

Hybrid Indoor Security System based on Millimetre Wave Radar, RFID, and Face Recognition

Mohammed Safy¹, Ahmad M. Nagm^{2,3}, Ahmed Abdelhafeez^{2,4}, Moshira A. Ebrahim^{5,*},
Mohamed Refaat Abdellah⁶

¹ College of Computing and Information Technology, Arab Academy for Science, Technology and Maritime Transport (AASTMT), Smart Village, B 2401, Giza, Egypt; msafy@aast.edu

² Faculty of Computers and Information Technology, Innovation University, Egypt

³ Department of Computer Engineering and Electronics, Cairo Higher Institute for Engineering, Computer Science and Management, Cairo 11845, Egypt; ahmadnagm@alazhar.edu.eg

⁴ Applied Science Research Center, Applied Science Private University, Amman, Jordan

⁵ Computer Engineering and Information Technology Department, Modern Academy for Engineering and Technology, Cairo, Egypt; mushira.ibrahim@eng.modern-academy.edu.eg

⁶ The Department of Computer Science, College of Information Technology, Misr University for Science and Technology, Cairo, Egypt; mohamed.refaat@must.edu.eg

Abstract Indoor security systems must guarantee availability, integrity, confidentiality, and traceability; however, traditional single-sensor techniques frequently fail in low-light, occluded, or loud environments. This work presents a hybrid indoor security system that integrates Ultra-High Frequency (UHF) RFID, millimeter wave (mmWave) radar (IWR1642), and YOLOv5-Tiny-based computer vision to detect, track, and identify persons in enclosed locations such as museums and exhibition halls. Each sensing modality serves a distinct purpose: RFID checks identity and counts authorized entries/exits; mmWave radar follows movement in all visibility conditions and offers blind-spot coverage; and the vision subsystem makes high-speed facial/person recognition to enforce blacklists. A Kalman filter-based fusion technique synchronizes asynchronous sensor streams, while TensorRT edge acceleration enables low-latency vision inference. Experimental results in dynamic indoor environments indicate that the proposed model increases detection accuracy by 14%, lowers false positives by 22%, and achieves real-time performance (24 FPS) when compared to baseline single-sensor models. This multi-sensor fusion proposed architecture overcomes the constraints of standalone systems by providing a scalable and privacy-preserving solution for current smart surveillance applications.

Keywords Human Tracking, Indoor Surveillance, Sensor Fusion, Mmwave Radar, Yolov5, UHF RFID, Real-Time Detection, Smart Security

DOI: 10.19139/soic-2310-5070-2680

1. Introduction

Non-specialists may struggle to upgrade or install security systems. While a correctly constructed surveillance system considerably improves emergency response, a poor configuration renders it ineffective. Prioritizing critical components improves property protection. In sensitive, confined areas, a simple camera recording system is sometimes insufficient; reliable, real-time tracking is required. Although camera-based systems are effective at face recognition, they can be obtrusive and often require user consent, which may be problematic in homes and workplaces [1].

*Correspondence to: Moshira A. Ebrahim (Email: mushira.ibrahim@eng.modern-academy.edu.eg). Computer Engineering and Information Technology Department, Modern Academy for Engineering and Technology, Cairo, Egypt.

Simple indoor public venues, such as museums, exhibition halls, and government buildings, require powerful security systems that can detect, track, and respond to unlawful entry or suspicious movement. Traditional surveillance methods, which depend exclusively on cameras or RFID, sometimes suffer from environmental restrictions such as low illumination, obstacles, and restricted detection range [2], [3]. Furthermore, the lack of real-time integration of sensory inputs results in uneven data interpretation and delayed reaction.

Radio frequency (RF) approaches [4] are less intrusive, using Wi-Fi signal fluctuations to identify people without requiring them to carry gadgets. However, these systems frequently need extra devices such as separate transmitters and receivers. Other alternatives include point cloud sensors such as LiDAR and depth cameras [5]. However, LiDAR is too costly for household usage, while depth cameras have limited range and accuracy. Both LiDAR and depth cameras face challenges with user acceptance.

