An Edge-Cloud Collaborative Urban Safety System for High-Altitude Object Dropping Detection, Alarm, and Record Filing

Zeyuan Zang Jichen Zhang Zishuo Wang Baibei Jin

Rentao Gu (Mentor)

School of Future, Beijing University of Posts and Telecommunications
Beijing 100876, China

{zangzeyuan,rentaogu}@bupt.edu.cn, tinyman0521@hotmail.com, 13902050688@163.com, jinbaibei30@gmail.com

Executive Summary

With the continuous advancement of urbanization, falling objects from high-rise buildings have become a prominent issue in urban public safety. These incidents are characterized by their sudden occurrence, concealment, and high potential for harm, which complicates evidence collection and governance. Although several countries have enacted laws and regulations and installed surveillance systems around buildings, the existing systems lack specialized detection algorithms and real-time alert mechanisms for falling objects, making it difficult to meet the demands of smart city safety governance. To address this challenge, this paper proposes the **HADAR (High-Altitude Object Dropping Detection, Alarm, and Record Filing)** system, which integrates detection, alarm, and record-keeping functions. The system aims to provide a widely deployable solution for falling-object detection and documentation around buildings, filling the existing technological gap. Its implementation will offer automated and real-time safety protection for urban communities, driving the intelligent development of falling-object management technologies, and contributing to the enhancement of urban public safety and the achievement of smart city goals.

1 General Description

In recent decades, with the continuous advancement of urbanization, the proliferation of high-rise buildings has made falling objects from height an increasingly concerning safety hazard. The danger primarily lies in two aspects: the severity of individual incidents and their high frequency of occurrence. Due to the rapid acceleration of objects during free fall, even relatively small items can cause serious injuries if they strike a person's head. For example, a 500-gram bottle of mineral water falling from a height of five stories (approximately 20 meters) would, under the assumption of an impact duration of 0.01 seconds, generate an impact force of about 500

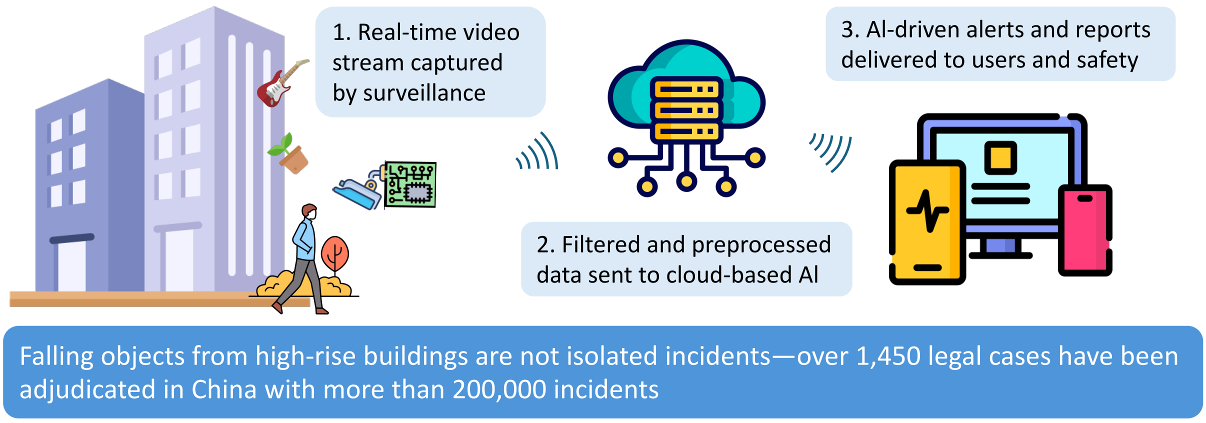


Figure 1: Overview of HADAR's procedure

kilograms—sufficient to inflict fatal injuries. At the same time, falling-object incidents have been reported with alarming frequency. In one month of 2023 alone, Shenzhen's "12345" Government Service Hotline received as many as 400 complaints related to falling objects. Based on this figure, it can be estimated that more than 200,000 such incidents may occur nationwide each year. According to a report by the U.S. Bureau of Labor Statistics, more than 50,000 injuries annually in the United States are attributed to being struck by falling objects.

In response to the growing urban safety challenge posed by falling objects from high-rise buildings, this article proposes the HADAR system, which integrates edge computing, deep learning, and Internet of Things (IoT) architecture to provide an automated solution for monitoring falling-object incidents around buildings and issuing timely alerts, thereby significantly enhancing the safety of urban communities. The implementation of this system involves four phases: data collection and model development, edge computing optimization, system integration and deployment framework, and field testing and refinement.

As is shown in Figure 1, the core methodology of this study lies in the combination of edge intelligence and edge-cloud collaboration. By deploying lightweight object detection models, the system achieves real-time detection and alerting of falling objects on edge devices such as Raspberry Pi. Specifically, the system uses a combination of NVIDIA Jetson Nano/Orin hardwares and cameras to process video streams locally, employing techniques such as Fourier transform, frame differencing, and object contour filtering to ensure fast and accurate detection of falling objects in complex environments. This edge computing solution significantly reduces the delay in detection and alerting, minimizing the reliance on cloud computing resources, and ensuring real-time performance and efficiency.

Furthermore, the IoT architecture of the HADAR system supports data interaction and remote management across multiple devices. The collaboration between cameras, edge computing units, cloud servers, and frontend interfaces ensures the system's scalability and compatibility. The frontend interface is designed using the Vue3 framework, while the backend is built with Node.js, ensuring the system's efficiency and flexibility. To enhance the system's robustness, the research also incorporates data augmentation and Kalman filtering techniques, ensuring stable operation in complex environmental conditions, such as adverse weather, lighting changes, and variations in building types. Particularly, through the use of adaptive thresholding, the system dynamically adjusts detection confidence based on scene complexity, further improving accuracy and stability.

