NEURAL SENTINAL -ADAPTIVE AI FOR REAL-TIME INTRUSION PREVENTION

K. Pragash[1] ,M. Sivapriyan[2] , M. Karthikeyan[3] , J. Aatif Ahamed[4] , M. Mohammed Feroz[5]

[1]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, pragashkaliyan@gmail.com;
[2]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, sivapriyan7siva@gmail.com;
[3]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, mkarthikeyan00100010@gmail.com;
[4]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, amal.aathif@gmail.com;
[5]Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College, Puducherry, feroz1522krish@gmail.com;

Abstract:

Unauthorized access to sensitive areas is a critical security issue in various sectors, including government, military, and private properties. This paper presents a computer vision-based surveillance system designed to prevent unauthorized access in real-time. Leveraging advanced techniques such as object detection, facial recognition, and motion analysis, the proposed system continuously monitors restricted zones, automatically detecting and responding to intrusions. The solution improves on traditional surveillance by offering real- time alerts, automated monitoring, and reduced human intervention. By utilizing machine learning models, the system adapts to new security threats, ensuring robustness across varied environments. This scalable solution aims to enhance security by integrating seamlessly into existing infrastructures.

Keywords: Computer vision, restricted area prevention, object detection, real-time monitoring, automated surveillance, facial recognition, anomaly detection.

1. Introduction:

Security breaches in restricted and sensitive areas represent a critical threat to assets, infrastructure, and personnel. Incidents of unauthorized access can lead to significant disruptions and potential harm, particularly in sectors like government, military, industrial, and private facilities where sensitive information or valuable assets are often stored. Traditional security measures such as CCTV and manual patrols have limitations; these systems primarily rely on human operators to detect and respond to incidents, often leading to delays and oversight [4][18]. Surveillance systems today still tend to be reactive rather than

proactive, addressing incidents only after they have occurred, and are limited by human fatigue and error in monitoring [8][11]. This paper introduces an intelligent, automated surveillance system that leverages computer vision techniques to continuously monitor restricted zones. By combining object detection, facial recognition, and motion analysis, the proposed system identifies potential intrusions in real-time and can distinguish between routine activities and suspicious behaviors [1][3].

2. Literature Survey:

Recent advancements in machine learning and computer vision have transformed the capabilities of automated surveillance and security systems, enabling more accurate and efficient monitoring [18][6]. A substantial body of research focuses on object detection and recognition to identify intruders or unauthorized objects in real time. Redmon et al.'s work on YOLO (You Only Look Once) revolutionized object detection by offering high-speed recognition, enabling practical, real-time applications [1]. Similarly, Schroff et al.'s development of FaceNet enhanced facial recognition, allowing systems to verify identities with high accuracy, which is particularly valuable in restricted area applications where access control is crucial [3].

In addition to object and facial recognition, behavioral analysis through anomaly detection plays a vital role in modern security systems. Techniques like Autoencoders and Isolation Forests have been explored to identify unusual patterns or behaviors [19]. These models analyze motion patterns to detect loitering or sudden, unexpected movements, which often precede unauthorized access attempts [4][12].

3. Proposed System Model:

The proposed system integrates cutting-edge computer vision and machine learning methodologies to develop a reliable and intelligent solution for restricted area surveillance.

Each component addresses a specific aspect of unauthorized access detection, collectively enabling continuous, real-time monitoring of restricted zones with minimal human intervention. This section details the key modules of the system, including object detection, motion analysis, facial recognition, and anomaly detection.

3.1 Object Detection

This system employs YOLO for real-time object detection, which is essential for identifying unauthorized individuals, vehicles, or other objects within surveillance zones. YOLO's architecture enables it to process entire images in a single forward pass, providing fast and accurate detection [1]. The YOLO model is particularly suitable for security applications due to its high speed and efficiency, allowing the system to monitor multiple zones simultaneously with minimal latency [1][12].

To optimize detection accuracy, the YOLO model is trained on large, diverse datasets that are representative of real-world conditions, including various lighting, weather, and environmental changes [18]. In addition, the training process includes both authorized and unauthorized entities, allowing the model to differentiate between potential intruders and regular personnel or vehicles. To further reduce false positives, the system incorporates additional contextual data, such as location and time, to fine-tune detection accuracy based on specific security protocols.

