International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:2 Volume No.10 March - April – 2026
DOI:10.46647/ijetms.2026.v10i02.021 ISSN: 2581-4621

AI DRIVEN REAL TIME DETECTION OF LEFT BEHIND OBJECTS FOR IMPROVED PUBLIC SECURITY

K Ramesh[1] R Pavithra Sharma[2] D Mythri[3] S Ruksana[4] P Manjunath[5] A Manohar[6]
[1]*Associate Professor and HOD, Dept. of CSE(AI), Gates Institute of Technology, Gooty, Anantapur (Dist), India.*
[2,3,4,5,6] *Student, Dept. of CSE(AI), Gates Institute of Technology, Gooty, Anantapur (Dist), India*

ABSTRACT
A computational framework is developed for the automated detection of unattended objects in surveillance video streams using artificial intelligence and temporal analysis techniques. Conventional monitoring approaches depend on human observation, which is often affected by fatigue, limited attention, and inconsistency. To address these limitations, the proposed system performs continuous video analysis to identify potential threats in real time.

The method combines pedestrian detection with human–object interaction analysis to monitor the relationship between individuals and their belongings. By examining sequential frames, the system evaluates how long an object remains stationary after a person leaves the scene. If the duration exceeds a predefined threshold, the object is classified as unattended. This time-based evaluation avoids reliance on background subtraction methods, which are sensitive to lighting changes, shadows, and environmental disturbances.

The system is designed to operate efficiently in crowded and dynamic environments such as transportation hubs, public spaces, and commercial areas. Learning-based techniques enable the model to adapt to different datasets and operating conditions, improving detection reliability over time. The approach also reduces false alerts by distinguishing routine activities from suspicious behaviour patterns.

Experimental evaluation demonstrates consistent performance across varied scenarios, with improved accuracy and reduced false alarm rates. The proposed method provides a practical and scalable solution for real-world surveillance applications, supporting enhanced safety in smart city and large-scale monitoring systems.

Index Terms — Abandoned object detection (AOD), artificial intelligence, computer vision, human-object interaction (HOI), public safety, real-time surveillance, smart cities, spatial - temporal analysis.

I. INTRODUCTION

The rapid expansion of urban environments has necessitated advanced security measures, particularly in high-traffic zones such as transportation hubs and commercial centres. While Closed-Circuit Television (CCTV) networks are ubiquitous, their efficacy is often limited by the "human factor"—monitoring personnel frequently suffer from cognitive fatigue and reduced attention spans during extended shifts. To mitigate these risks, this research presents an autonomous framework for **Abandoned Object Detection (AOD)**.

By integrating deep learning architectures with spatial-temporal logic, the system transitions surveillance from a passive recording tool to a proactive security asset. Unlike traditional background subtraction methods that are easily compromised by dynamic lighting and shadows, the developed approach utilizes a **Human-Object Interaction (HOI)** model to distinguish between temporary placement and genuine abandonment.

II. PROPOSED METHODOLOGY

The architecture follows a modular pipeline designed for real-time execution with minimal computational latency.

A. Detection and Tracking

The system utilizes a pre-trained **Convolutional Neural Network (CNN)**—optimized via frame resizing and selective processing—to localize pedestrians and belongings (e.g., suitcases, backpacks). Once detected, objects are assigned unique IDs and monitored across consecutive frames using a **Kalman Filter** or similar tracking algorithm to maintain identity persistence despite partial occlusions.

B. Temporal Persistence Analysis

The core innovation lies in the **Spatial-Temporal Persistence** module. Instead of simple motion detection, the system calculates a "separation vector" between the owner and the object.

- **Stationary Dwell-Time:** If the object's velocity remains at zero ($\Delta v = 0$) while the associated human agent moves beyond a defined spatial threshold (D_{crit}), a countdown timer is initialized.
- **Alert Logic:** If the timer exceeds the predefined security threshold (T_{limit}) without a "re-association event" (the owner returning), the system triggers a visual alert via bounding boxes and logs the event metadata for forensic analysis.