Generally speaking, the HADAR system, through its combination of edge computing and edge-cloud collaboration, along with innovative integration of deep learning and computer vision technologies, not only provides an efficient and accurate solution for detecting and alerting

(a) Photo with **Kathleen Kramer**, the President and CEO of Institute of Electrical and Electronics Engineers (IEEE).

(b) Photo with **Amandeep Singh Gill**, Under-Secretary-General and the UN Secretary-General's Envoy on Technology.

(c) Photo with **Zou Ciyong**, Deputy Director General and Managing Director of the United Nations Industrial Development Organization (UNIDO).

Figure 2: Our project was invited to exhibit at the WSIS+20 High-Level Event and the AI for Good Global Summit held in Geneva, where we briefed senior officials from the UN, ITU and IEEE, and gained unanimous recognition.

falling-object incidents but also offers strong technical support for the safety management of smart cities. This project won the Second Prize at the National Finals of the 18th National College Software Innovation Competition—Software Design Innovation Contest, and was showcased at the AI for Good Global Summit in Geneva, Switzerland, as well as the WSIS+20 High-Level Event 2025, where it received wide recognition and attention from officials of IEEE, the United Nations and other international organizations.

2 Problem Formulation

With the acceleration of urbanization and increasing density of high-rise buildings, **falling objects from height** have emerged as a critical threat to urban public safety. Statistical evidence reveals that Chinese courts adjudicate over 1,450 such cases annually, while unreported incidents may reach 200,000 per year. The severity extends far beyond common perception—for instance, a 500-gram water bottle falling from the 5th floor generates an impact force of 500 kilograms, far exceeding the tolerance threshold of the human skull.

Through comprehensive analysis of existing surveillance systems and falling-object detection requirements, we identify five fundamental technical bottlenecks that severely limit current approaches:

Bottleneck 1: Real-time Detection Latency Challenge Traditional computer vision approaches employing background modeling methods achieve low latency (< 20ms) but suffer from inadequate small falling-object detection accuracy below 30%. Conversely, deep learning algorithms demonstrate higher accuracy ($> 75\%$) but often exhibit inference latency exceeding 200ms, failing to provide timely warnings before impact occurs.

Bottleneck 2: Small Object Recognition Accuracy Limitation Falling objects typically occupy less than 0.1% of the image frame, with over 90% measuring below 30×30 pixels. Current detection algorithms struggle with feature loss and motion blur, achieving recognition accuracy below 50% for such small, fast-moving targets. This limitation severely compromises the reliability of automated detection systems.

Bottleneck 3: Cross-scenario Generalization and Adaptability Gap Commercial monitoring systems require per-building parameter optimization and manual calibration. Under adverse environmental conditions including rain, fog, and glare, false positive rates surge dramatically, severely constraining large-scale deployment feasibility and operational reliability.

Bottleneck 4: Evidence Documentation and Forensic Integrity Issues Current systems lack intelligent mechanisms for extracting critical event moments, specifically object release and pre-impact phases essential for legal proceedings. Manual video review is time-consuming and often fails to capture key evidence frames.

Bottleneck 5: Scalable System Architecture and Resource Optimization Challenge Existing solutions typically rely on centralized cloud processing, creating bandwidth bottlenecks and single points of failure. The lack of edge-cloud collaborative architectures results in inefficient resource utilization, with 24/7 video storage costs exceeding 10,000 RMB per camera annually while critical event segments constitute less than 0.003% of total footage, resulting in substantial resource waste and retrieval difficulties.

These contradictions lead to a dilemma of "failing to prevent before, and failing to trace after", calling for breakthrough technical solutions.

3 Technical Solution

Based on our comprehensive analysis of technical challenges, we propose **HADAR (High-Altitude Object Dropping Detection, Alarm, and Record Filing)** system. The system architecture employs a carefully curated technology stack organized into three primary categories: **Edge Computing Infrastructure**, **Machine Learning and Computer Vision**, and **System Integration and Communication**.

3.1 Architecture overview

Edge Computing Infrastructure The edge intelligence layer utilizes NVIDIA Jetson Nano/Orin hardware platforms running Ubuntu JetPack OS with specialized peripherals including IMX219 CSI cameras and GPIO-controlled alert systems. The compute acceleration stack comprises CUDA Toolkit for GPU programming, TensorRT for INT8 model optimization, cuDNN for deep learning acceleration, and Jetson-specific libraries (Jetson-GPIO, Jetson.Camera) for hardware interfacing. Performance optimization is achieved through cupy for GPU-accelerated computing and opencv-cuda for accelerated computer vision operations.

Machine Learning and Computer Vision The core detection pipeline leverages Ultralytics YOLOv8 for object detection, PyTorch with torchvision for deep learning model development, and Bi-LSTM networks for temporal sequence modeling. The inference deployment stack includes TensorFlow Lite for lightweight edge inference, ONNX Runtime for cross-platform model deployment, and standard computer vision libraries (opencv-python, PIL) for image processing. Mathematical computations are handled by numpy, scipy, and specialized filtering libraries (filterpy for Kalman filtering).

System Integration and Communication The distributed system architecture employs Node.js with Express framework for RESTful API development, supported by middleware including body-parser for request parsing, cors for cross-origin resource sharing, and multer for file upload handling. Frontend interfaces are developed using Vue3 ecosystem including vue-router for navigation, pinia for state management, axios for HTTP communication, and Element Plus for UI components. Desktop applications utilize PyQt5 framework with associated modules (PyQt5-Qt5, PyQt5-sip) for cross-platform compatibility. Database operations are handled through MySQL for centralized cloud storage and SQLite3 for edge-side local caching, with database connectivity managed by appropriate drivers.

