



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 3, March 2026



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

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

Cam4U - Object & Anomaly Detection

B. Niharika, B. Vijay, J. Raghava, G. Meenakshi, K. Rajini

Department of Artificial Intelligence & Data Science, Vasireddy Venkatadri Institute of Technology,
Andhra Pradesh, India

ABSTRACT: Security and surveillance have become critical components of modern infrastructure, demanding intelligent and automated systems capable of detecting threats in real time. The Cam4U system presents an AI-powered surveillance platform that continuously monitors live video streams to detect objects, anomalies, and suspicious activities with high accuracy and low latency.

The system leverages a custom-trained YOLO26m model for real-time object detection and an I3D (Inflated 3D Convolution) model for temporal action recognition, enabling the identification of complex human behaviours and motion-based anomalies across video frames. This combination enhances detection beyond static object recognition, providing a deeper understanding of dynamic events.

To improve usability, the platform integrates an AI-powered chatbot that allows users to query system logs, retrieve detection insights, and interact with the system efficiently. Upon detecting critical events, real-time alerts are triggered via browser notifications and automated email systems, ensuring immediate response.

A responsive web-based dashboard developed using React.js enables live monitoring, event tracking, analytics, and system control. All detection events are stored with timestamps, confidence scores, and metadata for analysis and reporting.

By combining deep learning, real-time analytics, and intelligent alerting, Cam4U provides a scalable and efficient surveillance solution that significantly reduces manual monitoring efforts while improving security response capabilities.

I. INTRODUCTION

In recent years, the rapid growth of urbanization and technological advancement has significantly increased the demand for efficient and intelligent surveillance systems. Traditional surveillance approaches primarily rely on Closed-Circuit Television (CCTV) systems, where continuous human monitoring is required to identify suspicious activities. However, such systems are inherently limited due to human fatigue, delayed response times, and the inability to process large volumes of video data effectively. These limitations highlight the necessity for automated, real-time surveillance solutions that can enhance security and reduce dependency on manual monitoring.

With the emergence of Artificial Intelligence (AI) and Deep Learning (DL), computer vision-based surveillance systems have gained considerable attention. Modern AI-driven systems are capable of analyzing video streams, detecting objects, and recognizing complex patterns in real time. Despite these advancements, many existing solutions are still limited to static object detection and fail to capture temporal dynamics such as human actions and anomalous behavior, which are critical in real-world security scenarios.

To address these challenges, this paper presents Cam4U – an AI-powered real-time object and anomaly detection system. The proposed system integrates advanced deep learning models to enhance surveillance intelligence. A custom-trained YOLO-based model is utilized for high-speed and accurate object detection, while an Inflated 3D Convolutional (I3D) model is employed for temporal action recognition. This dual-model architecture enables the system to not only detect objects but also understand motion patterns and identify suspicious activities across video frames.

Furthermore, the system incorporates real-time alert mechanisms, including browser notifications and automated email alerts, ensuring immediate response to critical events. A web-based dashboard provides centralized monitoring, while an AI-powered chatbot enables users to interact with the system efficiently by querying detection logs and retrieving



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

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

insights. The platform also supports multi-camera management and comprehensive event logging for analysis and reporting.

Overall, Cam4U aims to transform conventional surveillance systems into intelligent, automated, and scalable solutions capable of improving security, minimizing human effort, and enabling faster decision-making in critical situations.

II. LITERATURE REVIEW

Literature survey is a crucial step in any software development process. Before developing a surveillance and detection system, it is important to assess various factors such as computational feasibility, real-time processing requirements, and system scalability. Once these factors are determined, the selection of appropriate deep learning models, frameworks, and deployment architectures is made to ensure successful development. The Cam4U project builds upon existing research in AI-based object detection, anomaly recognition, temporal action classification, and intelligent alert systems to deliver a robust and automated real-time surveillance solution.

1. **Joseph Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection"** (YOLO) The YOLO family of models revolutionized real-time object detection by reframing detection as a single regression problem, enabling the model to predict bounding boxes and class probabilities directly from full images in a single pass. YOLO26m further improved accuracy and speed through advanced anchor-free detection and efficient backbone architectures. This approach forms the core detection engine of Cam4U, enabling real-time object classification from live webcam feeds. However, YOLO models may struggle with detecting very small objects or heavily occluded subjects, particularly in low-resolution or poorly lit video streams.

2. **João Carreira & Andrew Zisserman, "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset"** (I3D Model) This seminal paper introduced the Inflated 3D Convolution (I3D) model, which extends 2D convolutional filters into 3D to capture temporal motion patterns across video frames. By inflating pre-trained ImageNet filters into the temporal dimension, I3D achieves state-of-the-art performance on action recognition benchmarks. The Cam4U system leverages a custom-trained I3D model to recognize complex human behaviors and motion-based anomalies beyond what static frame detection can achieve. A key limitation is that I3D requires a minimum sequence of consecutive frames to classify actions accurately, making it less effective for detecting very brief or sudden events.

