Real-Time Abnormal Activity Detection and Alert System Using Computer Vision

Ajay U Bangera[1], Sahana Bandekar[2], C Jeevesh[3], Sharanya[4], Spoorthi[5]
[1,3,4,5]*Student, Department of Computer Science & Design, Srinivas Institute of Technology, Mangalore, Karnataka, India*
[2]*Assistant Professor, Department of Computer Science & Design, Srinivas Institute of Technology, Mangalore, Karnataka, India*
doi.org/10.64643/IJIRTV12I12-206692-459

***Abstract*—This paper presents a novel, multi-tenant Software-as-a-Service (SaaS) computer vision framework for the automated detection of abnormal behaviors. Historically, video surveillance has relied on manual observation or basic motion detection, which is labor-intensive and prone to severe false-positive alarm rates. Our proposed system addresses these inefficiencies by employing a dual-tier AI architecture. The system integrates YOLOv8 for high-speed, real-time anomaly detection at the edge, coupled with a secondary verification layer utilizing a Vision-Language Model (Gemini Vision API) to analyze snapshots and eliminate false alarms. Hosted on a multi-tenant architecture, the platform allows users to create isolated projects (e.g., for Educational Institutions or Government Systems), manage varied CCTV inputs, and enforce Role-Based Access Control (RBAC). The system logs verified alerts to a relational database and autonomously dispatches email notifications to authorized members. This demonstrates a scalable, highly accurate solution for proactive security deployments.**

***Index Terms*—Abnormal Activity Detection, Multi-Tenant Architecture, YOLOv8, Vision-Language Models, False Alarm Mitigation, Video Surveillance.**

I. INTRODUCTION

Abnormal activities and suspicious incidents in institutional and government environments create major security threats because they endanger safety and security operations. The traditional surveillance systems require operators to watch CCTV video because system performance declines when more surveillance cameras are added.

A. Problem Statement and Research Gap:

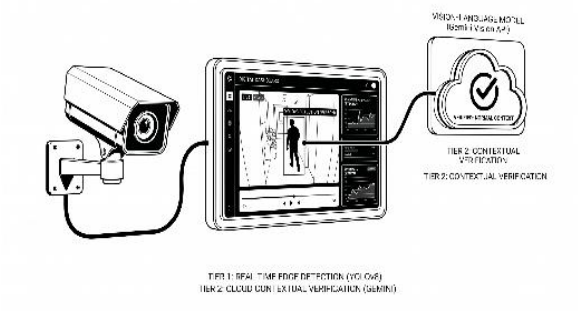
Automated systems use deep learning models such as YOLO to achieve real-time detection. However, these systems face operational problems because they produce excessive false-positive results. Security personnel experience alert fatigue because the system depends on a real-time object detection model to generate security alerts.

Existing solutions operate as single-node systems which restrict their ability to function across multiple tenants in large organizations. Our solution develops a cloud-based multi-tenant system which uses two-step AI verification to eliminate false positive results before notifying users.

B. Contributions

This research paper establishes its main findings through the development of a surveillance system which delivers both high precision and scalable performance. The system introduces a central platform which enables different project environments to operate with secure user authentication and multiple camera systems.

The system used YOLOv8 to execute real-time spatial detection which transferred suspected anomalies to the Gemini Vision API for additional contextual analysis. The system uses a secure notification matrix which delivers verified alerts and visual evidence through email to project members who have been authorized.



II. LITERATURE REVIEW

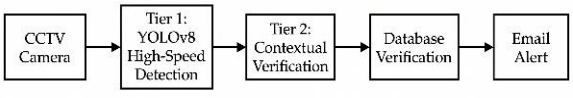
Evolution of Deep Learning in Video Surveillance Deep learning and Convolutional Neural Networks (CNNs [1][2] serve as the fundamental basis for automated abnormal activity detection systems which developed from these two technologies. The domain surveys show that deep learning functions as the main technological force which powers contemporary video anomaly detection methods according to the three surveys dedicated to this field of study. The research about visual surveillance and video action recognition shows that single-stage object detectors function as essential tools which enable real-time detection of anomalous actions according to the findings of the study. The YOLO (You Only Look Once) design may receive more adaptations to different specialized environments by the growth of improved YOLOv5 models that mixes CNNs for video surveillance substation anomaly detection purposes. The arrangement of YOLOv8 has achieved substantial success within highly organized environments which contain logistics warehouses. The growth of advanced criminal activity grouping methods uses anchor frame detection models as tools to achieve automatic criminal activity classification beyond static cameras while investigators investigate robotic systems which detect anomalies in unpredictable common areas.