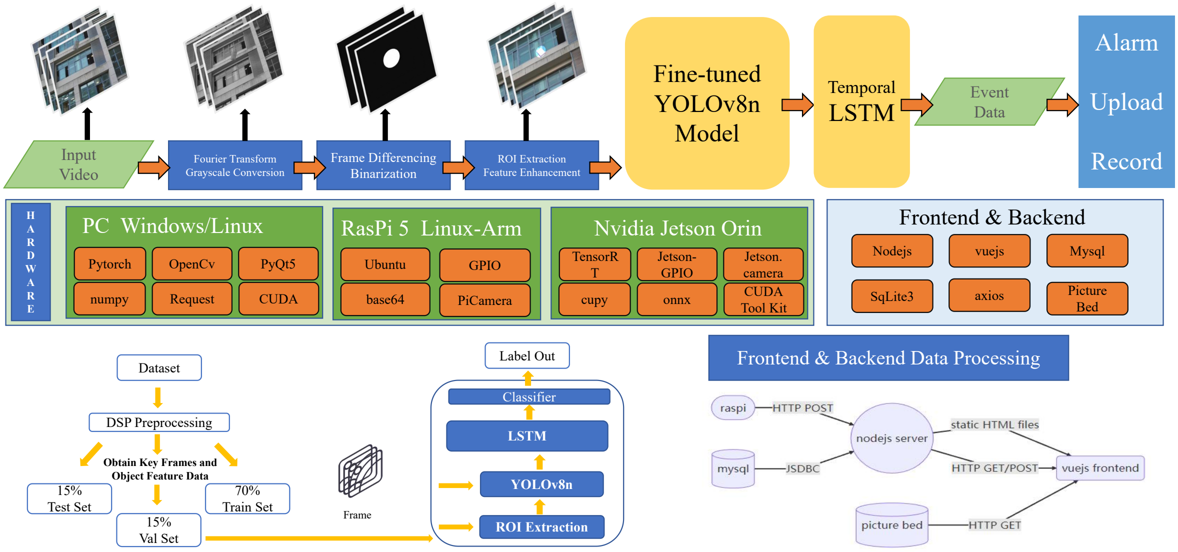


Figure 3: The hardware setup, software implementation and whole workflow of HADAR

These modules converge into a **hierarchical edge–cloud collaborative framework**, where intelligent perception at the edge is orchestrated with centralized services in the cloud. The whole implementation is illustrated in Figure 3.

A carefully curated falling-object dataset is used to train a YOLO + Bi-LSTM hybrid model, enhanced with data augmentation to improve robustness. After training, the model is quantized, converted into ONNX format, and deployed on a Jetson Nano device, where TensorRT is utilized to enable efficient edge inference. During real-time operation, incoming video streams undergo DSP preprocessing, including Fourier transformation, frequency correction, and grayscale differencing, to enhance detection accuracy.

When a falling object is detected, the system triggers two simultaneous responses. On the edge device, a local real-time alarm is issued via a speaker, while event metadata—including time, floor, severity level, and keyframes—is uploaded to a Node.js backend through an HTTP POST request and stored in a MySQL/SQLite3 database for future analysis.

For user interaction, the system provides two types of interfaces: a PC-based monitoring UI built with PyQt5 for live visualization, and a web/mobile application that allows remote users to query and review historical events using filters such as time, location, floor, and keyframes.

3.2 Core Technical Innovations

To address the five identified technical bottlenecks systematically, the HADAR system introduces the following five innovative solutions that directly target each specific challenge.

Innovation 1: Hybrid Architecture for Real-time Detection (Addressing Bottleneck 1)

To satisfy the strict real-time requirement of falling-object detection within **50 ms**, we designed a hybrid detection architecture that integrates traditional DSP techniques with a lightweight deep learning model. The pipeline consists of four stages shown in Figure 4.

In the preprocessing stage, Fourier transform and frequency-domain correction are used to suppress background jitter, while frame differencing highlights moving targets. Candidate regions are extracted via binarization and adaptive ROI cropping, which reduces computation and filters out irrelevant objects. After that, these regions are processed by a pruned YOLOv8n