Recent improvements in sensing technology and embedded intelligence have created new opportunities for hybrid systems that use complementary sensors to overcome individual constraints. For example, mmWave radar detects the mobility of objects using a single device regardless of visibility conditions [6], UHF RFID allows identification verification at entry points [7]. Furthermore, deep learning-based computer vision approaches such as YOLOv5 provide facial recognition and visual tracking [8]. Security operations are divided into securing open areas and securing closed areas. In this work, the focus is on securing closed places such as museums, theatres, cinemas, and others. However, most present systems run most object tracking sensors in isolation or lack real-time synchronization and data fusion, limiting their usefulness in dynamic contexts [9].

The main objective of this paper is to develop a unified, real-time, hybrid security tracking and identification system tailored for indoor surveillance. The main contributions of this work are as follows:

- Hybrid multi-sensor architecture that integrates UHF RFID, mmWave radar, and YOLOv5-Tiny-based computer vision to provide identity verification, blind-spot motion tracking, and real-time facial/person identification.
- Real-time sensor fusion framework with a Kalman Filter, time-synchronization protocol, and adaptive sensor weighting to address environmental changes.
- Optimized vision processing with TensorRT acceleration for high frame rates (24 FPS) without affecting detection accuracy.
- Experimental evaluation included baseline comparisons with single-sensor approaches, occlusion-resilience testing, and performance analysis in various illumination and interference scenarios.
- Considerations for scalability and privacy, including multi-room installations and GDPR-compliant face data management. This combination enables timely detection of anomalies such as unauthorized access, overcrowding, or sensor failure, thereby enhancing security responsiveness.

The remainder of this paper is organized as follows: Section 2 reviews related works. Section 3 details the system architecture and methodology. Section 4 presents experimental results and performance evaluation. Section 5 concludes the paper with future research directions.

2. Literature Review

Among the many research investigations looking at the use of particular technologies for interior tracking and security are vision systems, millimeter-wave (mmWave) radar, and RFIDs. Most of these techniques, however, operate alone and are therefore less effective in uncertain or blocked surroundings. The pertinent literature is reviewed in this part, together with the main flaws that our suggested unified framework tries to treat.

Zhao et al. [6] suggested an mmWave radar-based system for monitoring and recognizing humans in smart environments that use sparse point clouds and trajectory linkage. While the radar was excellent at detecting mobility, the system lacked authenticity verification and was vulnerable to error messages without auxiliary sensors. In the field of RFID-based applications, Fadzir et al. [10] developed an RFID-GSM system to track kids on school vehicles and notify parents. Dias et al. [11] used RFID cards for automated attendance in university classrooms. Similarly, Ouyang et al. [12] used RFID for vehicle management and theft detection. These techniques illustrate RFID's usefulness for presence logging, but they do not take into consideration in-room mobility or

behavior analysis. Furthermore, they are frequently constrained by short-range HF RFID devices, which our approach addresses by employing UHF RFID for longer-range recognition.

Khalid et al. [13] evaluated security problems of RFID and identified flaws in authentication schemes. This study also noted the lack of complete privacy guidelines for various RFID-based authentication algorithms. These studies underline that RFID devices alone cannot ensure safe identification and that multi-factor authentication methods such as facial recognition and positioning tracking are required, which our solution delivers via camera and radar integration. Yang et al. [14] demonstrated an intelligent door system with smart cameras for access control, whereas Zhang et al. [15] examined the installation of a camera surveillance system for smart city transformation. Cameras are beneficial for visual tracking, but they are restricted in dark, smoky, or occluded settings, which is why we utilize mmWave radar as a complementary sensor.

Li et al. [16] developed an imaging-based indoor positioning system that employs deep belief nets and geographical landmarks. Li L. [17] used visual SLAM and lasers for autonomously navigating. Both technologies demonstrate the promise of vision systems for mapping and tracking, but they are computationally costly and susceptible to visual distortions. To increase accuracy, Deng H. [18] presented a multi-camera system for real-time drone monitoring with minimal fish-eye distortion. However, these systems are sometimes expensive and hard to operate in public indoor settings. In contrast, our proposed system utilizes a single wide-angle camera calibration for fish-eye correction, reducing both cost and complexity.