3.2 Motion Analysis

The motion analysis module uses optical flow and background subtraction techniques to track movement within the surveillance area [4]. Optical flow algorithms calculate the motion of objects between consecutive frames, capturing directional and velocity information [4][8]. This is particularly useful for identifying suspicious movement patterns, such as loitering or rapid, erratic motion that might indicate a security breach.

Background subtraction enhances motion detection by isolating moving objects from static elements in the scene [8][14]. This technique subtracts the current frame from a reference background image, identifying changes that may signify unauthorized activity. Combined with optical flow, background subtraction allows the system to continuously monitor movement and detect anomalies. For example, the system can distinguish between normal pedestrian flow and individuals lingering in restricted areas, which may warrant further inspection.

3.3 Facial Recognition

To ensure only authorized individuals have access to restricted zones, the system integrates a facial recognition model using FaceNet [3]. FaceNet maps facial features to a high- dimensional space, where similar faces are clustered closer together, enabling effective identity verification [10]. FaceNet maps facial features to a high-dimensional space, where similar faces are clustered closer together, enabling effective identity verification.

The facial recognition module operates in real time, identifying individuals by matching captured facial images against a pre-existing database of authorized personnel. When an individual's face is detected, FaceNet extracts their unique facial embeddings and cross- references them with the database. If a match is found, access is granted; otherwise, an alert is generated. This component is crucial for preventing unauthorized access, especially in high-security areas. The system also employs multiple-angle face capture and normalization techniques to maintain accuracy in varying lighting conditions and camera angles.

3.4 Anomaly Detection

Anomaly detection algorithms add an additional layer of security by identifying unusual or suspicious behavior. In this module, techniques such as Autoencoders and Isolation Forests are employed to detect activities that deviate from the norm [19][7]. Autoencoders are neural networks trained to learn normal patterns within the surveillance footage. When the system encounters an activity that does not match these learned patterns, it flags the event as a potential anomaly.

4. Implementation:

The implementation of the proposed restricted area prevention system involved multiple stages, including data collection, system architecture setup, and front-end interface development. Each component was carefully designed to ensure the system's effectiveness, scalability, and ease of use for security personnel.

4.1 Data Collection

To train the model effectively, data collection focused on capturing video footage under various scenarios that represent real-world conditions in restricted areas [18][1][3]. This included footage of both authorized and unauthorized access attempts, as well as diverse environmental interferences, such as changes in lighting, weather, and movement of non-

threatening objects (e.g., animals, environmental elements like leaves). By simulating a wide range of intrusion and non-intrusion events, the training dataset provides a robust foundation for object detection, facial recognition, and anomaly detection.

To ensure comprehensive coverage, the dataset includes:

- **Authorized Access Events**: Videos of personnel with legitimate access entering the restricted area, allowing the model to recognize authorized behaviors.
- **Unauthorized Entry Attempts**: Simulated footage of individuals attempting to breach the restricted area without authorization, which helps in training the object detection and facial recognition components to distinguish intruders.
- **Environmental Interferences**: Conditions such as nighttime, heavy rain, fog, and variations in camera angles to make the model resilient to changing environmental conditions.
- **Non-Static Backgrounds and Movement Patterns**: Scenarios with both stationary and moving backgrounds, as well as typical movement patterns of individuals within restricted areas.

All collected data was labeled and annotated to include timestamps, object types, access status, and behavior indicators. The data was divided into training and validation sets to assess the model's performance in various scenarios, enhancing detection accuracy and minimizing false positives.

