

Surveil Guard: Intelligent Surveillance - Detecting, Reporting, and Alerting

S. Hariprasath¹, A. Harihara Sudhan², S. Keerthi Vasagan³, P. Roy Sudha Reetha⁴

^{1,2,3}UG Scholar, Department of AI & DS, PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India.

⁴Assistant Professor, Department of AI & DS, PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India.

Emails: r.venkatakrishna@lords.ac.in¹, roshanhashaik@gmail.com², bprameeth123@gmail.com³, safoorayasmeen17@gmail.com⁴

Abstract

As security challenges evolve, traditional surveillance systems often fall short in effectively identifying and responding to real-time threats. This paper introduces SurveilGuard, a novel, AI-powered surveillance framework designed to autonomously detect, report, and respond to abnormal activities such as fighting, smoking, hugging, and other predefined behaviors in real time. Leveraging motion-triggered camera activation via the ESP32-CAM and powerful anomaly detection models like YOLO for activity recognition, SurveilGuard offers a seamless integration of video capture, behavior analysis, and incident reporting. When abnormal activities are detected, the system sends real-time alerts using Twilio and automatically generates detailed reports enriched with video evidence and contextual data using CLIP for image-text matching. These reports are securely stored and easily accessible via a web interface for authorized personnel, enhancing situational awareness and operational response. Additionally, Power BI is employed for data visualization, allowing for comprehensive reporting and interactive dashboards. SurveilGuard represents a significant advancement in automated surveillance, offering a scalable solution for real-time security monitoring with minimal false alarms, empowering security teams to respond quickly and effectively.

Keywords: Intelligent Surveillance, Anomaly Detection, Real-Time Reporting, Automated Security, Video Analytics, YOLO, Twilio, CLIP, Power BI.

1. Introduction

In today's world of evolving security challenges, traditional surveillance systems relying on static cameras and passive monitoring often fall short. They struggle with real-time threat detection, generate high false alarm rates, and lack autonomous responses, overwhelming security personnel and delaying critical incident resolution. To address these issues, SurveilGuard offers an AI-powered surveillance framework that revolutionizes security through motion-triggered camera activation (ESP32-CAM), real-time anomaly detection (YOLO models), and instant notifications via Twilio. It detects predefined behaviors like fighting and smoking, sending timely alerts and actionable insights to security teams. SurveilGuard analyzes live video feeds in real-time, triggering instant notifications and generating comprehensive incident reports enriched with video

evidence and contextual data using CLIP for image-text matching. This automation minimizes human error, accelerates responses, and reduces workload. A secure web interface enhances situational awareness, allowing access to real-time alerts, reports, and analytics, while Power BI integration offers interactive dashboards for data-driven decision-making. Scalable and customizable, SurveilGuard is suitable for diverse environments, from public spaces to critical infrastructure. This report outlines its design, implementation, and future enhancements, such as Wi-Fi-based bullying detection for expanded coverage in non-camera zones. SurveilGuard redefines modern surveillance, empowering security teams to stay ahead of threats effectively [1-4].

1.1. Problem Definition

Traditional surveillance systems face significant

limitations that hinder their effectiveness in addressing modern security challenges. These include high false alarm rates caused by non-threatening activities such as pets or lighting changes, delayed response times due to manual reviews, and the inability to provide real-time reporting or automated incident documentation. Additionally, these systems lack contextual awareness, often failing to distinguish normal activities from threats, and suffer from blind spots due to fixed camera coverage. Continuous recording also leads to excessive storage consumption, while the absence of AI-based analysis limits the identification of complex behaviors. Furthermore, delayed threat detection slows response times, and the lack of data visualization tools prevents insights into trends or vulnerabilities. SurveilGuard addresses these issues through real-time detection of abnormal behaviors using AI models like YOLO, automated reporting enriched with video evidence and metadata, instant alerts via Twilio, secure cloud-based storage accessible through a web interface, and interactive Power BI dashboards for comprehensive trend analysis and insights [5-7].

