



# Intelligent Room Surveillance: AI-Based Object Tracking and Missing Item Detection

Meena Kumari Bugatha. N\*<sup>1</sup>, Nasrrin Kalifathulla Khan<sup>1</sup>

<sup>1</sup>Department of Artificial Intelligence and Machine Learning Engineering, St.Joseph's College of Engineering, OMR, Chennai, India.

meenakumaribugathan@stjosephs.ac.in

**Abstract.** In an age of artificial intelligence and automation, intelligent surveillance systems have become essential in asset management and safety. The system proposed here, a smart surveillance system, utilizes AI-enabled computer vision algorithms that can monitor objects and recognize the absence of an object or objects, or determine it (or them) were removed or misplaced in real time. The system continuously monitors the environment of a room, recognizes objects, tracks their movement, and notifies alert users if an object is displaced or removed. The system uses deep learning models of detection, classification, and tracking of objects of interest with advanced algorithms, including YOLO and DeepSORT, to accurately track and analyze objects' movement. When an object is detected to be unbound, the system identifies that object's last known location and alerts users either via an intuitive dashboard or throughout a push notification. The system can integrate with IoT devices, including but not limited to smart cameras, smart sensors, and forms of smart technology to make the system accurate and efficient. The purpose of the system was to be able to monitor for security across various settings such as home, office space, warehouse, or any other surveillance setting. The proposed system can be implemented in a scalable and cost-effective way to provide enhanced monitoring capabilities and to reduce the amount of manual monitoring in the same day-to-day observational settings as manual observations. Using AI, IoT, and real-time analytics, the proposed system develops a proactive and predictive outreach of modern surveillance to improve security, reduce losses from assets and improve monitoring protocol efficiency.

**Index Terms:** AI-Based Surveillance, Object Tracking, Missing Item Detection, Computer Vision

## 1. Introduction

The fast improvement in Artificial Intelligence (AI) and imaging technologies has transformed many sectors, providing new solutions for real-time monitoring, object detection, and automated surveillance. AI and imaging processing techniques establish an effective infrastructure to better understand and analyze visual data. The ability to recognize, categorize, and track objects has large implications in the area of security, healthcare, autonomous cars, and smart surveillance. The current humanities face new security challenges, and AI equipped intelligent surveillance systems provide a critical tool to manage safety and assets. A Role of AI in Imaging and Object Detection : AI has transformed the discipline of computer vision and imaging processing by allowing machines

to interpret and analyze visual information precisely. One of the more significant advancement has been in object recognition, where AI algorithms, specifically deep learning models like Convolutional Neural Networks (CNNs), have enabled automated systems to detect and classify objects in images and in video streams. In particular, the potential applications for security surveillance present significant efficiencies in real time tracking and lost item detection. Traditional methods of object detection relied on manual feature extraction, which often lacked robustness in dynamic environments. However, AI-powered solutions leverage self-learning algorithms, allowing systems to adapt and improve over time. Deep learning frameworks such as YOLO (You Only Look Once) and DeepSORT (Simple Online and Realtime Tracking) have emerged as highly effective tools for real-time object tracking, ensuring improved accuracy and efficiency.

Intelligent surveillance systems return AI-based object tracking and anomaly detection as a means of bolstering secure surveillance safety. These intelligent surveillance systems rely on AI-based computer vision technology for interpreting footage to observe and learn about patterns and ongoing comments of suspicious activity. AI-based computer vision technology assist in reducing human resourceful factors by reengineered the human dependency process object tracking process when, for an example, a child is missing or misplaced object, consequently reducing and human error factors. AI surveillance can be used in a variety of contexts and systems including, but not limited to, the following • Home Security Systems: Cameras/sensors supported by AI can scan home environments and/or analyze conditions while driving, providing the homeowner possible alerts if movement is detected or something is disturbed. • Work/Office Surveillance: Businesses could engage AI enabled Security Monitoring and Surveillance, which may include the monitoring of behavior in the workplace, as well as, verification of safety protocols for individuals to effect business loss by means of theft of the company will monitor the activity.]Applications of AI in Surveillance and Security Traditional methods of object detection rely upon structured manual feature extraction, which could be less reliable in a dynamic environment. In contrast, AI-driven solutions illustrate a distinct self-learning algorithm that enables a system to build up and adapt over time. Emerging deep-learning frameworks, such as YOLO (You Only Look Once) and DeepSORT (Simple Online and Realtime Tracking), for example, are regarded as the top options for real-time object tracking in a more efficient and accurate environment. Intelligent surveillance systems return AI-based object tracking and anomaly detection as a means of bolstering secure surveillance safety. These intelligent surveillance systems rely on AI-based computer vision technology for interpreting footage to observe and learn about patterns and ongoing comments of suspicious activity. AI-based computer vision technology assist in reducing human resourceful factors by reengineered the human dependency process object tracking process when, for an example, a child is