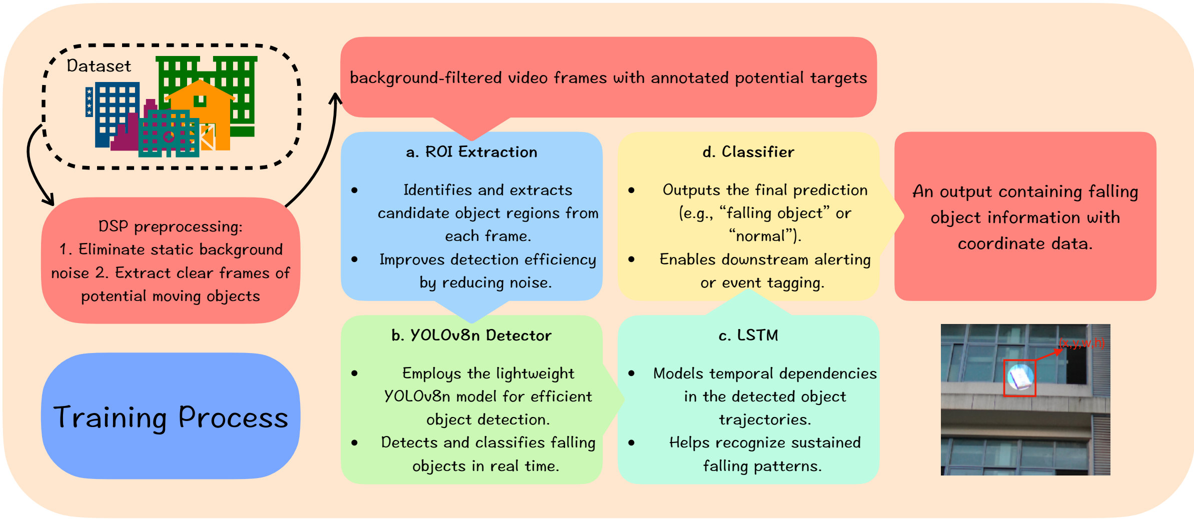


Figure 4: Training pipeline of the real-time detection algorithm

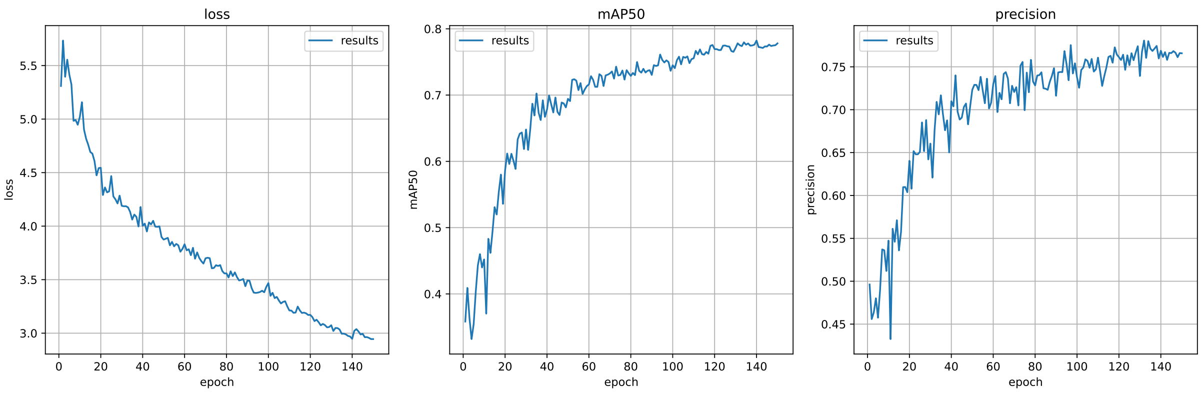


Figure 5: Training metrics over epochs for the falling-object detection model.

model that focuses on small object detection and performs inference only on candidate areas, enabling batch processing for higher efficiency. Finally, a temporal validation mechanism verifies whether the trajectory matches gravitational acceleration, effectively reducing false positives.

To achieve the targeted latency, we apply model slimming, TensorRT-based INT8 quantization, CUDA-accelerated preprocessing, and a pipelined strategy where DSP operations on the current frame overlap with inference on the next. The end-to-end detection latency was reduced from **193 ms** to **57 ms**, satisfying real-time requirements while maintaining high detection accuracy. The latency distribution for each processing stage in a single falling-object event is shown in Table 1.

Figure 5 illustrates the evolution of three key training metrics—loss, mean Average Precision at IoU threshold 50 (mAP@50), and precision—over 150 epochs during the training process of falling-object detection model based on the YOLOv8n backbone architecture. The model was trained using the Adam optimizer with a learning rate of 0.01, a batch size of 4, and input image dimensions of 640×640 pixels.

Innovation 2: Adaptive Multi-scale Feature Enhancement (Addressing Bottleneck 2)

Most falling objects are extremely small, with over 90% below 30×30 pixels, making detection challenging due to feature loss and motion blur. To address this, we enhance small-object representation using multi-scale feature fusion through YOLO's Feature Pyramid Network and apply DSP-based sharpening within adaptive ROIs to enlarge object proportions and reduce blur. These strategies significantly improve detection precision for small, fast-moving objects.

After applying the optimizations, the model's small object detection performance improved significantly. Results on the curated dataset show that the detection accuracy (Acc.) reached **74.51%**, and that the Intersection over Union (IoU) rate reached **61.02%**.

Innovation 3: Cross-scenario Adaptive Learning Framework (Addressing Bottleneck 3)

Since deployment environments vary, the system is designed for high adaptability. We collected data across 18 architectural scenarios, four weather conditions, and eight categories of falling objects, combined with extensive data augmentation to improve robustness. Dynamic thresholding and gravity-parameter adaptation ensure stable performance under different lighting conditions and camera heights. Moreover, incremental learning allows the model to efficiently fine-tune itself using small batches of newly collected data, improving adaptability in unseen environments.

The proposed methods significantly enhanced model robustness across diverse deployment scenarios. Without scene-specific fine-tuning, the average accuracy drop decreased from **38.7%** to **24.2%**. After applying incremental fine-tuning, the drop was further reduced to as low as **8%**, demonstrating strong adaptability to unseen environments.

Innovation 4: Intelligent Keyframe Evidence Extraction (Addressing Bottleneck 4)

For legal investigation and forensic purposes, we developed an intelligent keyframe extraction algorithm to identify two critical moments: when the object is released and just before ground impact. The algorithm leverages velocity vector variation analysis and a sliding time-window clustering method to locate segmentation points accurately.