1.2. System Overview

- Real-Time Detection:** AI models like YOLO for abnormal behavior recognition.
- Automated Reporting:** Detailed reports with video evidence and metadata.
- Alerting and Response:** Instant alerts via Twilio for quick action.
- Data Storage and Access:** Secure, cloud-based storage accessible via web interface.
- Data Visualization:** Power BI dashboards for trends and analysis.

1.3. Assumptions and Dependencies

1.3.1. Assumptions

- Stable Wi-Fi for communication.
- High-quality dataset for AI training.
- Strong security protocols for data protection.

1.3.2. Dependencies

- Reliable hardware (e.g., ESP32-CAM).
- Twilio for real-time alerts.
- Power BI for data visualization.
- Regular updates to AI models like YOLO.

- Compliance with legal and ethical regulations [8-11].

2. System Architecture

The system integrates ESP32-CAM, motion sensors, YOLO, CLIP, Twilio, and Power BI to provide a cost-effective and real-time surveillance solution. Motion sensors activate the ESP32-CAM only when motion is detected, while YOLO identifies activities and CLIP generates detailed reports. Twilio sends real-time alerts, and Power BI visualizes data for analysis and decision-making, creating a scalable and efficient security system (as described in figure 1).

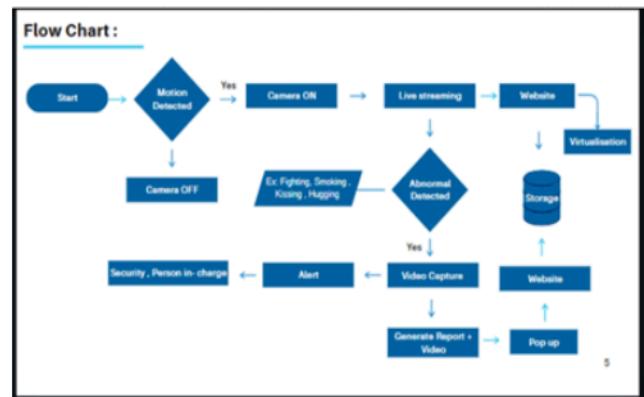


Figure 1 System Architecture

2.1. Module Explanation

2.1.1. Camera and Motion Detection Module

Uses ESP32-CAM and motion sensors to capture the video only during activity, reducing power and the storage usage while ensuring essential events are recorded.

2.1.2. Activity Recognition and Anomaly Detection Module

YOLO detects activities like smoking, while AI/ML models classify them as normal or abnormal, ensuring accurate and fast detection with minimal false positives.

2.1.3. Alerting and Notification Module

Twilio sends alerts via SMS, email, or calls for detected anomalies, with configurable triggers and event details like time, location, and activity type.

2.1.4. Report Generation and Storage Module

CLIP creates detailed, automated reports with visual

evidence and descriptions, securely stored for easy access and analysis.

2.1.5. Data Visualization Module

Power BI visualizes trends and patterns in incidents, providing actionable insights for security teams through interactive dashboards.

2.1.6. Web Interface and Access Control Module

A secure web interface offers live video, reports, and analytics, with role-based access ensuring data confidentiality and efficient management.

2.1.7. System Integration and Communication Module

APIs and middleware ensure the seamless communication between components, enabling real-time data flow, synchronization, and system reliability. This system architecture and its modules deliver an intelligent, responsive, and resource-efficient surveillance solution.

3. System Specifications

3.1. Hardware Requirements

As displayed in the figure 2, figure 3, figure 4 you can observe,

- **ESP32-CAM:** Motion-triggered video recording.
- **PIR Motion Sensor:** Activates camera upon detection.
- **Power Supply:** Battery or USB.
- **Computer/Server:** Hosts AI models and web apps. Minimum specs: Intel i5, 8GB RAM, 250GB SSD, NVIDIA GTX 1050 (for local models).
- **Network:** Wi-Fi for connectivity and internet for cloud integration. The illustration in figure 3 classifies the live stream.