The Shift Toward Vision-Language Models (VLMs) While YOLO building excel at spatial finding, recent literature highlights a critical shift in direction to Vision-Language Models (VLMs) to resolve the contextual deficits and false-positive rates inherent in traditional method [9]. Research demonstrates that fine-grained prompting is essential for unlocking the capability of VLMs to handle complex video anomaly detection [10]. The VERA framework uses VLMs to enable verbalized learning which generates explainable anomaly detection results that include contextual information instead of using basic bounding box outputs [11]. Recent research findings support the use of dual-model or ensemble systems because uncertainty-aware VLMs improve detection performance through their complete integration into the detection process [12]. Specialized sectors which include traffic anomaly detection are currently developing a new model of VLM-based surveillance.

The implementation of a dual-tier AI pipeline requires organizations to shift their operations from on-premises edge nodes to cloud-based systems which can accommodate their expanding needs according to research evidence found in reference [13]. The research study examined new security technologies for multi-tenant IoT-cloud systems which showed how organizations must handle the conflicting demands of operational efficiency and system protection when they deploy distributed surveillance technologies on shared infrastructure according to reference [14]. The systems depend on multiple external application programming interfaces which include Vision-Language Model endpoints therefore organizations must establish monitoring systems to track software as a service application programming interface access violations and data theft incidents to protect their tenant security and maintain data protection standards according to reference [15].

III. METHODOLOGY

Instead of having on a simple model setup, our system uses a more reliable two-step verification process to ensure accuracy.



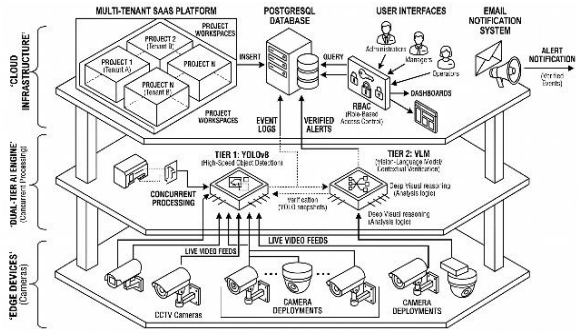
The organization uses the Ultralytics YOLOv8 model to track live video feeds through continuous real-time monitoring. The model has been developed to detect specific objects while also identifying initial indicators of abnormal behavior. The system begins its response process when it identifies any suspicious activity which violates our established baseline security protocols by capturing a high-resolution frame that they store as an active incident in their database.

The system has developed functions that enable it to transfer stored snapshots through Gemini Vision API by using its special prompt system. The system uses Gemini to evaluate images which functions as its primary security assessment tool. The system uses its capacity to detect different room situations to distinguish between activities that are allowed in examination rooms and government corridor spaces.

IV. SYSTEM ARCHITECTURE

We have designed this system to operate as a secure multi-performing platform. Means this system can

serve multiple organizations, clients or teams at a time, while that everyone's data is fully isolated and secured.



1. Logins & Project Management

Users can start creating their first project by accessing the platform and logging into their account. The project owner grants access to others through Role-Based Access Control (RBAC) which allows him to send invitations via email to team members who need access. The system provides enhanced security through RBAC because project owners choose which users can view video dashboards and receive alerts based on their job requirements.

2. The Project Workspace

System allows users to register and connect different types of CCTV cameras. Our backend system maintains its performance capabilities while handling multiple live video streams that require processing.

3. Enhanced Database Storage

The system moved from basic data storage which used local computer files to more sophisticated storage solutions. The system uses PostgreSQL as its dependable relational database which functions as its active database management system. Our database system operates as the records engine which links and arranges all components of our system including Users Projects Team Members Cameras and Incident Logs into a unified structure. The system creates a fresh secure backup whenever someone documents a security incident that links to the project and camera from which the incident was recorded to protect all data from potential loss.