3. **Mohammadi et al., "Deep Learning for Anomaly Detection: A Survey"** This comprehensive survey examines various deep learning approaches for anomaly detection in video surveillance, including autoencoders, convolutional neural networks, and recurrent architectures. The paper highlights that motion-based anomalies are best detected through temporal models that analyze sequences of frames rather than individual images. It also identifies the challenge of defining "normal" behavior across diverse environments. The Cam4U system addresses this by combining YOLO object detection for spatial analysis with I3D temporal modeling for behavioral anomaly recognition. However, anomaly detection systems trained on specific environments may produce false positives when deployed in new or different settings.

4. **Sreenu G. & Saleem Durai M.A., "Intelligent Video Surveillance: A Review through Deep Learning Techniques for Crowd Analysis"** This research reviews deep learning-based intelligent surveillance systems with a focus on crowd behavior analysis, person detection, and multi-camera tracking. The study highlights that CNN-based architectures significantly outperform traditional computer vision methods in complex real-world surveillance scenarios. It also emphasizes the importance of real-time processing pipelines and efficient alert mechanisms for practical deployment. While deep learning models achieve high accuracy, they demand significant computational resources, often requiring GPU acceleration for real-time operation — a consideration addressed in the Cam4U system architecture.

5. **Chen et al., "Real-Time Notification Systems for Intelligent Surveillance Using IoT and Cloud Computing"** This paper proposes a multi-channel notification framework for AI surveillance systems, integrating browser-based push notifications and email alert pipelines triggered by detection events. The study demonstrates that automated alert systems significantly reduce human response time compared to manual monitoring. The proposed architecture uses event-driven triggers linked to detection confidence thresholds, allowing configurable alert sensitivity. Cam4U adopts a similar approach, triggering browser and SMTP-based email alerts upon anomaly detection. A noted limitation is that notification delivery depends on network availability, which may introduce latency in low-bandwidth environments.



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

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

6. **Shridhar et al., "AI-Powered Chatbot Interfaces for Intelligent System Interaction"** This study explores the integration of conversational AI chatbots as user interfaces for complex technical systems, including surveillance platforms. NLP-driven chatbots allow non-technical users to query system logs, retrieve specific event records, and configure system parameters through natural language commands. The research shows that chatbot integration significantly improves user accessibility and reduces the learning curve of feature-rich systems. Cam4U integrates an AI chatbot to enable users to interact with detection logs and system settings conversationally. However, chatbots may struggle with highly specific or ambiguous queries involving complex multi-parameter filters.

III. REQUIREMENTS & ARCHITECTURE

3.1 Functional Requirements

User Authentication & Session Management

The platform enables users to register and log in securely via email and password, utilizing JWT-based session management to ensure robust authentication and data protection. Once authenticated, users maintain persistent sessions and are seamlessly redirected to their personalized dashboard for an optimized monitoring experience. To safeguard system resources, role-based access control is implemented, restricting administrative functionalities exclusively to authorized personnel. Furthermore, a secure logout mechanism ensures that active sessions are terminated and tokens are invalidated, effectively protecting sensitive user information.

1. Real-Time Object Detection
2. Anomaly Detection & Action Recognition
3. Alert & Notification System
4. AI Chatbot Integration
5. Camera Management
6. Database & Event Logging
7. Settings & Configuration

3.2 Non-Functional Requirements

Performance Requirements

1. Reliability & Availability
2. Usability & User Experience (UX)
3. Security & Data Integrity
4. Scalability

Technology Stack

- Frontend: React.js, TypeScript, CSS3
- Backend: Python (Flask / FastAPI)
- AI Detection Model: YOLO26m — Custom Trained
- Action Recognition Model: I3D (Inflated 3D Convolution) — Custom Trained
- Database: MongoDB
- Authentication: JWT (JSON Web Token)
- Alert System: Browser Push Notifications, SMTP Email
- Video Processing: OpenCV, WebRTC
- AI Chatbot: Grok / Gemini AI / OpenAI API
- Model Training: Python, PyTorch, Ultralytics

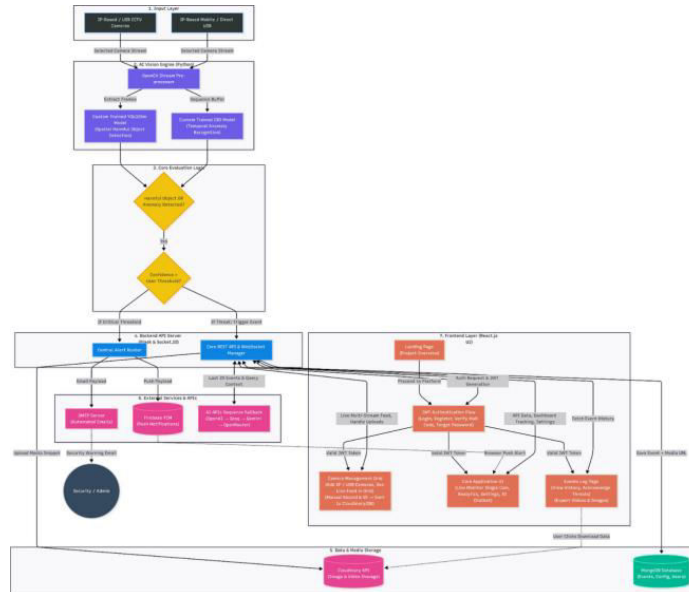
3.4.1 Architecture

The Block Diagram of the Cam4U system represents its key functional components and their structured interactions