Figure 2 IoT Kit Images



Figure 3 IoT Kit Images

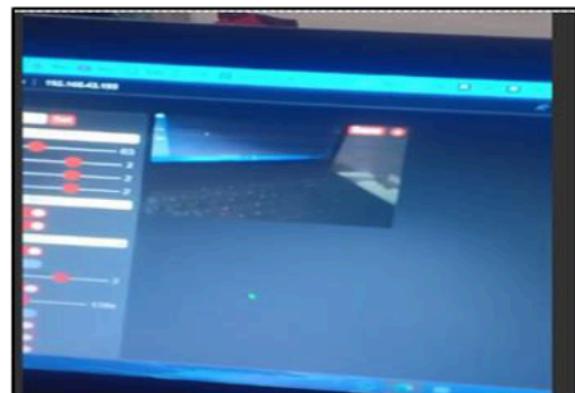


Figure 4 Live Stream

3.2. Software Requirements

- **OS:** Windows 10/11 or Ubuntu 20.04+.
- **Languages:** Python (AI/backend), JavaScript (frontend).
- **Frameworks:** Django (web), OpenCV (video processing), YOLO (activity recognition), CLIP (reporting), Twilio API (alerts), Power BI (visualization).
- **Database:** MongoDB, Firebase.
- **Tools:** VS Code, APIs for integration.

3.3. Software Description

- **Camera and Motion Detection** - ESP32-CAM streams video upon motion, processed via OpenCV.
- **Activity and Anomaly Detection** - YOLO identifies activities, and custom models classify anomalies.
- **Alert and Notification** - Twilio API sends real-time alerts via SMS or email.
- **Web Interface** - Built with Django to manage video reviews, alerts, and visualizations.



Figure 5 Sample Image

The sample image in Figure 5 demonstrates real-time object detection with bounding boxes highlighting identified individuals

- **Report Generation** - CLIP converts detected anomalies into descriptive, actionable reports.
- **Data Visualization** - Power BI provides interactive dashboards for trend analysis.

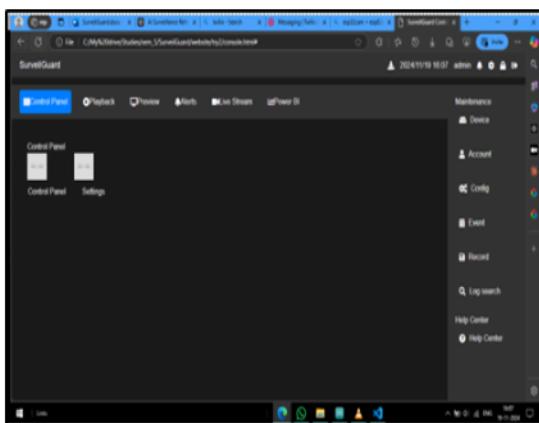


Figure 6 Website

Figure 6 illustrates the analysis dashboard on the website, showcasing metrics such as accuracy, timeline, and predictions.

3.4. Dataset Description

SourceVisual datasets sourced from Roboflow; textual datasets from MS COCO, Visual Genome, and custom annotations. Structure

- **Visual Data:** 2,579 images (train: 77%, validate: 15%, test: 8%) annotated with bounding boxes for activities.
- **Textual Data:** Captions align with images for context.

Examples:

- Image: A group of people fighting. Text: "Multiple individuals in a public altercation." Figure 7 outlines the dataset composition, including visual and textual data with annotations.

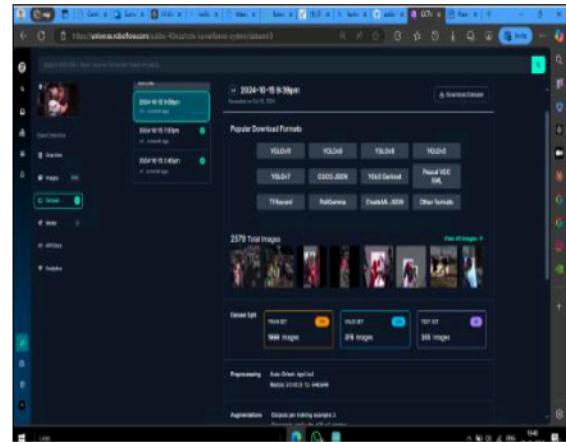


Figure 7 Dataset

4. Results

4.1. Anomaly Detection

- **Accuracy:** Training: 96.8%, Validation: 94.3%, Testing: 92.1%.
- **Precision (mAP@0.5):** 91.5%