missing or misplaced object, consequently reducing and human error factors. AI surveillance can be used in a variety of contexts and systems including, but not limited to, the following • Home Security Systems: Cameras/sensors supported by AI can scan home environments and/or analyze conditions while driving, providing the homeowner possible alerts if movement is detected or something is disturbed. • Work/Office Surveillance: Businesses could engage AI enabled Security Monitoring and Surveillance, which may include the monitoring of behavior in the workplace, as well as, verification of safety protocols for individuals to effect business loss by means of theft of the company will monitor the activity.

## 2. Related Works

Liu et al. [1] (2024) studied strategies for using multidomain stereo identification and manipulation strategies for engagement in counterterrorism. The authors integrated stereo vision with intelligent decision-making protocols to improve accurate threat detection for multiple environments. The study illustrated considerable adaptability to different environments, but high processing requirements limited the real-time applicability of the algorithms for mobile surveillance capabilities. DeLucia et al. [2] (2025) analyzed research on LaserInduced Breakdown Spectroscopy (LIBS) as a chemistry/sensory technology for hazardous materials detection. Their chemistry approach showed promise for rapid, remoterange identification of dangerous substances based on analysis of plasma emissions. The approach is highly versatile and accurate, though less effective under variable environmental conditions (e.g., changing humidity and dust). Nelson et al. [3] (2024) devised a metal detection system intended for increased security in stadium and event settings. The detection system incorporated electromagnetic sensors and real-time processing to categorize items as safe or unsafe. The study reported improved scanning efficacy but also presented challenges to resolving false alarms when scanning in crowded environments. Schieber et al. [4] (2024) examined the progress of polycrystalline mercuric iodide (HgI) detectors in applications for X-ray imaging. The focus of their study was on improving charge collection efficiency as well as detector uniformity for improved image quality. While the study identified image quality as improvements over current X-ray energy-absorbing materials, the study conceded that attrition of the costs for these detectors as well as the complicated crystal growth process would diminish its large-scale adoption. Mery et al. [5] (2025) performed a broad survey of computer vision-based X-ray baggage inspection systems. The authors examined deep learning models, such as convolutional neural networks (CNNs), for detecting contraband objects in cluttered images. The authors found that AI-based approaches out performed classical inspection approaches, but problems of model explainability and interpretability remained as problematic for high-stake surveillance approaches. Kong and Dai [6] (2024) reported a technique for