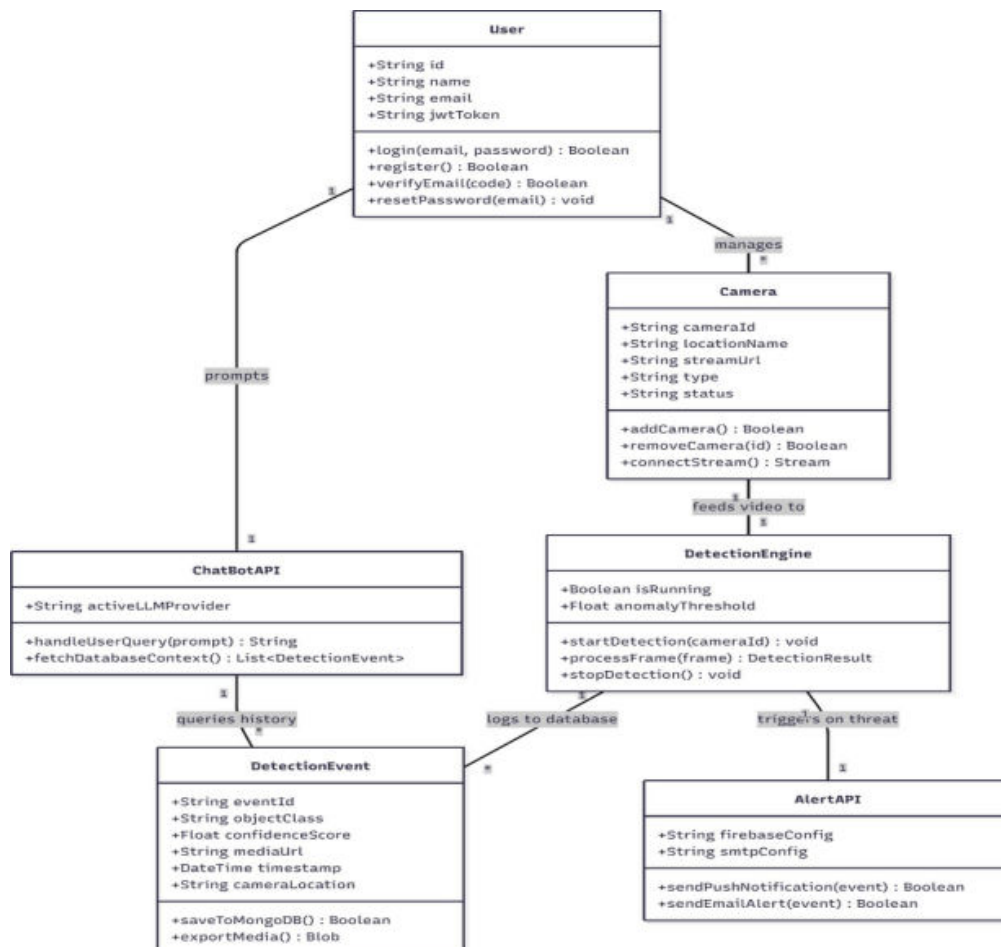
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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



Class Diagram

The Class Diagram of the Cam4U project represents the core entities and their relationships within the system.