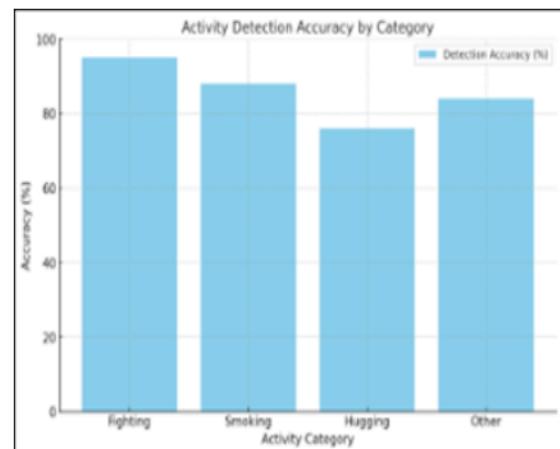


Figure 8 Detection Accuracy Graph

Figures 8 and 9 present performance metrics, showcasing detection accuracy and precision vs. recall for activity detection." Figure 10 depicts the model's accuracy through the mean average precision (mAP) graph.

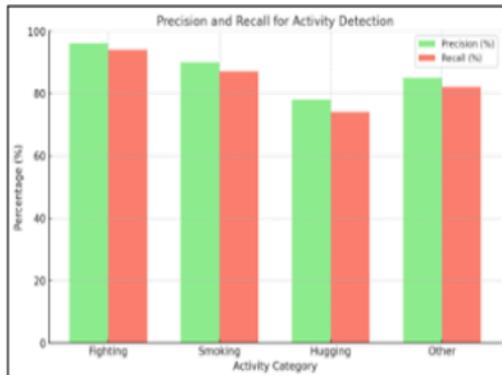


Figure 9 Precision vs Recall Graph

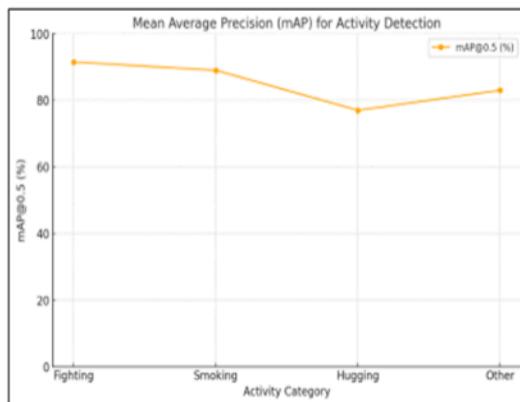


Figure 10 Mean Average Precision (mAP) Graph

4.2. Report Generation

- Accuracy:** CLIP achieved 90.2%.
- Example:** "Individual smoking in a restricted area. Immediate attention required."

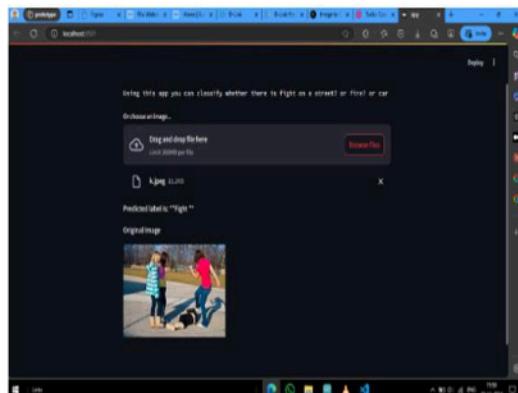


Figure 11 Detection Using Clip

Figure 11 highlights CLIP's detection capabilities, generating alerts for restricted area violations.

4.3. Real-Time Alerting

Alert delivery: 1.2 seconds, 99% reliability.

4.4. Data Visualization

Power BI dashboards analyze trends, map incidents, and log user access.

4.5. Testing

- Test Case 1:** Activity Detection
- Input:** Video of a fight.
- Output:** Alert and report with contextual details. Result: Passed.