object-detection-based image compression specifically purposed for wireless surveillances. The authors proposed only encoding video frames that contained objects of interest, which optimized bandwidth for streaming while preserving image quality. While the deliverables improved transmission efficiency, the need for accurate and timely object detection diminished adequate encoding of video frames, especially when scenes were dynamic. Sheen et al. [7] (2024) created a three-dimensional millimeter-wave imaging system aimed at detecting concealed weapons. The groundwork of their method was based on synthetic aperture radar (SAR) methods for reconstructing three-dimensional object profiles hidden under clothing. While the system could provide reliable detection methods with current limitations of metallic reflection interference and expensive hardware component availability. Chen et al. [8] (2025) presented an intelligent caching strategy for vehicular named data networks incorporated timespace localization features to reduce latency in data access. Their implementation was effective in reducing latency and providing more actionable transmissions of surveillance data in smart city applications, particularly in reducing congestion by reducing the network strains of larger use cases. Some barriers to an expanded application were a lack of performance at scale in larger vehicular coverage areas and variability in vehicular coverage. Wan et al. [9] (2025) presented a framework for AI-based video analysis algorithms designed for detecting abnormal events in intelligent transportation systems. Their approach was based on a motion trajectory and scene understanding algorithms, producing high levels of detection performance for accidents and abnormal behavior. However, their paper limited the application of their algorithms in real-time detection applications, due to the computational load and feature extraction requirement. Wilson [10] (2025) considered the role of modeling and simulation for research to support airport security operations. His work showed how virtual environments can be used to test security checkpoint layouts, personnel deployments, and more to impact operations. Although simulations could improve the efficiency of fixed checkpoints often encountered at airports, the gaps in the range of real-world unpredictability for a live security scenario limited their complete applicability. Rappaport et al. [11] (2023) examined the practicability of mobile millimeter-wave communications for the 5G network. They concluded that the mmWave has potential in providing ultra-fast connectivity, allowing security and surveillance systems of the future to operate in real-time. However, the limitation of signal attenuation in non-line-of-sight conditions remained a barrier. Meng et al. [12] (2024) suggested millimeter-wave holographic imaging to provide advanced security checks. Their method reconstructed three-dimensional holographic images to visualize hidden objects in a non-invasive manner. While achieving high precision, the system was dependent on extensive calibration and control of the environment to maintain quality of imaging. Myint et al. [13] (2024) advanced concealed object recognition by

utilizing active millimeter-wave imaging with deep learning techniques. The system effectively determined concealed objects from body contours, improving reliability of detection when distinguishing between layers. The major limitation to portability was high energy requirements and demand for data processing. Zhang et al. [14] (2024) introduced StackGAN, a stacked generative adversarial network capable of transforming text-to-image synthesis by acceptance of written descriptions to create photorealistic images. The authors demonstrated that the multistage image synthesis improved synthesis quality but required vast amounts of training data on a large scale. As an idea, the technology could be valuable for improving AI construction of images based on security monitoring. Katsuyama et al. [15] (2025) developed an AI-based suspicious object detection system operating in the W-band using GANs and reported their system outperformed existing technology in detecting suspicious objects and significantly enhanced the ability to detect suspicious objects when it was in motion, and assisted in standardizing AI-based imaging for security purposes. Note that processing time was hindered for real time purposes. Yu et al. [16] (2020) developed a distributed storage framework for surveillance cameras preserving privacy while managing stored data. Their solution utilized encrypted decentralized storage creating a secure method of to ensure data integrity. However, although secure, a performance bottleneck developed in synchronizing the nodes of the distributed model. Ding et al. [17] (2024) proposed an anti-forensics model for detection of face-swapped videos based on adversarial training. This method can further enhance resilience for multimedia security against tampering using deepfake techniques. The model operated well on controlled datasets, however, its performance significantly dropped in low-resolution formats when compressed videos were mixed into the dataset. Xiao et al. [18] (2025) researched millimeter-communication technologies for future mobile communications noting it will be suitable for security and autonomous systems to high transfer rates it allows. Yujiri et al. [19] (2023) provided an extensive overview of passive millimeter-wave imaging technology aimed at noninvasive threat detection. They displayed excellent potential in the detection of concealed weapons at a distance without contact; however, a number of challenges remained with regard to image noise and environmental factors still affecting sensitivities. Goshi et al. [20] (2025) successfully employed an active W-band millimeter-wave sensor for cable imaging and structural inspection. The sensor achieved high spatial resolution imaging of embedded materials, however further adjustments will be needed to enable human threat detection [21].

