Billard, "3D point cloud and RGBD of pedestrians in robot crowd navigation: detection and tracking", IEEE Dataport (2021). https://doi.org/10.21227/ak77-d722

INNO SPACE
SJIF Scientific Journal Impact Factor

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com