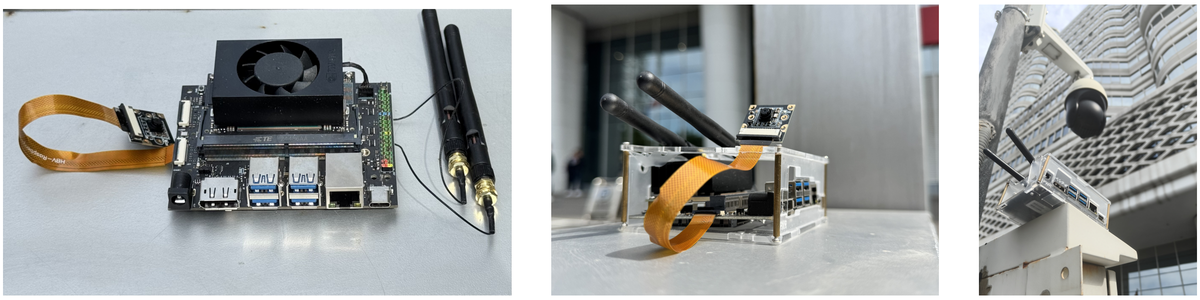
After applying the optimizations, the system automatically extracts complete event video clips along with 1-4 key frames for each throwing incident. The key-frame extraction accuracy reached **93.7%**, ensuring reliable visual evidence for law enforcement and forensic investigations.

Innovation 5: Scalable Edge-Cloud Collaborative Architecture (Addressing Bottleneck 5)

We design a heterogeneous edge-cloud coordination framework that optimizes resource utilization while maintaining system reliability and reducing operational costs. It implements local processing on NVIDIA Jetson platforms with TensorRT optimization, reducing cloud dependency and bandwidth requirements. On-device, detected keyframes and metadata are buffered locally in SQLite3 to ensure functionality under unstable network conditions. Once connectivity is available, the data is synchronized with a Node.js-based RESTful API server, where images

Table 1: Latency Distribution Across Processing Stages

Processing Stage	Latency (ms)	Ratio (%)
DSP Preprocessing	22.4	39.0
Candidate Target Extraction	5.6	9.8
Deep Learning Inference	19.8	34.5
Temporal Verification	9.6	16.7
Total	**57.4**	**100.0**



(a) (b) (c)

Figure 6: Illustrations of the edge computing hardware in HADAR: **(a)** dedicated hardware design, **(b)** standalone deployment, and **(c)** integration with existing surveillance cameras.

receive unique 16-digit UUIDs and are stored as static resources, while metadata is saved in the database. A Vue3-based web interface supports multi-criteria queries by time, location, and severity level, providing concise event summaries, detailed metadata, and zoomable keyframes for assisted verification.

Using the optimized workflow, a complete pipeline for **event metadata upload, storage, classification, and retrieval** was established, with user interfaces developed for PC, Web, and Mobile App platforms.

Thanks to the lightweight architecture, the system demonstrates high availability, easy maintenance, and low deployment cost. Performance testing showed the following capabilities:

- Supports more than **10,000+ events per day**.
- A single server instance supports **10+ monitoring nodes**.
- Scalable for deployment in medium- and small-scale communities.

These technical innovations systematically address all identified bottlenecks, providing a comprehensive solution that advances the state-of-the-art in automated falling-object detection and management systems. The integrated approach not only resolves individual technical challenges but also creates synergistic effects that enhance overall system performance and reliability.

3.3 Real-World Performance

The system runs on an Nvidia Jetson Orin Nano board equipped with an IMX219 8MP CSI camera and a GPIO-controlled buzzer for real-time alarms. The software stack is built on `JetPack 6.2` with `CUDA 12.6` for GPU acceleration, while `Jetson-GPIO` and `Jetson.Camera` manage peripherals and optimized image capture. To maximize efficiency, we utilize GPU-accelerated `cupy` computations, `opencv-cuda` for optimized preprocessing, zero-copy memory access to

Table 2: Performance of Edge-side Computing Architecture

Metrics	Value	Notes
Average Inference Latency	57 ms	Single-frame inference (960×720)
End-to-End System Latency	61 ms	From detection to alarm trigger
Memory Usage	2.7 GB	Includes OS and application load
Average Power Consumption	8.7 W	During 30 min with $\sim$1 event detected

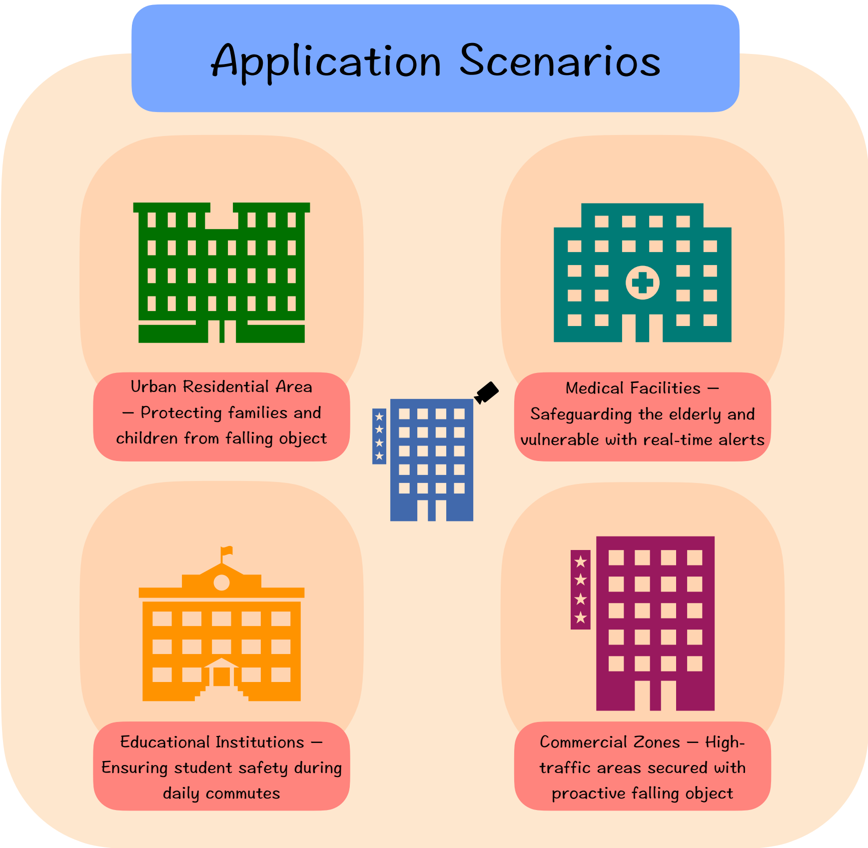


Figure 7: HADAR's deployment scenarios in different urban environments

minimize data transfer overhead, and TensorRT-based INT8 quantization to accelerate YOLOv8n inference while maintaining high accuracy.