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

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

IV. METHODOLOGY

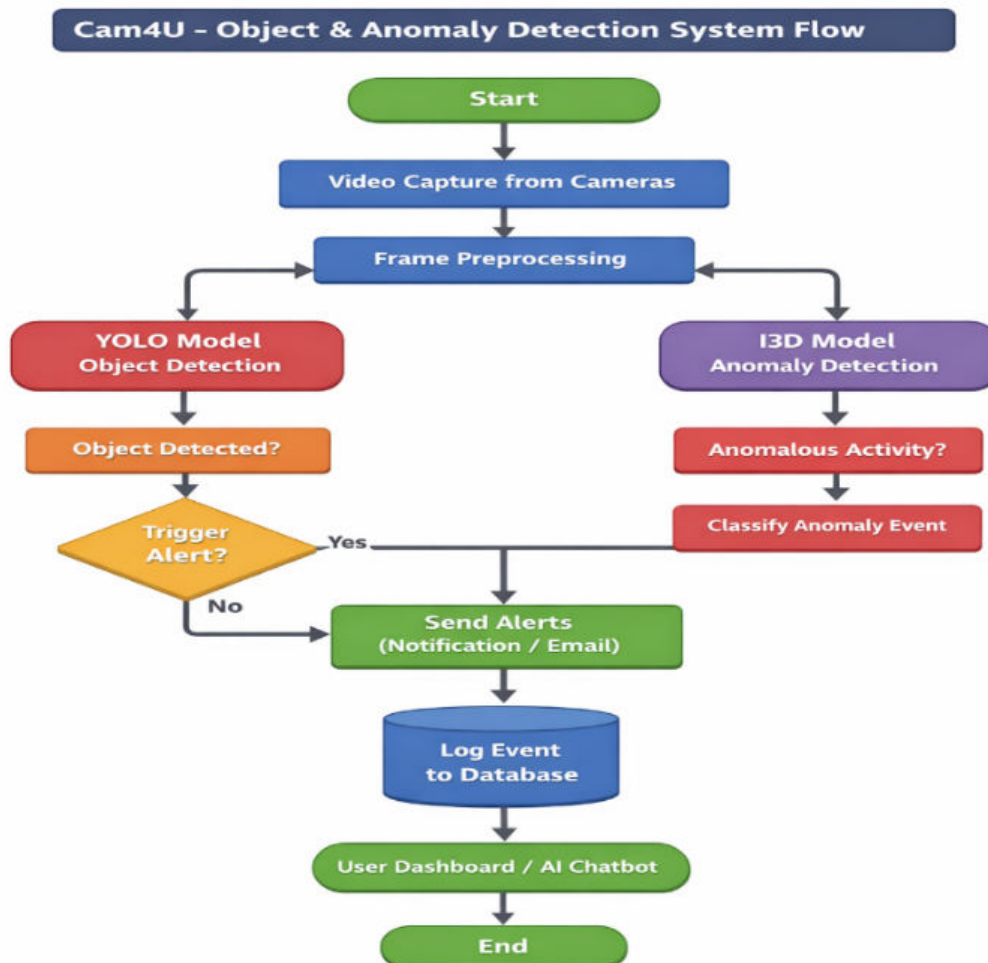
The core workflow is:

1. **Data Acquisition:** Capturing live video streams from cameras.
2. **Preprocessing:** Optimizing frames (resizing, noise reduction) for efficient computation.
3. **Object Detection (YOLO):** Identifying static objects like persons and vehicles.
4. **Temporal Analysis (I3D):** Analyzing sequences of frames to detect complex, abnormal actions (e.g., fighting, running).
5. **Decision Making:** Combining YOLO and I3D outputs with threshold logic to classify events as normal or anomalous.
6. **Alert Generation:** Sending real-time browser push notifications and email alerts for critical events.
7. **Data Storage:** Logging all event details (timestamp, class, confidence) for audit and analysis.
8. **User Interaction:** Providing a dashboard for live monitoring and an AI chatbot for system queries.

The key advantages of this dual-model approach are: combining spatial and temporal analysis, ensuring high accuracy, enabling real-time detection with low latency, and minimizing human intervention for scalable surveillance.

V. RESULT

The proposed system, **Cam4U – Object & Anomaly Detection**, was evaluated based on accuracy, efficiency, and real-time performance. The results demonstrate strong performance in object detection, anomaly recognition, and alert generation.





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

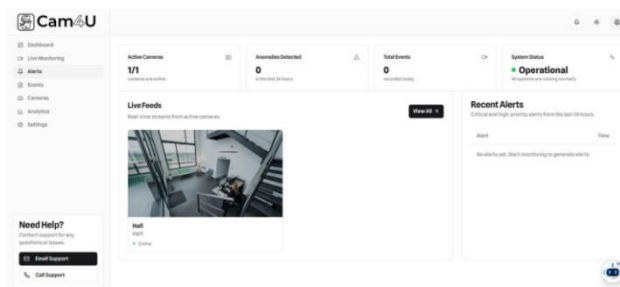
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. Performance Evaluation Table

Metric	YOLO Model (Object Detection)	I3D Model (Anomaly Detection)
Precision	86.7%	90.38%
Recall	74.7%	98.95%
Accuracy	—	90.00%
mAP (Detection Quality)	82.1%	—
Inference Time	~35 ms/frame	Sequence-based
FPS	25–30 FPS	Real-time (batch processing)

This table shows that:YOLO performs well in **fast object detection**

Fig. 5.1 – Cam4U User Interface & Dashboard)



- I3D performs exceptionally in **capturing anomalies with high recall**

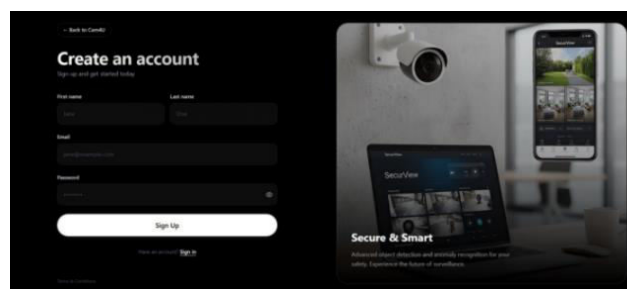
SYSTEM DESIGN

5.1 User Interface & Dashboard

The Dashboard is the core operational hub of the Cam4U platform, providing users with a centralized view of all surveillance activities.