Figure 12 showcases the alert notification system, providing real-time alerts with contextual details based on detected activities.

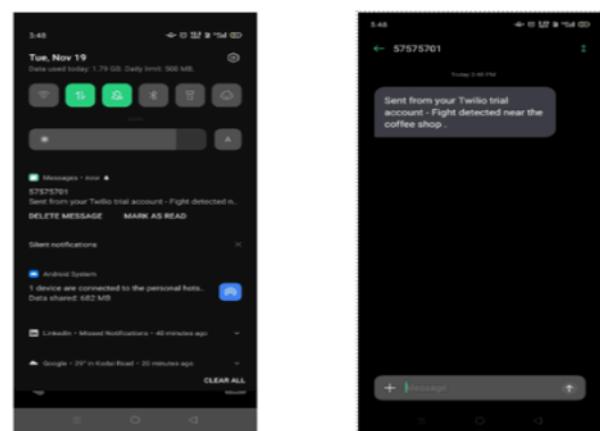


Figure 12 Alert Notification

- Test Case 2:** Web Access Control
- Input:** Login attempts by authorized/unauthorized users.
- Output:** Access granted/denied appropriately. Result: Passed.

These results demonstrate system reliability in real-time surveillance and reporting. Marketing Website (Figure 12): Features the system's capabilities and use cases, shown in Figure 13.

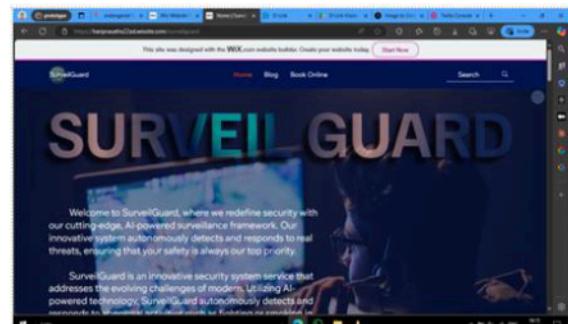


Figure 13 Marketing Website

Conclusion

SurveilGuard is an advanced surveillance system designed to improve security monitoring using cutting-edge technology. It combines real-time threat detection (YOLO), automated alerts (Twilio), intelligent reporting (CLIP), and insightful data visualization (Power BI). These features create a reliable, efficient, and scalable solution for diverse environments like small businesses, schools, and large industries. The system optimizes performance with motion-triggered video recording and reduces unnecessary notifications, saving storage and enhancing efficiency. SurveilGuard also prioritizes privacy and complies with legal standards, ensuring sensitive data is secure and accessible only to authorized users. With automated reporting and data insights, security teams can respond to incidents faster and analyze long-term trends effectively. SurveilGuard is a robust, ethical, and future-ready solution for modern security challenges [12].

Literature Survey

The literature on modern surveillance systems highlights advancements in integrating IoT, AI, and deep learning to enhance security and real-time monitoring. Technologies like the ESP32-CAM enable low-cost, motion-triggered video surveillance, while deep learning algorithms like YOLO allow for accurate real-time object detection of behaviors such as fighting or smoking. Real-time alerting systems, using tools like Twilio, improve response times by notifying security teams instantly. Models like CLIP generate contextual reports from video footage, aiding decision-making with detailed incident descriptions. Data visualization platforms like Power BI provide interactive dashboards for tracking trends and improving operational insights. Hybrid systems combining motion sensors and AI reduce false alarms, and IoT integration enables scalable, flexible monitoring across large areas. Together, these innovations form the backbone of intelligent, automated surveillance solutions for enhanced security and responsiveness.

Future Enhancements

A significant upgrade for SurveilGuard is the integration of Wi-Fi-based activity detection, especially for identifying physical bullying. This

method uses changes in Wi-Fi signals caused by movement to detect aggressive behaviors, even in areas without cameras. This enhancement, combined with existing video analytics, transforms SurveilGuard into a hybrid system capable of broader, non-invasive surveillance. It is ideal for schools, workplaces, and public spaces where privacy and cost are significant concerns. By incorporating this feature, SurveilGuard will offer enhanced adaptability, making it an even more comprehensive and efficient security solution.