4. Smart Alert System

The system demonstrates its most impressive capabilities through this specific function. The database changes the log status from pending to verified when our Tier 2 AI Gemini model detects a genuine anomaly in the system. The system creates an automatic email for the project owner and approved personnel, which includes both the event details and snapshot evidence. The security team receives no irrelevant alerts because Gemini classifies the incident as a false alarm and the system marks it with a false positive.

V. IMPLEMENTATION AND TESTING

We tested our double AI engines for their efficacy and the capacity of the database to accommodate many tenants and tenants' projects concurrently.

Execution Scenarios:

1. People who developed the AI system tested it using special test cases which they selected because those cases usually produce failures in standard AI systems. The Tier 1 component instantly reacted by capturing images of perceived problems. The Tier 2 AI engine Gemini assessed the situation and determined that all observed events did not pose any threat to security. The system automatically eliminated those alerts which security staff members had not been informed about.
2. The system successfully dispatched alerts to their correct recipients. The system worked correctly throughout all real security events.

The AI system confirmed the detection as authentic and then used the role-based access control database to identify project-specific authorized recipients who received email alerts that included image evidence which was sent only to the designated security personnel of the tenant. The system prevented any data from being shared between different tenant workspaces.

VI. CONCLUSION AND FUTURE SCOPE

The solution demonstrates an exceptional example which shows how to build a multi-tenant anomaly detection system that operates with extreme efficiency.

The fast anomaly detection capabilities of YOLOv8 work independently from the slow complex reasoning processes which belong to the Gemini Vision API system.

Future Scope:

1. The team will create intelligent message queuing systems which use RabbitMQ and Redis for handling unexpected API traffic surges to the Gemini API.
2. Project owners will gain the capability to create their own natural language prompts which will function with the Gemini Vision API.

REFERENCES

[1] Research on Anomaly Detection in Substation Video Surveillance Combining Convolutional Neural Network and Improved YOLOv5," *Proc. SPIE*, 2026.

[2] Research on Anomaly Detection in Substation Video Surveillance Combining Convolutional Neural Network and Improved YOLOv5," Proc. SPIE, 2026. (Duplicate of the above reference.)

[3] "Anomaly Detection in Logistics Warehouses Based on YOLOv8," IEEE Xplore, 2024.

[4] "Generalised Robotic Anomaly Detection in Dynamic Public Environments," arXiv preprint, 2025.

[5] "VERA: Explainable Video Anomaly Detection via Verbalized Learning of Vision-Language Models," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2025.

[6] "Unlocking Vision-Language Models for Video Anomaly Detection via Fine-Grained Prompting," in Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV), 2026.

[7] "Two Are Better Than One: Uncertainty-Aware Vision-Language Models for Video Anomaly Detection," in Proc. Int. Conf. Learn. Represent. (ICLR), 2026.

[8] "Video-Based Traffic Anomaly Detection with Vision-Language Models: A Survey," in Proc. IEEE Int. Conf. Intell. Transp. Syst. (ITSC), 2025.

[9] "SaaS API Monitoring for Data Leakage and Access Anomalies," IEEE Xplore, 2026.

[10] "Performance Trade-Offs in Multi-Tenant IoT–Cloud Security: A Systematic Review of Emerging Technologies," MDPI, 2026.

[11] "Survey of Anomaly Detection Methods in Surveillance Videos Based on Deep Learning," J. Vis. Commun. Image Represent., 2025.

[12] "Video Anomaly Detection: A Comprehensive Survey of Deep Learning Approaches," Artif. Intell. Rev., 2026.

[13] "Abnormal Activity Detection Using Video Action Recognition: A Review," IEEE Access, 2024.

[14] "Automatic Classification of Criminal Activities for Security Surveillance by Keyframes Detection and Advanced Inception Techniques," PubMed Central (PMC), 2025.

[15] "Abnormal Activity Recognition for Visual Surveillance," *Expert Syst. Appl.*, 2024.