4.2 System Architecture

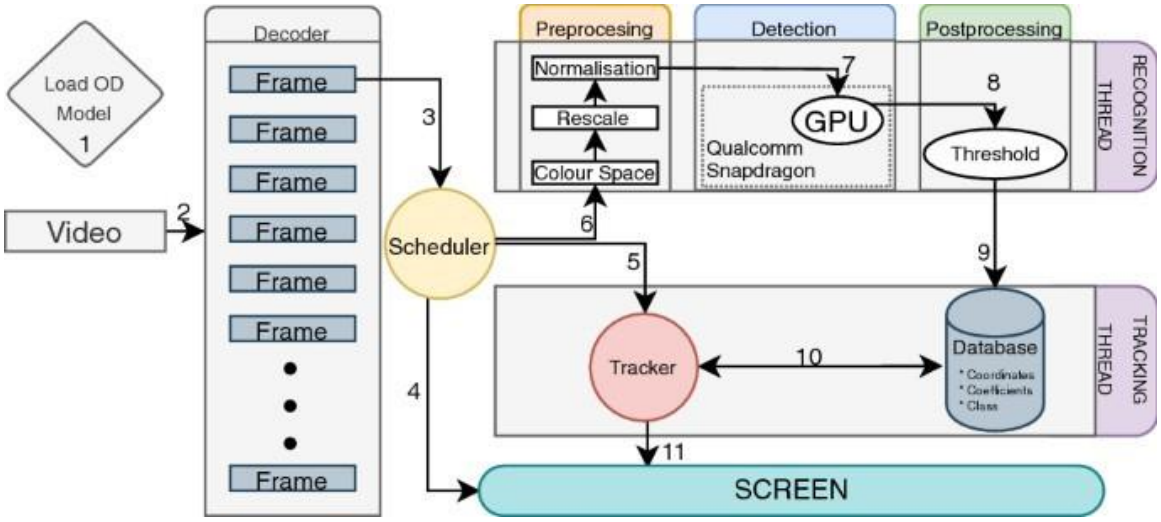
The system's architecture was designed to prioritize fast processing, scalability, and reliability for real-time applications. The core system is hosted on **high-performance servers** equipped with GPUs (Graphics Processing Units), essential for handling intensive video processing tasks required by deep learning models. GPU support is particularly beneficial for accelerating tasks like object detection, facial recognition, and real-time video analysis.

Key components of the system architecture include:

- **Edge Computing**: To minimize latency, edge computing devices are deployed close to the camera sources, handling pre-processing tasks such as frame extraction, initial motion detection, and basic anomaly screening. This approach offloads some

computational tasks from the central server, ensuring faster response times and reducing the bandwidth required for data transmission to the central server.

- **Central Server Processing**: The primary processing of object detection, facial recognition, and advanced anomaly detection occurs on central servers, which aggregate data from multiple cameras and sensors. The server processes incoming video streams in parallel, using containerization (e.g., Docker) to isolate and manage different tasks effectively. The server setup is designed for scalability, allowing additional servers to be added as needed for larger facilities.
- **Cloud Integration and Data Storage**: The system integrates with a cloud storage solution for archiving footage and model checkpoints, enabling historical analysis and further training. Data collected from each restricted area is stored in a secure, encrypted format, with access controlled by user roles, ensuring data privacy and compliance with security policies.
- **Alert System**: A real-time alerting mechanism is integrated into the architecture, allowing the system to notify security personnel immediately upon detecting unauthorized access. Alerts are routed through WebSocket or similar real-time communication protocols, ensuring that notifications are delivered instantly across devices.



4.3 Front-End Interface

The front-end interface was designed with usability and accessibility in mind to ensure that security personnel can operate the system efficiently. The interface is web-based, making it accessible from desktops, tablets, or mobile devices connected to the secure network, enabling remote monitoring and control.

Key features of the front-end interface include:

- **Real-Time Video Feed**: The interface displays live video feeds from multiple cameras in the monitored areas, allowing security personnel to observe ongoing activities. Each feed is updated in real time, and the interface supports switching between feeds, zooming, and adjusting the display layout.
- **Alert Notifications**: When the system detects unauthorized access or suspicious behavior, an alert notification immediately appears on the interface. The alert includes a timestamp, camera location, and the nature of the event (e.g., unauthorized entry, loitering). Security personnel can acknowledge the alert or take immediate action, such as locking doors or notifying on-site security.
- **Footage Review and Playback**: Security personnel can review historical footage through a dedicated playback feature. This feature allows personnel to search for events by date, time, or type, making it easier to conduct investigations and analyze past incidents. The footage review section also includes tools for annotating and flagging specific moments of interest.
- **Dashboard with Analytics**: The interface includes a dashboard displaying real-time analytics on security activities, including intrusion frequency, types of detected behaviors, and high-risk times of day. These insights allow security teams to identify patterns in unauthorized access attempts and adjust their protocols accordingly.
- **User Authentication and Access Control**: To maintain security, the interface requires user authentication, with access levels tailored to each user's role. For instance, administrative users have full access to settings and data management features, while general security personnel may only access live feeds and alert notifications.

This comprehensive implementation approach ensures that the proposed system is not only accurate in detecting unauthorized activities but also efficient, responsive, and user-friendly for security teams responsible for managing restricted areas.