References

- [1]. Sayanti Adhya, Gobinda Sen " IOT Based Home Security System," 2023 7th International Conference on Electronics, IEMENTech, DOI:10.1109 /IEMEN Tech 604 02.2023.10423552, Electronic ISBN: 979-8-3503-0551-7, PoD ISBN: 979-8-3503-2894-3.
- [2]. Epaminondas S. Lage, Rodolfo L. Santos, Sandro M.T. Junior, Fernando Andreotti " Low-Cost IoT Surveillance System Using Hardware-Acceleration and Convolutional Neural Networks " 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland , DOI: 10.1109/WF-IoT.2019.8767325, Electronic ISBN:978-1-5386-4980-0 , PoD ISBN:978-1-5386-4981-7
- [3]. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi "You Only Look Once: Unified, Real-Time Object Detection," arXiv:1506.02640 [cs.CV] , (or arXiv:1506.02640v5 [cs.CV] for this version)
- [4]. Kanyifechukwu Jane Oguine, Ozioma Collins Oguine, Hashim Ibrahim Bisallah " YOLO v3: Visual and Real-Time Object Detection Model for Smart Surveillance Systems(3s),"arXiv:2209.12447 [cs.CV] (or arXiv:2209.12447v1 [cs. CV] for this version)
<https://doi.org/10.48550/arXiv.2209.12447>
- [5]. Kennedy Okokpujie, Imhade P. Okokpujie Fortune T. Young ,Roselyn E. Subair " Development of an Affordable Real-Time IoT-Based Surveillance System Using ESP32

- and TWILIO API," © 2023 IIETA. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>). DOI: <https://doi.org/10.18280/ijssse.130609>
- [6]. Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, Ilya Sutskever , "Learning Transferable Visual Models From Natural Language Supervision," arXiv:2103.00020 [cs.CV], (or arXiv: 2103.00020v1 [cs.CV] for this version), <https://doi.org/10.48550/arXiv.2103.00020>
- [7]. Alpay, T., Magg, S., Broze, P. et al. Multimodal video retrieval with CLIP: a user study. *Inf Retrieval J* 26,6(2023).<https://doi.org/10.1007/s10791-023-0942> 5 -
DOI:<https://doi.org/10.1007/s10791-023-09425-2>
- [8]. Sarita Sanap, Khishamuddin Syed " Real-Time Data Visualization in PWAs using Power BI Embedded " 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). DOI:10.1109/ICCCNT61 001 .2024.10723938 , Electronic ISBN:9 79- 8-3503-7024-9 ,PoD ISBN:979-8-3503-70 27-0, Electronic ISSN: 2473-7674 ,PoD) ISSN: 2162-7665
- [9]. Prof. Krunal Pawar , Prof. S.K.Honawad ,Prof. S.S.Chinchali , Prof. Pradeep Deshpande " Smart Home Security Surveillance System Using Motion Detection and IOT," IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727 PP 21-24
- [10]. Huu-Thanh Duong, Viet-Tuan Le and Vinh Truong Hoang " Deep Learning-Based Anomaly Detection in Video Surveillance: A Survey " © 2023 by the authors. Licensee MDPI, Basel, Switzerland. Sensors 2023, 23(11), 5024; <https://doi.org/10.3390/sensors2311502>
- [11]. Abhijeet Singh, Parag Gupta, Vaidehi Sankhe, Maya Patil " An IoT Based Approach to Smart Surveillance System," 2023 International Conference on IoT, Communication and Automation Technology (ICICAT) , Electronic ISBN: 979-8-3503-0282-0 , PoD ISBN: 979-8-3503-0283-7, DOI: 10.1109/ICI-CAT57 735.2023.10263602
- [12]. Bo Lan, Fei Wang, Lekun Xia, Fan Nai, Shiqiang Nie, Han Ding , " Detecting School Physical Bullying With Wi-Fi and Deep Wavelet Transformer ", IEEE Internet of Things Journal , DOI: 10.1109/JIOT.2024.3486071, Electronic ISSN: 2327-4662, CD: 2372-2541