The system achieved the edge-side performance summarized in Table 2 on the **Jetson Nano** platform.

These experimental results confirm that the proposed edge-side computing architecture enables efficient inference under resource-constrained conditions while maintaining low power consumption and high stability. The system meets real-time requirements, achieving latency **<60 ms** and inference speed **>16 FPS**, thus providing a reliable hardware foundation for high-altitude falling-object detection and real-time alerting.

4 Social Impact

4.1 Key issues and application scenarios

This study addresses the limitations of current urban safety systems in detecting and managing falling objects from high-rise buildings. Our HADAR system integrates real-time detection, alarm triggering, and incident documentation, improving monitoring accuracy, response speed, and management efficiency for falling-object incidents in urban environments.

As illustrated in Figure 7, HADAR can be deployed in diverse urban environments, including high-density residential areas, commercial complexes, schools, hospital zones, and public plazas. In older communities or open-style residential layouts, the risks of falling objects are heightened due to renovations, aging infrastructure, and overcrowding. HADAR provides real-time monitoring of key areas and delivers audio alarms, LED displays, and mobile notifications to promptly alert residents, pedestrians, and managers, enabling them to take preventive actions. Vulnerable groups, particularly children and the elderly, benefit most from this realtime intervention, significantly reducing injuries and fatalities. Beyond preventing accidents, HADAR enhances the overall sense of security in communities and builds public confidence in safety management.

4.2 Societal Value and Governance Implications

Enhancing Public Safety and Social Trust The high-precision falling-object detection feature of HADAR enables residents to trust that the system can promptly identify and effectively address potential safety hazards. The effectiveness of this mechanism is demonstrated not only through real-time risk alerts but also through a systematic data traceability function that shifts governance models from "post-incident remediation" to "prevention and real-time control." Leveraging an edge-cloud collaborative architecture, HADAR records and synchronizes event metadata in real-time, including timestamps, object characteristics, and image evidence. The accumulation of such data provides law enforcement and property managers with detailed event background information, enabling them to quickly identify responsible parties and establish a comprehensive chain of evidence, thereby reducing legal disputes caused by insufficient proof or unclear accountability. This data-driven governance model enhances public trust and engagement in safety management, further advancing the strengthening of social trust and innovation in urban governance.

Driving Smart City Development Our project deploys sensors across urban areas using IoT technology, creating an intelligent and comprehensive monitoring system for falling objects that enhances the reliability and responsiveness of urban safety management. This innovative solution not only provides more reliable protection for high-rise areas but also contributes valuable experience to the development of smart cities. Through intelligent monitoring, data analysis, and real-time alerts, urban management can make more precise and efficient decisions, advancing the transformation of urban governance toward an intelligent, data-driven model.

Strengthening Legal Awareness and Social Responsibility By providing verifiable evidence for judicial rulings, HADAR improves the efficiency and fairness of law enforcement. Furthermore, the visibility of HADAR's monitoring and alerting system educates residents about safety regulations and legal responsibilities. Over time, this fosters stronger public participation in community safety initiatives, enhances collective social responsibility, and contributes to a safer and more harmonious society.

4.3 Economic Value and Sustainable Urban Development

Reducing Accident Losses and Social Costs Falling-object incidents can result in high medical expenses, property damage claims, and legal disputes. HADAR minimizes these losses through proactive detection and rapid responses, saving significant costs for residents, insurers, and local governments. In addition, by lowering the number of incidents, the system reduces societal burdens and enhances public health outcomes.

Improving Law Enforcement and Governance Efficiency Automated monitoring, event tracing, and evidence chain generation significantly reduce the need for manual inspections and

Figure 8: HADAR aligns with 4 UN sustainable development goals (SDGs)

investigations. By lowering operational costs and improving enforcement efficiency, HADAR provides a scalable and cost-effective solution for urban safety management, setting a new benchmark for smart governance in modern cities.

Driving Industrial Growth and Enhancing Urban Competitiveness HADAR connects multiple industrial ecosystems, including sensors, IoT hardware, edge computing, cloud storage, and AI-driven analytics. Its deployment promotes the advancement of smart security and smart city industries, generating new economic growth opportunities. Moreover, public safety is a critical factor influencing urban livability and investment attractiveness. By improving safety infrastructure, HADAR enhances cities' ability to attract talent, foster innovation, and stimulate sustainable economic development.

4.4 Contribution to the UN Sustainable Development Goals

As shown in Figure 8, the HADAR system contributes to four United Nations Sustainable Development Goals (SDGs). For **Goal 3: Ensure healthy lives and promote well-being for all at all ages**, the system utilizes deep learning and IoT technologies for real-time monitoring and early warning of falling objects, reducing injury risks and enhancing public health. For **Goal 9: Build resilient infrastructure, promote sustainable industrialization and foster innovation**, HADAR promotes the development of smart infrastructure and provides data-driven decision support for urban planning and safety management. For **Goal 11: Make cities inclusive, safe, resilient and sustainable**, the system enhances urban safety management, builds smart communities through intelligent monitoring, and encourages public participation in safety governance. For **Goal 16: Promote just, peaceful and inclusive societies**, HADAR strengthens social governance, improves legal compliance, and fosters public trust by enabling transparent and efficient safety management.