Other radar-based studies include Kanno et al. [19] and Guan et al. [20], who investigated radar for 3D mobility mapping and removing foreign objects and debris (FODs) from airport runways to provide high-level safety in aircraft movement and remove any obstacles in the path. Although these demonstrate the resilience of mmWave radar, their use cases differ dramatically from indoor person tracking. Guan J. [21] focused on improving radar signal precision in low-resolution circumstances, leveraging the fact that mmWave signals have favorable propagation characteristics in low visibility conditions. Whereas Long N. [22] advocated merging Red, Green, and Blue-Depth (RGB-D) sensors with mmWave radar to identify obstacles for visually impaired users, emphasizing the benefits of sensor fusion.

In addition, Song et al. [23] developed a robust curved-road tracking model using an adaptive Kalman filter in autonomous systems, which influenced our decision to use a Kalman filter to combine sensor data from radar, camera, and RFID streams. Cui H. [24] investigated mmWave radar-based recognition for person tracking and discovered substantial false-positive rates. This study demonstrated that using two mmWave radars can enhance the system precision and eliminate the large number of false alarms that were raised due to unstable data and noise. Huang X. [25] proposed a rapid indoor people tracking model that employs recursive Kalman filtering (RKF) to outperform the exponential Kalman filter (EKF) drawbacks in terms of complexity and computation time. Also, the study presents two clustering strategies to offer high accuracy and low processing time. While accurate, it lacked multimodal verification and a user identification method.

According to recent studies, mmWave radar can reliably locate and count people indoors with sub-meter accuracy and continue to function even in the presence of partial occlusions; multi-radar systems also make fall detection and multi-person tracking possible [26]. While broader fusion frameworks (e.g., DRL-driven SLAM and hybrid radar-wearable tracking) demonstrate the advantages of combining modalities, complementary research in sensor fusion focuses on calibration, alignment, and deep radar-camera fusion [27]. On the compute side, YOLO-based pipeline edge deployments show real-time object detection on limited hardware, which encourages on-device analytics for surveillance [28].

However, recent studies seldom quantify accuracy-latency trade-offs for multi-sensor pipelines on the edge and usually lack identity-aware access control (RFID-based authorisation) within the fusion loop. Furthermore, it is uncommon for technical PPFR studies to link their designs to functional GDPR controls in interior security systems. By adopting a Kalman-based framework that is accelerated on-device to combine RFID identification with radar/vision kinematics and by putting in place privacy protections that comply with current regulatory advice, our approach fills these gaps.

The shortcomings mentioned in the literature reveal a common issues such as single-sensor systems lack resilience, whereas sensor fusion systems frequently ignore real-time integration and identity verification. Our proposed hybrid overcomes the constraints of single-sensor systems by integrating UHF RFID, mmWave radar, and

YOLOv5-Tiny vision for real-time integration and identity verification. It implements real-time multi-sensor fusion with a Kalman filter for consistent person tracking, optimizes latency and processing with edge-based acceleration for the camera module, and includes an integrated GUI for live monitoring, alerting, and anomaly detection, which was not present in previous works.

3. System Architecture and Methodology

This section describes the revised design of the proposed hybrid indoor security system. The system combines three subsystems—RFID, mmWave radar, and computer vision (camera-based) to allow real-time person detection, tracking, and identification in confined spaces like museums and exhibition halls.

3.1. System Overview and Sensor Fusion Framework

Figure 1 shows the proposed architecture for a multi-sensor modular, synchronized framework. Unlike earlier research that only vaguely links sensor devices, this system makes use of a centralized processing module with real-time data fusion capability. The layered architecture was employed as the following:

- The data acquisition layer includes a UHF RFID reader, mmWave radar (TI IWR1642), and cameras (YOLO-tuned vision system).
- The synchronization and preprocessing layer synchronizes time-stamped data with software-based buffer queues and shared system clocks.
- Fusion and Decision Layer: Uses a Kalman Filter-based data fusion method to correlate detections from several sensors.
- Interface Layer: A MATLAB-based GUI enables real-time tracking, occupancy validation, and anomaly detection.