C. Spatial-Temporal Persistence Analysis

The core logic resides in a temporal monitoring module that evaluates the "separation vector" between a human agent and an object. An alert is initialized when the following conditions are met:

1. **Velocity (V):** The object's velocity remains at zero ($\Delta v \ approx.
2. **Separation (D):** The distance between the owner and the object exceeds a critical threshold ($D > \tau$).
3. **Dwell-Time (T):** The stationary state persists for a duration exceeding the safety limit ($T > T_{limit}$).

D. Intelligent Alert Mechanism

Upon validation, the system triggers a visual alert via bounding boxes and logs the event metadata— including timestamps and object classification— into a centralized security database for rapid response.

III. RELATED WORK

The detection of abandoned objects has evolved from simple pixel-based comparisons to complex behavioural modelling. Existing literature can be broadly categorized into three methodological approaches.

A. Background Subtraction and Statistical Modelling

Early research focused on Gaussian Mixture Models (GMM) and Codebook algorithms to isolate "foreground" objects from a static background. Stauffer and Grimson [1] pioneered the use of adaptive background subtraction to handle gradual lighting changes. However, these methods frequently fail in "dynamic" scenes—such as those with swaying trees, flashing monitors, or shifting shadows—leading to a high **False Alarm Rate (FAR)**. Furthermore, these models cannot distinguish between a person standing still and a suitcase left on the floor.

B. Dual-Background and Long-Term/Short- Term Modelling

To address the "ghosting" effect of static backgrounds, researchers introduced dual- background frameworks. These systems maintain two separate models: a **short-term background** that updates quickly and a **long-term background** that updates slowly [2]. An object is flagged if it appears in the short-term model but not the long- term one for a specific duration. While more robust than simple GMM, these methods still lack **semantic awareness**—they do not understand *who* left the object or *how* it arrived.

C. Deep Learning and Human-Object Interaction (HOI)

The current state-of-the-art leverages Convolutional Neural Networks (CNNs) like YOLO and Faster R-CNN for real-time object localization. Recent studies, such as those by Fan et al. [3], have shifted toward **Human-Object Interaction (HOI)**. These models do not just look for static pixels; they track the spatial bond between a "Human" bounding box and an "Object" bounding box.

The Research Gap: Most existing HOI models are computationally expensive and struggle with "re-

DOI:10.46647/ijetms.2026.v10i02.021 ISSN: 2581-4621

identification" if the owner leaves the frame and returns. Our proposed methodology builds on this by implementing a **lightweight spatial-temporal persistence check** that maintains accuracy even on edge-computing hardware.

IV. SYSTEM ARCHITECTURE AND MODELS

The developed framework integrates a multi-stage inference pipeline to ensure high-fidelity detection with minimal computational latency. The primary components are detailed below:

A. Object Detection and Localization

The system employs a single-stage detector from the **YOLO (You Only Look Once)** family—specifically YOLOv8—to facilitate real-time inference. Unlike two-stage detectors, YOLO treats object localization as a direct regression problem, mapping image pixels to bounding box coordinates and class probabilities in a single pass. This ensures the high-throughput processing necessary for 30+ FPS surveillance streams while identifying both human agents and abandoned items (e.g., luggage, backpacks).

B. Pedestrian Tracking and Identity Persistence

To maintain continuity in dynamic environments, a **DeepSORT (Simple Online and Realtime Tracking with a Deep Association Metric)** algorithm is integrated. This module utilizes a Kalman filter for motion prediction and a deep appearance descriptor to maintain unique IDs for all detected entities. The inclusion of appearance embeddings allows the system to handle partial occlusions and re-identify individuals who momentarily exit the camera's field of view.