5.1.1 User Registration

User authentication in Cam4U ensures secure access to the surveillance platform while providing a straightforward onboarding experience. The first step involves creating an account manually via email and password credentials.



(Fig. 5.2 – Cam4U User Registration)



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

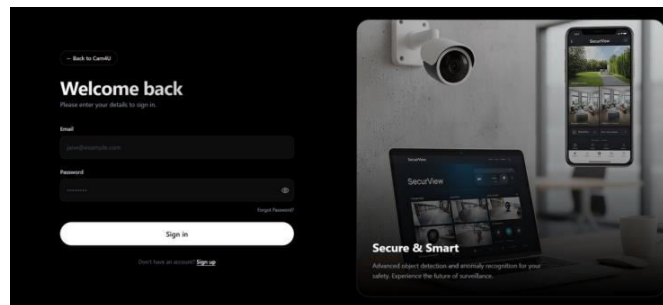
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

5.1.2 User Login & Authentication

Once an account is registered, the system provides a streamlined login process to grant users authenticated access to their respective dashboards.

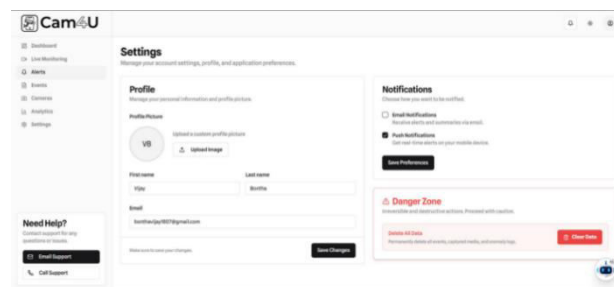
- Login Process
- Session Management
- Role-Based Access Control

(Fig. 5.3 – Cam4U User Login & Authentication)

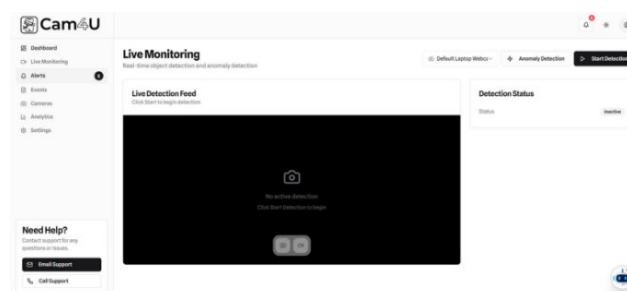


5.1.3 Settings & Configuration Panel

The Settings Page provides users with full control over system configuration and personalization. From this panel, users can manage alert thresholds, configure email recipients for automated alerts, adjust detection sensitivity levels, manage camera source configurations, and update personal account details.



(Fig. 5.4 – Cam4U Settings)



(Fig. 5.5 – Cam4U Real-Time Surveillance & Live Monitoring Dashboard)

5.1.4 YOLO26m Model Training

The YOLO26m model was custom-trained for real-time object detection using a domain-specific dataset consisting of nine classes relevant to surveillance scenarios.



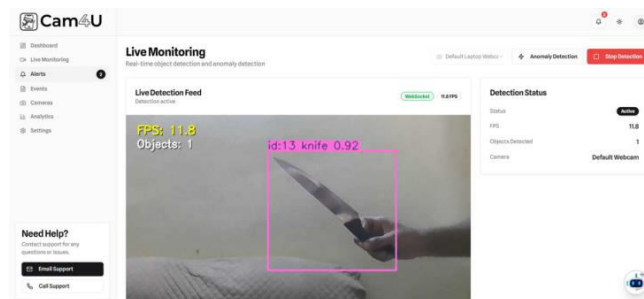
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

The classes include:

- Face Mask
- Hammer
- Axe
- Cricket Bat
- Handgun
- Knife
- Mobile
- Person
- Sickle

The dataset contains annotated images for each class with varying object scales and real-world conditions. The model was trained using the Ultralytics YOLO framework on an NVIDIA L4 GPU with an input image size of 768×768 . The model demonstrates strong detection performance across multiple object categories, particularly for safety-related objects such as weapons and human presence.



(Fig. 5.6 – Custom YOLO26m Detection Output)

5.1.5 I3D Model Training

The I3D (Inflated 3D Convolution) model was trained using the UCF-Crime dataset, a large-scale real-world surveillance dataset designed for anomaly detection in videos.

The dataset consists of multiple anomaly categories, including:

- Explosion
- Robbery
- Shooting
- Stealing
- Vandalism
- Abuse
- Arson
- Shoplifting
- Arrest
- Assault
- Burglary
- Fighting
- Road Accidents
- Normal Videos (non-anomalous events)

Each video is processed as a sequence of frames to capture temporal motion patterns. The I3D model learns to distinguish between normal and abnormal activities by analyzing spatio-temporal features across consecutive frames.