In this proposed model, a scenario is proposed for the security operation. A list of the names and pictures of unwanted persons must be prepared to identify and deal with them by the concerned security services before entering the place to be secured. The following steps show the proposed building algorithm:

Algorithm 1 Proposed Building Security Algorithm

- 1: **External Camera Unit:**
 - 2: Capture an image of the visitor.
 - 3: Compare the visitor image with the blocked image database.
 - 4: **RFID Unit:**
 - 5: Issue an RFID card to the visitor.
 - 6: Security officer verifies visitor ID.
 - 7: **Millimeter Wave Radar Unit:**
 - 8: Count and track the people in the room.
 - 9: Provide a motion map.
 - 10: **Internal Camera Unit:**
 - 11: Detect, track, and count people in real time inside the room.
 - 12: **Control Unit:**
 - 13: Compare the counting results of the three units.
-

This proposed algorithm is suggested to improve the building's security and safety based on the combination of three units: the camera, RFID, and millimeter-wave radar.

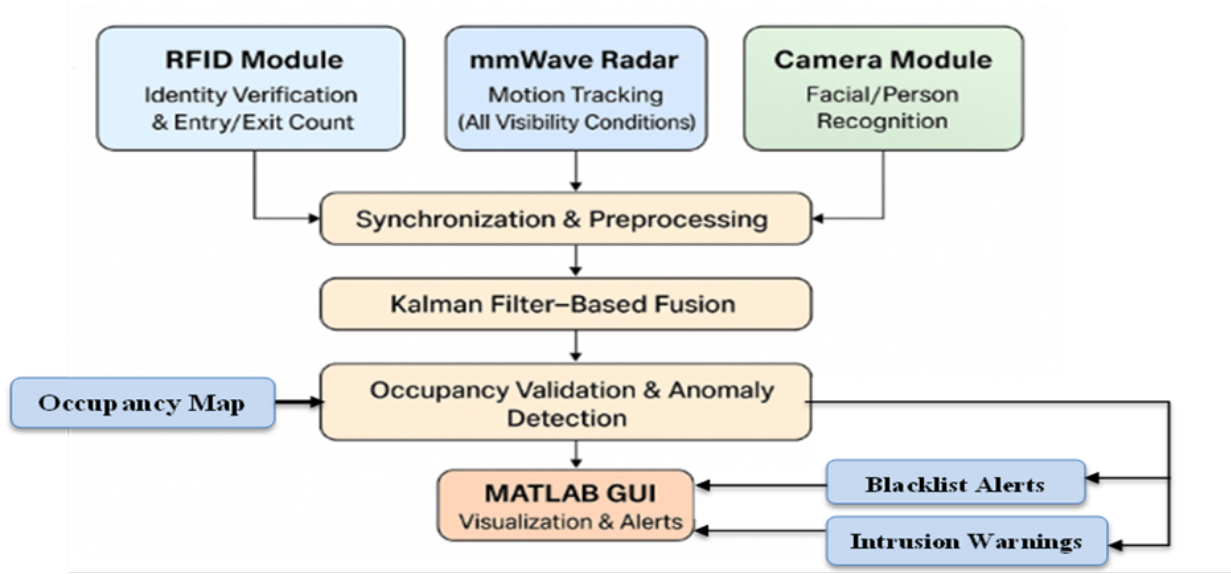


Figure 1. Proposed System Architecture Diagram.

3.2. RFID Module (Enhanced UHF Implementation)

To overcome the restricted range issue of the original 13.56 MHz RFID design, the system utilized Ultra High Frequency (UHF) RFID (860-960 MHz) with passive tags. These have an effective range of 1-5 meters, allowing for accurate monitoring of entry and departure locations. The utilized hardware is an Impinj Speedway R420 reader with circularly polarized antennae. It is positioned at entry and exit gates, with each tag scan time-stamped and recorded. The studied case monitors legitimate entry and departures. Each tag corresponds to a user profile in the central system. The RFID module aims to provide some legal entries from the entrance of the indoor area by giving every visitor an RFID card. The coupling coefficient K [29] is the most important parameter. K is affected by the distance between the separated two-loop antennas. To determine the coupling coefficient K :

$$K = \frac{\mu_0 n_1 n_2 \pi d_1^2 d_2^2}{2 \left(\sqrt{L_{\text{READER}} L_{\text{TAG}}} \cdot \sqrt{(z^2 + d_2^2)^3} \right)} \quad (1)$$

Where μ_0 is the magnetic constant ($4\pi \times 10^{-7}$). n_1 and n_2 are the number of turns of each antenna coil. $d_1 = 7$ cm and $d_2 = 8$ cm are the diameters of Tag and Reader antennas respectively. z is the separating distance between the two coil antennas. L_{READER} is the reader inductance. L_{TAG} is the tag inductance.