5. Experimental Results:

The proposed surveillance system was tested in a controlled environment replicating a restricted area to assess its performance, with various test cases involving authorized access, unauthorized entry, and environmental interferences.

5.1 Performance Metrics

The system was assessed using the following metrics:

- **Detection Accuracy**: Percentage of correctly detected unauthorized access.
- **Recall**: Proportion of true intrusions detected.
- **False Positive Rate**: Rate at which authorized personnel were flagged as intruders.
- **Response Time**: Time taken to process and trigger an alert after detecting an intrusion.

5.2 Results Summary

METRIC	VALUE
Detection Accuracy	**92%**
Average Response Time	**1.2s**
False Positive Rate	**3%**
Recall	**94%**
Frame Processing Speed	**30 fps**

The system achieved 92% detection accuracy, with 94% recall, indicating high effectiveness in identifying intrusions. The false positive rate was minimized to 3%, and the system processed video at 30 fps with an average response time of 1.2seconds, making it suitable for real-time monitoring.

5.3 Environmental Sensitivity

The system demonstrated robustness against environmental factors such as lighting fluctuations and moving background objects. While facial recognition performance slightly dropped in low-light conditions, object detection and motion analysis remained effective.

5.4 Comparison to Traditional Systems

The proposed system outperformed traditional CCTV systems in terms of reducing false alarms and providing faster, automated alerts. Traditional systems often require manual monitoring, leading to delays in response time.

6. Conclusion:

Our computer vision-based surveillance system significantly enhances security in restricted areas by providing continuous, automated monitoring and real-time alerts for unauthorized access. Through the integration of advanced techniques such as YOLO for object detection, FaceNet for facial recognition, and anomaly detection algorithms, the system is capable of accurately identifying intruders and minimizing false positives. This system offers a key advantage over traditional surveillance methods, which often rely on manual monitoring and simple motion detection algorithms, by providing faster response times and reducing the need for human intervention.

The experimental results demonstrate that the system performs effectively in real-world scenarios, with high detection accuracy (92%), low false positive rate (3%), and rapid response times (1.2 seconds). The system's capability to adapt to environmental factors, including changing lighting conditions and dynamic background elements, further underscores its robustness and practicality for deployment in diverse environments.

However, there remains scope for improvement, with future efforts aimed at increasing the accuracy of the facial recognition model, particularly under low-light conditions, to ensure reliable performance across a wider range of environments. Additionally, expanding the system’s anomaly detection capabilities will allow it to better recognize and react to subtle or complex threats that may not be immediately obvious, thus improving overall security.

Moreover, scalability will be a key focus in future developments, as large-scale deployments may require further optimization for processing speed and resource management, particularly

with high-resolution video feeds. The potential integration of edge computing will help address latency and processing challenges, enabling real-time performance even in large, complex environments.

In conclusion, this system marks a significant advancement in the field of security surveillance, offering a highly effective, automated solution for monitoring restricted areas. Its real-time, AI-powered capabilities have the potential to revolutionize security protocols across various sectors, providing enhanced protection with minimal human oversight.

Reference:

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[2] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. European Conference on Computer Vision (ECCV).

[3] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[4] Farnebäck, G. (2003). Two-Frame Motion Estimation Based on Polynomial Expansion. Scandinavian Conference on Image Analysis.

[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[6] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems (NeurIPS).

[7] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS).

[8] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2019). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[9] Bulat, A., & Tzimiropoulos, G. (2017). How Far Are We from Solving the 2D & 3D Face Alignment Problem? IEEE International Conference on Computer Vision (ICCV).

[10] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep Face Recognition. British Machine Vision Conference (BMVC).

[11] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[12] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[13] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large- Scale Image Recognition. International Conference on Learning Representations (ICLR).

[14] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment Using Multi-task Cascaded Convolutional Networks. IEEE Signal Processing Letters.

[15] Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[16] Sun, Y., Wang, X., & Tang, X. (2014). Deep Learning Face Representation from Predicting 10,000 Classes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[17] Papandreou, G., Zhu, T., Chen, L. C., Gidaris, S., Tompson, J., & Murphy, K. (2018). PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part- Based, Geometric Embedding Model. European Conference on Computer Vision (ECCV).

[18] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[19] Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. International Conference on Learning Representations (ICLR).

[20] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative Localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).