The use of the UCF-Crime dataset enables the system to detect complex real-world anomalies such as violent activities, theft, and accidents with high recall and accuracy.

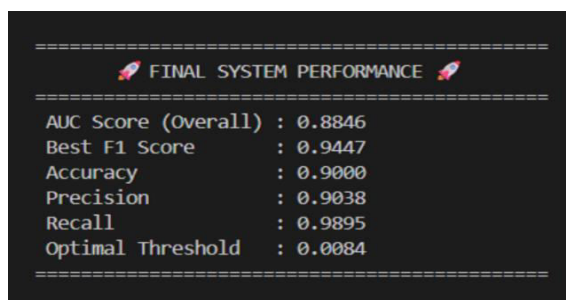


International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

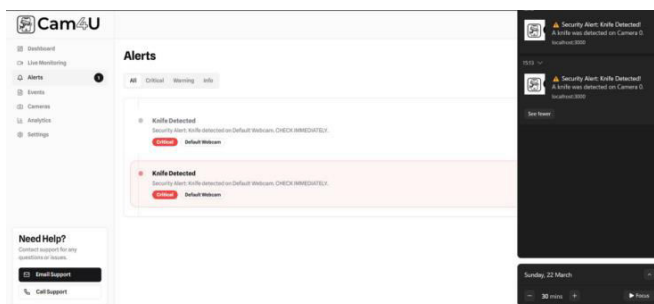
5.1.5.1 I3D Performance Metrics

- Accuracy: 90.00% (Overall correct classification rate).
- Recall: 98.95% (Extremely high recall, showing that the system successfully identifies nearly every anomaly instance occurrence in the test set).
- Precision: 90.38% (Indicates high reliability in flagging anomalous events with minimal false positives).
- Best F1 Score: 0.9447 (The weighted harmonic mean of precision and recall, showing a balanced and robust performance).
- AUC Score (Overall): 0.8846 (Area Under the Curve, representing the model's overall capability to distinguish between classes at various threshold settings).
- Optimal Evaluation Threshold: 0.0884 (The specific confidence threshold where the system achieves the optimal balance for security detection).



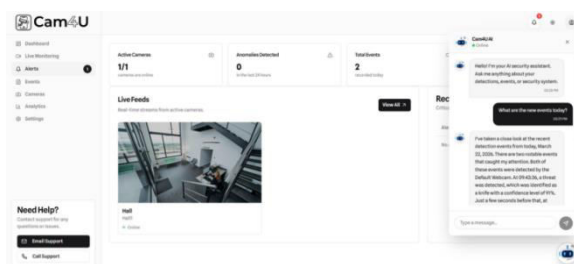
(Fig. 5.7 – Cam4U Performance)

Results



(Fig. 5.8 – Alerts Page & History View)

Browser Push Notifications



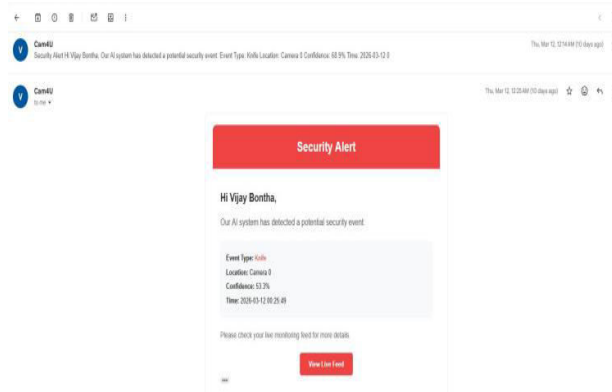
(Fig. 5.9 – Browser Push Notification)



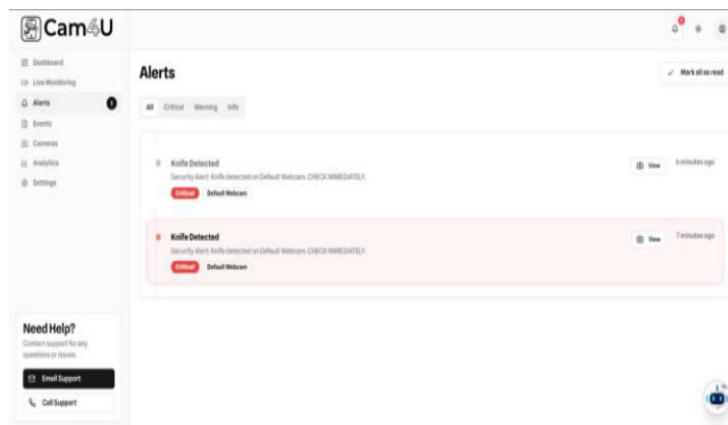
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