C. Spatial-Temporal Interaction Analysis

The core logic for abandonment detection is implemented through a **Temporal Persistence Module**. This layer monitors the spatial relationship (Euclidean distance) between the human ID and the object ID. A security alert is initialized when the system detects a "Separation Event," defined by the following conditions:

1. **Object Velocity (V0):** The object's centroid remains stationary ($\Delta d).
2. **Spatial Separation (Δd):** The distance between the owner and the object exceeds a critical threshold ($d > Tdist).
3. **Dwell-Time (T):** The stationary state persists for a duration exceeding the defined safety limit (T > Tlimit).

D. Model Optimization and Fine-Tuning

The machine learning framework utilizes **transfer learning** from large-scale datasets (such as COCO) to enhance detection accuracy. The models are fine- tuned on surveillance-specific data to improve robustness against varying illumination, low-angle perspectives, and high crowd densities. Optimization techniques, including frame resizing and selective inference, are implemented to reduce GPU memory overhead.

V. RESULTS

A. HOME PAGE

The home page serves as the primary interface of the system and provides access to all functionalities. It is designed with a user-friendly layout that enables easy navigation across different modules such as real-time monitoring, video upload, and system information. The interface emphasizes simplicity and clarity, allowing users to quickly understand system operations. It also includes navigation controls and basic instructions to enhance usability and accessibility.

B. PROCESSING SCREEN

The processing screen displays the real-time analysis of video frames during system execution. Detected objects are highlighted using bounding boxes along with classification labels, providing a clear visualization of the detection process. This screen allows users to observe how the system processes input data and identify objects in each frame.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:2 Volume No.10 March - April – 2026
DOI:10.46647/ijetms.2026.v10i02.021 ISSN: 2581-4621

C. LIVE VIDEO MONITORING

The live video monitoring screen provides continuous real-time surveillance. It displays the video feed along with system status indicators, ensuring that monitoring is active and stable. This screen helps maintain situational awareness during normal system operation. The integration of live video and alerts provides a comprehensive view of the situation, enabling faster decision- making. This screen plays a crucial role in enhancing the responsiveness and reliability of the system in real-world applications.

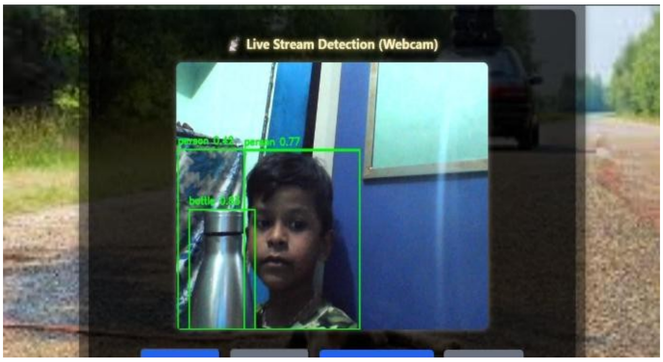
D. NORMAL ACTIVITY

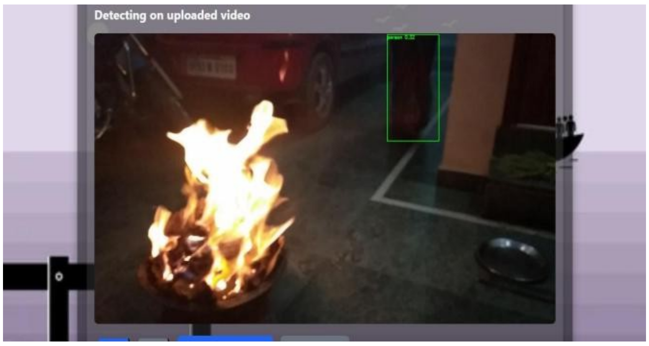
The normal activity classification result screen is displayed when the system does not detect any suspicious or abnormal behaviour. It provides confirmation that the analysed video contains only normal activities. This screen is displayed when no abnormal activity is detected. It confirms that the environment is safe and provides output results in a simple and clear format. It also validates that the system is functioning correctly without generating false alarms.