5 Conclusion

In summary, HADAR delivers comprehensive social and economic benefits through its integration of advanced sensing, intelligent analytics, and real-time warnings. Socially, it improves public safety, strengthens social trust, innovates governance models, and fosters greater community participation. Economically, it reduces accident-related losses, improves enforcement efficiency, stimulates industrial growth, and enhances urban competitiveness. The deployment of HADAR represents not only a technological advancement but also a dual embodiment of social responsibility and economic value, contributing to safer, smarter, and more sustainable urban environments.

Acknowledgements

The team would like to express our sincere gratitude to Prof. Rentao Gu for his valuable guidance on system design and overall project direction. All system development, algorithm implementation, and experimental evaluations were completed independently by the student team.

A Technical Specifications

For each module we choose a particular set of development frameworks and libraries, with the full list shown in Figure 9.

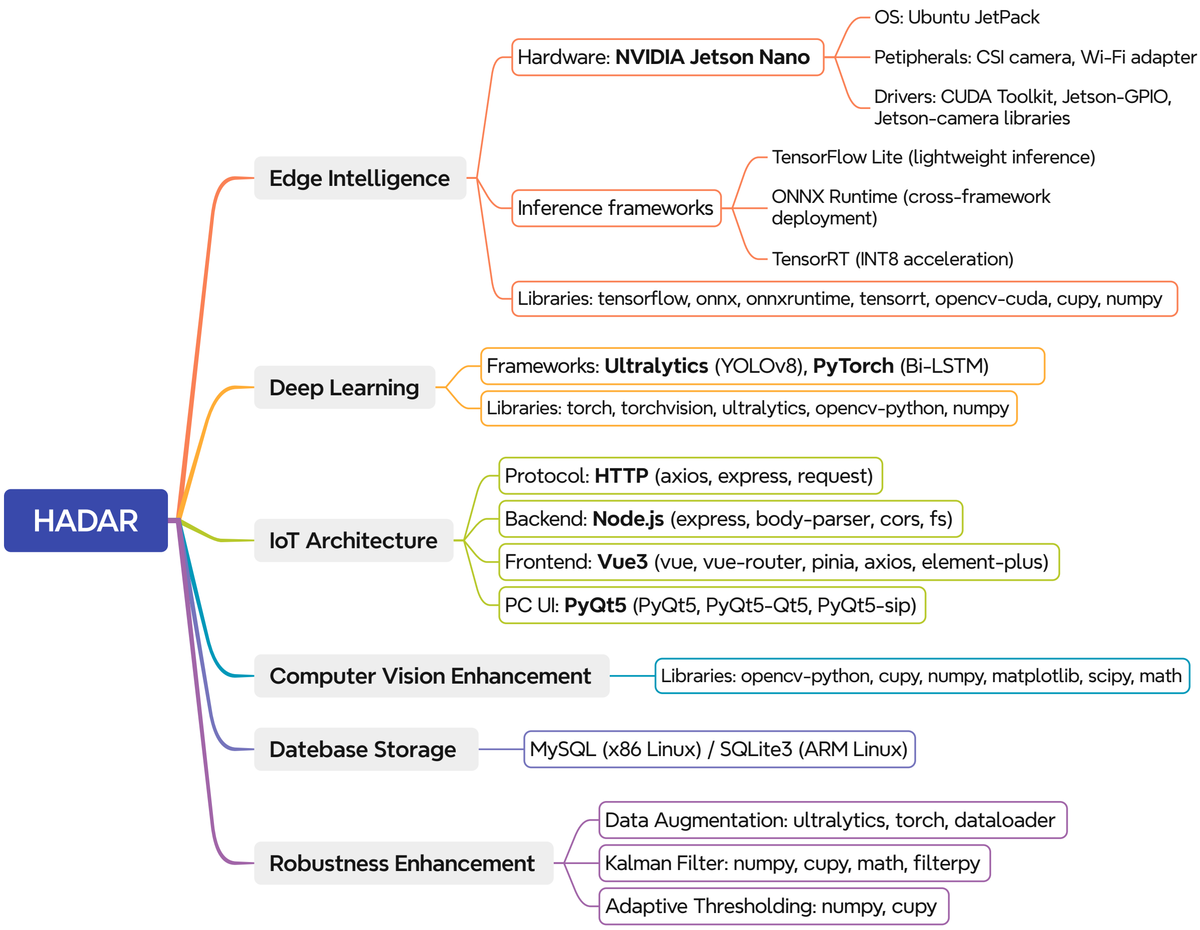


Figure 9: Development frameworks and libraries used in the HADAR system

B Cloud-based Frontend

The user interface of HADAR's cloud-based frontend is adapted for both desktop and mobile platforms.