### 3. Methodology

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

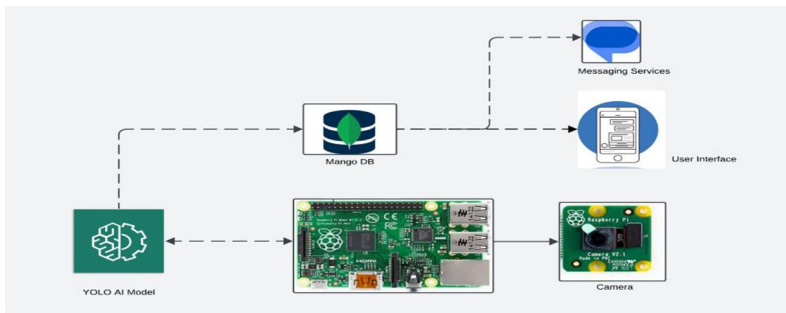
The methodology of the project, "Intelligent Room Surveillance: AI-Based Object Tracking and Missing Item Detection (1)," is set up to be efficient and effective in

monitoring objects in a defined space. The system operates step by step, where each section lays out the function of the surveillance system. Here is a detailed description of each block in the methodology.

$$S_m = \sum_i (1 - d_i) M \quad (2)$$

The system starts in "Initial System State," to be inactive and not to consume many computer resources (2). This state is highly important as it keeps the system functional and efficient as it only operates when necessary. The system constantly watches for triggering conditions to go live, such as object presence or changes to the environment. When there are no triggering conditions, the system stays idle to not execute unnecessary tasks by only running while it has triggering conditions. The initial system state also ensures that no unnecessary data is obtained by the system, therefore decreasing data clutter within the database and allowing the system to focus on getting relevant input conditional to triggering object presence or changes in the environment.

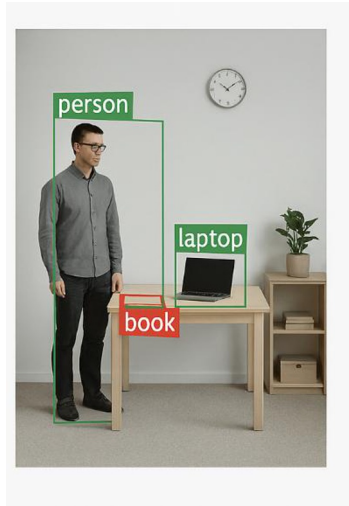
Specify a Monitoring Threshold At this stage, the system will define a monitoring threshold. This threshold will dictate when to transition from the initial state to a monitoring state. The monitoring threshold is adjustable based on user requirements in the observation environment. For instance, the threshold for a smaller room or area may be smaller than that of a larger area or room. The protocol of the system is to continually compare the current number of items/people in the room with this threshold to determine whether further actions should be taken. Activate Monitoring Mode Once it is determined that the threshold has been passed, the monitoring system active monitoring mode. It will active monitor a specified sub-region of the room, known as a Region of Interest (ROI) watching and tracking objects, and looking for any anomalies. During this active monitoring state, the system will update the visual aspect of the ROIs to indicate that the ROIs is enter the active monitoring state. Note that this changed visual will help those who are observing the state of the monitoring system or any other live use case to better understand that the system is receiving and monitoring data in an active state.]Evaluate Room Activity The system Fig 1 constantly verifies that the amount of objects or people in the room does not exceed the predetermined monitoring threshold, and compares both the object count and the count against this threshold in order to determine when to transition the monitoring system from the initial state to an active monitoring state. As long as the amount of objects oder people is below the threshold, the monitoring system will remain in its initial state Fig. 1. However, Green or Red indicates that the amount of objects or people exceeds the threshold, and the monitoring system will move on to determine the next step which includes passive monitoring and tracking the objects.



**Fig 1.** Architecture Diagram

**Activate Monitoring Mode** Once it is determined that the threshold has been passed, the monitoring system active monitoring mode. It will active monitor a specified sub-region of the room, known as a Region of Interest (ROI) watching and tracking objects, and looking for any anomalies. During this active monitoring state, the system will update the visual aspect of the ROIs to indicate that the ROIs is enter the active monitoring state. Note that this changed visual will help those who are observing the state of the monitoring system or any other live use case to better understand that the system is receiving and monitoring data in an active state. **Recognizing Objects and Tracking in Real Time** In this phase of the application, the system will recognize and begin tracking certain objects within the ROIs. The second phase will also recognize and categorize the objects based on this room using the YOLO AI model. Each object which crosses the ROI gets assigned a unique ID, which lets the overall system keep track of the object's track of travel over a certain period of time. The objects in that space will remain constantly updated in the system so that the system has the best information in as close to real time as possible over the objects in that room were located initially. This is a very important operation to keep continual track of the object in question both as a way to notify that something is missing from the original observation, as well as to track anything that is moving in this space that should not be.