Automated Email Alerts

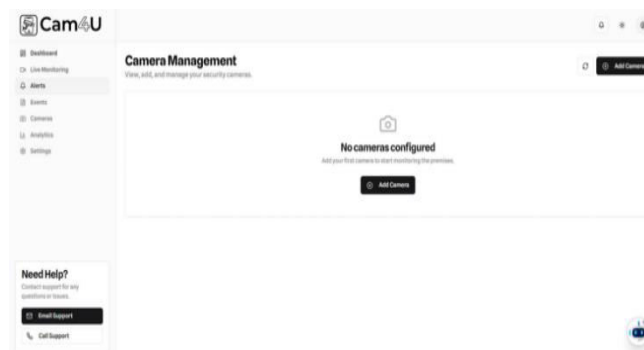


(Fig. 5.10 – Automated Email Alert)



(Fig. 5.12 – AI Chatbot Interface)

MODULE 5 – CAMERA MANAGEMENT SYSTEM

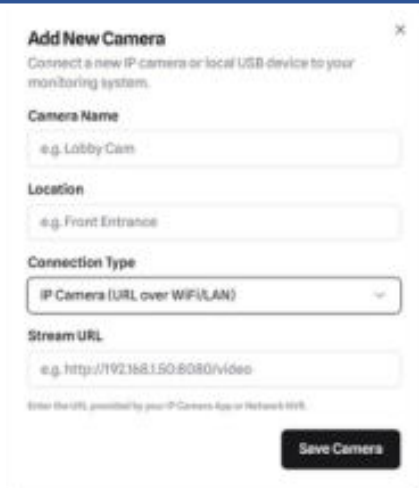


(Fig. 5.13 – Camera Management Panel)



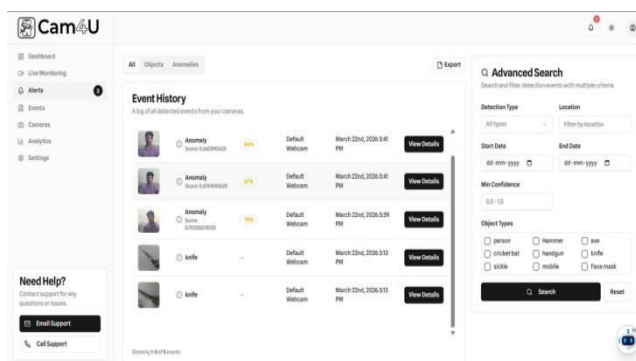
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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



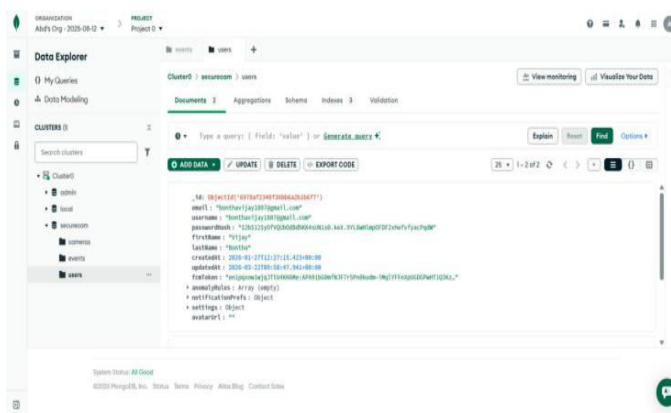
(Fig. 5.14 – Add / Configure New Camera)

Events Page & Detection History



(Fig. 5.11 – Events Page & Detection History)

5.1.6 Database Schema & Storage



(Fig. 5.15 – Database Schema Design)

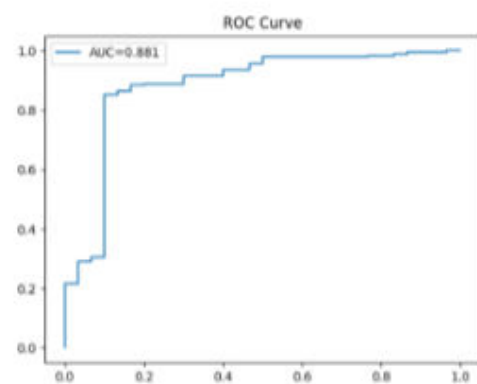
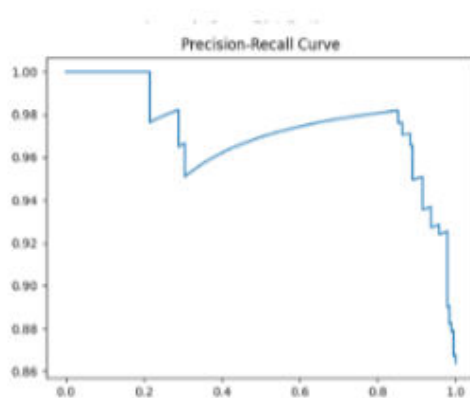
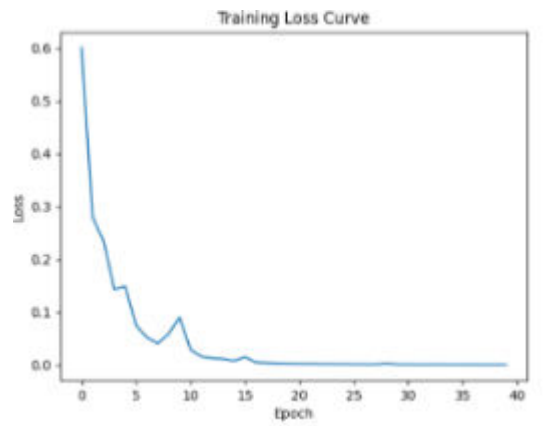