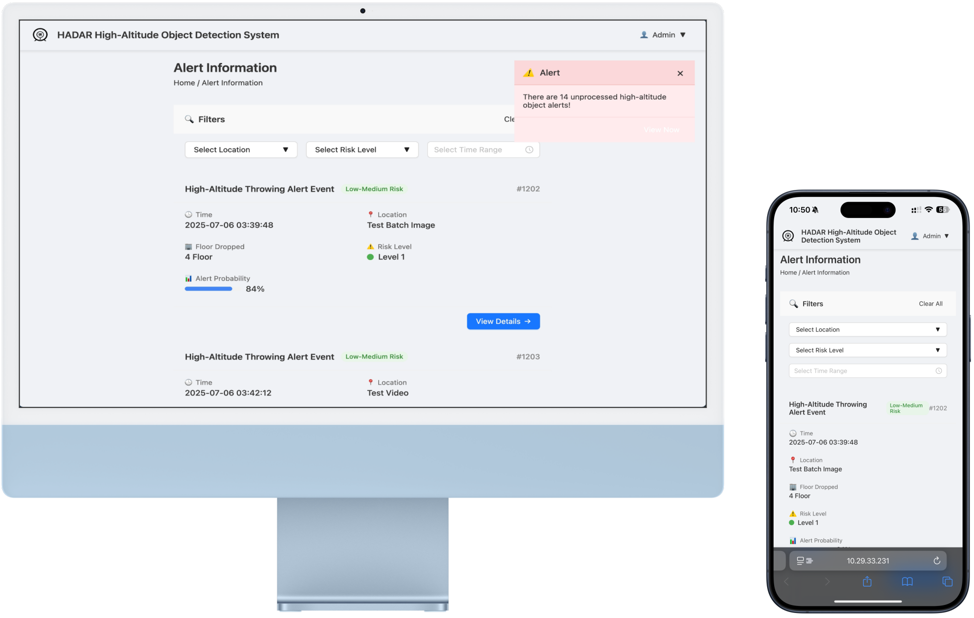


Figure 10: The alert interface, which displays a list of high-altitude object dropping events.

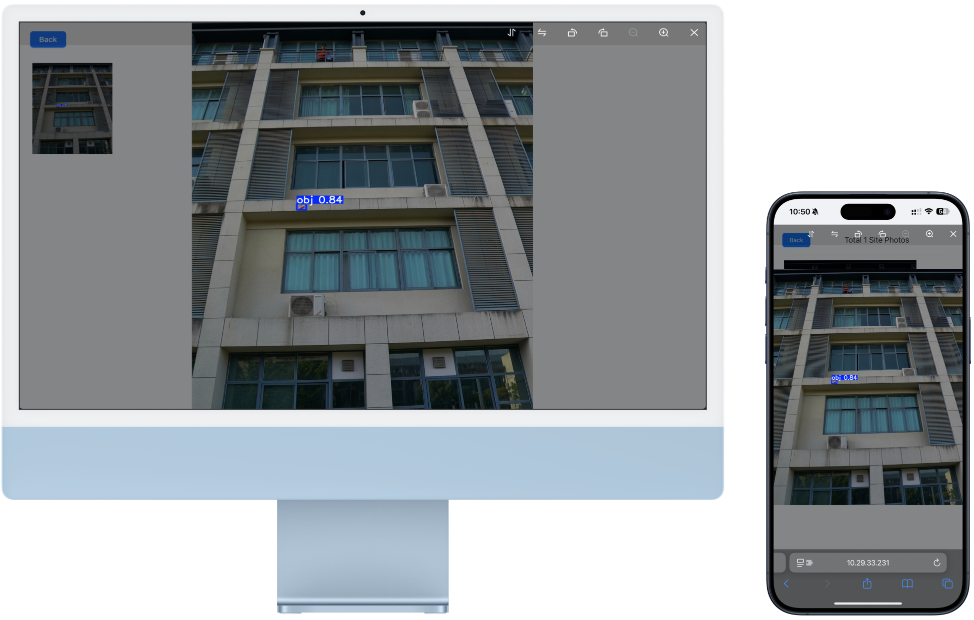


Figure 11: The details interface, which allows viewing one or multiple keyframes corresponding to each event and supports zoom operation.

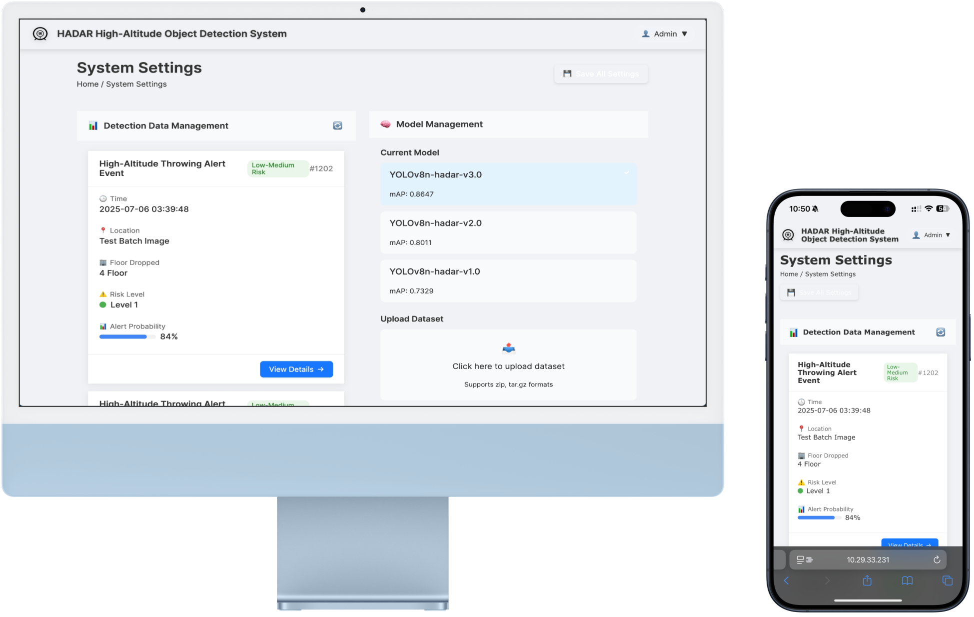


Figure 12: The system settings interface, which supports data management and remote parameter adjustment of specific edge computing devices.