**Detecting Anomalies** Once the system is continually tracking the objects in the ROIs, the system will check if there are any anomalies, such as missing items or unauthorized movement of tracked items. The system determines whether any anomalies have taken place by comparing the current state of the room to the reference state of the room. If the object which is being tracked is not in the location it was supposed to be, the system will flag it as missing. This operation is very important to efficiently flag anomalies in time Fig. 2. This operation is very important to our efficiency relative to the overall goal of prevention of theft, or items lost.



**Fig. 2.** Intelligent room surveillance

Indicate Anomaly Location When an anomaly is detected, the system can indicate the ROI where the anomaly was detected. This prompts the user the location of the anomaly. This is helpful in evaluating the scene further. The system can change the colour of the ROI, or border the ROI, allowing the user to readily identify the location of the anomaly. Record Data and Send to the Server When an anomaly is detected, the system captures video frames of the anomaly and transmits video frames to the server for observation. The frames indicate the location of the anomaly and it may contain other objects, or persons, that are important to the context of the observed anomaly. The frames are recorded in a database for later observation. This is useful for record keeping, investigation, or auditing purposes. Central Data Processing The server serves a vital purpose in the architecture of the system, acting as the main core for data storage and processing.

**Table 1.** Object Tracking Status

<b>Object</b>	<b>Status</b>
Person	Tracked
Laptop	Tracked
Book	Missing

After the anomaly captures frames of a captured event, the frames of the event are uploaded to the server for storage. The server is responsible for other functions as well, such as sending out notifications to users and storing details in a database Table 1. These functions are all centralized to assist with easier system functionality and data access. Store Event Data The server stores the captured frames along with any related metadata in a database, thus capturing the anomaly event in a permanent record. The frames can be stored with metadata, such as timestamps and object ID, which will allow users to easily search and retrieve events. The database is also designed to store large numbers of events, accommodating a large increase in users and objects as necessary. Notify Users Once the anomaly event is recorded in the database, the system will send a notification to the user. The notification can come via an SMS, email, or push notification, depending on how the user would like to receive notifications. The notification will include information related to the anomaly, including the location and time.

#### **4. Conclusion**

Monitoring Cycle The process is completed after the notification is sent and the anomaly event is recorded in the system to return to its original state monitoring the environment for new triggering conditions. This process demonstrates a systematic methodology for the intelligent room surveillance system "Intelligent Room Surveillance: AIBased Object Tracking and Missing Item Detection." By implementing these procedures, the system will engage in a method that promotes efficient and effective monitoring of the object(s) in the intended area. The paradigm demonstrated how advanced AI model types, like YOLO, can work in conjunction with data storage and notifications and eliminate time delays when they system detects and responds to detect the anomaly. This aspect is advantageous to not just extending security, additionally, offers an easy to use and reliable room to the manufacturer or user.