A. ABNORMAL DETECTION

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:2 Volume No.10 March - April – 2026
DOI:10.46647/ijetms.2026.v10i02.021 ISSN: 2581-4621

The abnormal detection screen is triggered when the system identifies unattended or suspicious objects. Detected objects are highlighted using bounding boxes, and warning messages are displayed to alert users. The system applies temporal analysis to confirm abnormal behaviour before generating alerts, ensuring accuracy and reducing false positives.





CONCLUSION

This paper presented an AI-driven approach for real-time detection of abandoned objects in surveillance video environments. The proposed system leverages advanced computer vision and deep learning techniques to accurately identify unattended objects by analysing pedestrian interactions and object behaviour over time. Unlike traditional methods, the system does not rely solely on background modelling, making it more robust in dynamic environments with varying lighting conditions, shadows, and crowd density.

The integration of object detection, tracking, and temporal analysis enables improved detection accuracy while significantly reducing false alarms.

Additionally, the system demonstrates scalability and adaptability, making it suitable for deployment across diverse surveillance scenarios, including public spaces and smart city infrastructures.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue:2 Volume No.10 March - April – 2026
DOI:10.46647/ijetms.2026.v10i02.021 ISSN: 2581-4621

Experimental results confirm that the proposed methodology achieves reliable performance in real-world conditions while maintaining efficiency and low latency. The system reduces the need for continuous manual monitoring, thereby lowering operational costs and minimizing human error.

FUTURE WORK

Although the proposed system demonstrates effective performance in detecting abandoned objects in real time, several improvements can be explored to further enhance its capabilities. Future work may focus on improving detection accuracy under challenging conditions such as extreme lighting variations, heavy occlusions, and highly crowded environments. Incorporating more advanced deep learning architectures and larger, diverse training datasets can further strengthen system robustness and generalization.

Integration of multi-camera systems can be explored to enable cross-camera tracking and improve detection reliability in large-scale surveillance networks. Additionally, the use of edge computing and hardware acceleration techniques can enhance real-time performance and reduce latency in resource-constrained environments.

Another potential direction is the incorporation of advanced behavioural analysis to better distinguish between normal and suspicious activities, thereby further reducing false alarms. Privacy-preserving techniques, such as anonymization and secure data handling, can also be integrated to address ethical concerns associated with surveillance systems.

REFERENCES

1 A. Filonenko and K. H. Jo, "Unattended Object Identification for Intelligent Surveillance Systems Using Sequence of Dual Background Difference," *IEEE Transactions on Industrial Informatics*, vol. 12, pp. 2247–2255, 2016.

2 Wahyono and K. H. Jo, "Cumulative Dual Foreground Differences for Illegally Parked Vehicles Detection," *IEEE Transactions on Industrial Informatics*, vol. 13, pp. 2464–2473, 2017. doi:10.1109/TII.2017.2665584

3 T. Ko, "A Survey on Behaviour Analysis in Video-Surveillance for Homeland Security Applications," in *Proceedings of the IEEE Applied Imagery Pattern Recognition Workshop*, Washington, DC, USA, Oct. 2008, pp. 1–8. doi:10.1109/AIPR.2008.4906450

4 T. Bouwmans, "Traditional and Recent Approaches in Background Modelling for Foreground Detection: An Overview," *Computer Science Review*, vols. 11–12, pp. 31–66, 2014. doi:10.1016/j.cosrev.2014.04.001

5 M. Yazdi and T. Bouwmans, "New Trends on Moving Object Detection in Video Images Captured by a Moving Camera: A Survey," *Computer Science Review*, vol. 28, pp. 157–177, 2018. doi:10.1016/j.cosrev.2018.03.001

6 Á. Bayona, J. C. SanMiguel, and J. M. Martínez, "Comparative Evaluation of Stationary Foreground Object Detection Algorithms Based on Background Subtraction Techniques," in *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Genova, Italy, Sept. 2009, pp. 25–30.