International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

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

Graphs

Metric	Value
AUC	0.8810526316
F1	0.9514066496
Accuracy	0.9136363636
Precision	0.9253731343
Recall	0.9789473684



VI. CONCLUSIONS

The proposed system, Cam4U – Object & Anomaly Detection, successfully demonstrates the effective application of Artificial Intelligence and Deep Learning in real-time surveillance systems. By integrating a custom-trained YOLO model for object detection and an I3D model for temporal action recognition, the system achieves a robust dual-layer detection mechanism capable of identifying both static objects and dynamic anomalies in live video streams.

The system overcomes the limitations of traditional surveillance methods by eliminating continuous manual monitoring and enabling automated, intelligent analysis of video data. The real-time alert mechanism, including browser notifications and email alerts, ensures immediate response to critical events, significantly improving security efficiency and reducing response time.

Experimental results indicate that the system achieves high accuracy and real-time performance, with strong precision in object detection and exceptionally high recall in anomaly detection. The integration of a web-based dashboard and AI chatbot further enhances usability, allowing users to monitor, analyze, and interact with the system effectively.

Additionally, the scalable architecture supports multiple camera inputs and efficient data management through centralized event logging, making the system suitable for various real-world applications such as homes, offices, educational institutions, and public spaces.

In conclusion, Cam4U provides a smart, automated, and scalable surveillance solution that improves detection accuracy, reduces human effort, and enables faster decision-making in security-critical environments. The system lays a strong foundation for future enhancements, including predictive analytics, edge deployment, multi-model integration.



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

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

REFERENCES

[A] Core Research & AI Models

- [1] W. Sultani, C. Chen, and M. Shah, "Real-world Anomaly Detection in Surveillance Videos," CVPR, 2018. (Foundation for your MIL anomaly detection).
- [2] J. Carreira and A. Zisserman, "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset," CVPR, 2017. (The I3D architecture for feature extraction).
- [3] M. Ilse, J. M. Tomczak, and M. Welling, "Attention-based Deep Multiple Instance Learning," ICML, 2018. (Basis for your Attention-Hybrid MIL model).
- [4] UCF-Crime Dataset, University of Central Florida, "Real-world Anomaly Detection in Surveillance Videos." (Used for training/validation).
- [5] Chuan Guo et al., "On Calibration of Modern Neural Networks," ICML, 2017. (Relevant to your AnomalyDetector calibration logic).
- [6] Andreas Rossler et al., "FaceForensics++: Learning to Detect Manipulated Facial Images," ICCV, 2019.

[B] Backend Infrastructure & AI Services

- [7] Ultralytics YOLO Framework, Glenn Jocher et al., "Real-time Object Detection and Tracking," <https://github.com/ultralytics/ultralytics>.
- [8] Flask Web Framework, Pallets Projects, <https://flask.palletsprojects.com>, Version 3.0+. (Core REST API engine).
- [9] OpenCV (OpenSource Computer Vision Library), <https://opencv.org>. (CRITICAL: Used for all webcam streaming and frame processing).
- [10] MongoDB Documentation, MongoDB Inc., <https://www.mongodb.com/docs>. (Used for user profiles and event storage).
- [11] Cloudinary SDK, <https://cloudinary.com/documentation>. (Used for detection media storage).
- [12] Scikit-learn, "Machine Learning in Python," <https://scikit-learn.org>. (Used for ROC-AUC and model evaluation metrics).
- [13] Matplotlib, "Visualization with Python," <https://matplotlib.org>. (Used to generate your evaluation plots and curves).
- [14] OpenRouter AI / OpenAI, <https://openrouter.ai/docs>. (Powers your security chatbot feature).
- [15] PyTorch & TorchVision, Meta AI, <https://pytorch.org/docs>. (Deep learning backend).

[C] Frontend & Real-time Integration

- [16] Next.js Documentation, Vercel Inc., <https://nextjs.org/docs>, Version 14.x.
- [17] Tailwind CSS & Shaden UI, <https://tailwindcss.com>.
- [18] Recharts, <https://recharts.org>. (Used for your analytics dashboard visualizations).
- [19] Firebase Cloud Messaging (FCM), Google, <https://firebase.google.com/docs/cloud-messaging>. (Used for push notifications).
- [20] Framer Motion & GSAP, <https://www.framer.com/motion>. (Used for UI transitions).
- [21] Socket.IO, <https://socket.io/docs>. (Used for real-time detection alerts in the UI).
- [22] Lucide React, <https://lucide.dev>. (UI icons).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

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

www.ijmrset.com