## References

1. Liu, Y., Wang, Z., and Zhang, G. The Use of Multi-domain Stereo Technology in Terrorism Combating Applications. *Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 230-230. Shenzhen, China. (2019).
2. Gupta, S., Kartha, R.S., Ramirez-Asis, E., Flores-Albornoz, J., Vilchez-Vasquez, R. and Singh, S.K. A Novel Uncalibrated Camera Based Traffic Surveillance System. *Journal of Advanced Transportation*, 2022 ID 5867524. (2022).
3. DeLucia, F., Samuels, A., Harmon, R., Format, L., Gunter, L. and Kline, L. Laser-induced Breakdown Spectroscopy (Libs): A Viable Versatile Chemical Sensor Technology for Hazardous Material Detection Utilizing Hazardous Material Detection Using Laser Induced Breakdown Spectroscopy. *IEEE Sensors Journal*, 5(4), 681-689. (2005).
4. Nelson, V. Chaudhary, J. Edman, P. Kantor. (2016). Walkthrough metal detectors for stadium security. *IEEE Symposium on Technologies for Homeland Security (HST)*, 1-6. Waltham, MA, USA. 2016
5. M. Schieber, A. Zuck, H. Gilboa, G. Zentai. Review of Polycrystalline Mercuric Iodide X-ray Detectors. *IEEE Transactions of Nuclear Science*, 53(4): 2385-2391. (2006).
6. Mery, D. Saavedra, M. Prasad. Computer Vision Based X-ray Baggage Inspection - Overview. *IEEE Access*, 8: 145620-145633. (2020).
7. H. Yin. Public Security Video Image Detection Construction Platform in Cloud Computation Environment. *Computational Intelligence and Neuroscience*, 13p. (2022).
8. L. Kong, R. Dai. Object Detection Based Video Compression Method for Wireless Surveillance Systems. *IEEE Multimedia*, 24(2): 76-85. (2017).
9. D. Sheen, D. McMakin, and T. Hall. Three-dimensional Millimeter-Wave Imaging for Concealed Weapon Detection. *IEEE Transactions on Microwave Theory and Techniques*, 49(9), 1581–1592. (2001).
10. Zhao, S. Li, B. Ren, Q. Qiu, and H. Sun. Cylindrical Three-Dimensional Millimeter-Wave Imaging via Compressive Sensing. *International Journal of Antennas and Propagation*, Article ID 218751, 6 pages. (2015).
11. T. Hong, J. Yao, C. Liu, and F. Qi. Mmwave Measurement of RF Reflectors for 5G Green Communications. *Wireless Communications and Mobile Computing*, Article ID 8217839, 10 pages. (2018).
12. C. Chen, J. Jiang, R. Fu, L. Chen, C. Li, and S. Wan. An Intelligent Caching Strategy Considering Time-Space Characteristics in Vehicular Named Data Networks. *IEEE Transactions on Intelligent Transportation Systems*, 1–13. (2021).

13. S. Wan, X. Xu, T. Wang, and Z. Gu. An Intelligent Video Analysis Method for Abnormal Event Detection in Intelligent Transportation Systems. *IEEE Transactions on Intelligent Transportation Systems*, 22(7), 4487–4495. (2021).
14. D. Wilson. Modeling and Simulation to Support Airport Security. *IEEE Aerospace and Electronic Systems Magazine*, 20(8), 3–6. (2005).
15. T. S. Rappaport, S. Sun, R. Mayzus et al. Mmwave Holographic Imaging for Security Check. *2019 IEEE Asia-Pacific Microwave Conference (APMC)*, 569–571. Singapore. (2019).
16. Vanitha, V., Joe, S.B., Krishnan, R., Fletcher, A.S.A., Anju, M., Akila, V.: Cognitive Threats Detection Model using Nature Inspired Chimpanzee Optimization for IoT Networks (CCM-COM). In: Atlantis highlights in engineering/Atlantis Highlights in Engineering. pp. 629–637 (2025). [https://doi.org/10.2991/978-94-6463-754-0\\_55](https://doi.org/10.2991/978-94-6463-754-0_55).
17. R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. Columbus, Ohio. (2014).
18. R. Girshick. Fast R-CNN. *2015 IEEE International Conference on Computer Vision (ICCV)*, 1440–1448. Santiago, Chile. (2015).
19. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779788. Las Vegas, Nevada.
20. Z. Cheng and F. Zhang. (2020). Flower end-to-end detection based on YOLOv4 using mobile device. *Wireless Communications and Mobile Computing*, Vol 2020, Article ID 8870649, 9 pages.
21. J.F., Y.S. and X.J. (2021). High Resolution ISAR Imaging Based on An Improved Sparse Signal Recovery Algorithm. *Wireless Communications and Mobile Computing*, Vol 2020, Article ID 5541116, 7 pages.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