To determine the magnetic coupling between the two antennas [30, 31]:

$$z = \sqrt{x - d_2^2} \quad (2)$$

$$x = \frac{\sqrt[3]{(\mu_0 n_1 n_2 \pi d_1^2 d_2^2)^2 \cdot 3}}{2^{2/3} k^{2/3} \sqrt{L_{\text{READER}} \cdot L_{\text{TAG}}}} \quad (3)$$

Figure 2 presents the variation of coupling coefficient and mutual inductance as a function of variable distance [32]. The design is based on two RFIDs, one on the entrance door, while the other is on the exit door. In order to achieve this design, there are two important aspects the first one is how to sync the two RFIDs on the same MCU, while the second is the PCB design.

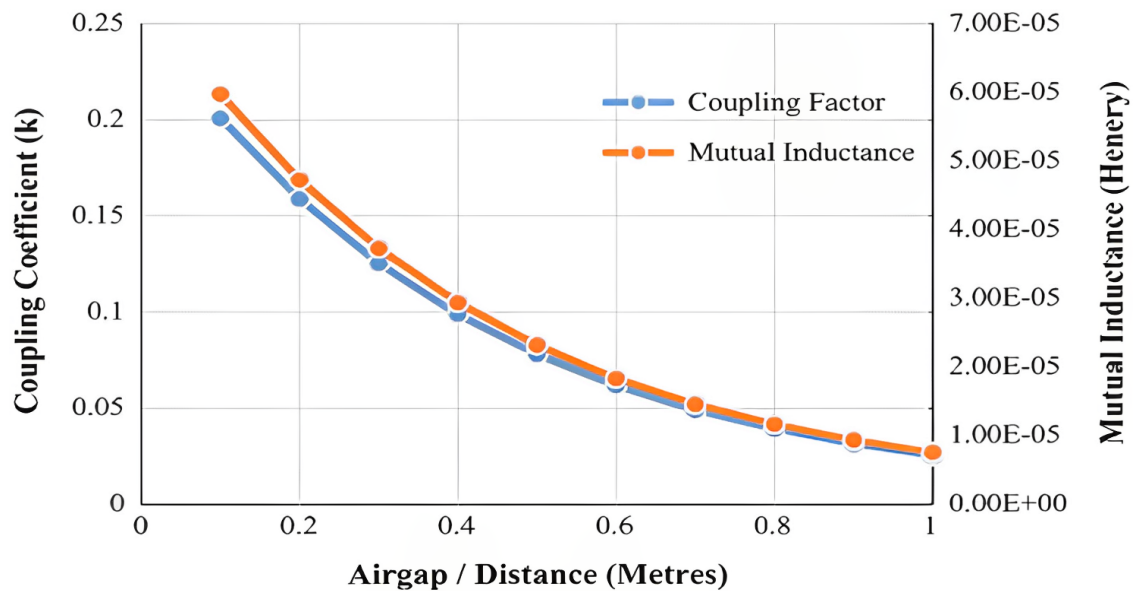


Figure 2. The Variation of Coupling Coefficient and Mutual Inductance[32].

Popular 5V microcontrollers such as Arduino and Basic Stamp typically require level shifting in sensors, displays, and streak cards when interacting with newer 3.3V devices. A bidirectional logic level converter, suitable for protocols like I2C and TTL serial, overcomes this by providing voltage conversions from 1.8V to 10V. It serves as a voltage switcher and communication bridge, enabling an Arduino to communicate with two RFID modules.

This allows for simultaneous scanning and speedy data transfer to the Arduino. While 5V microcontrollers such as Arduino and the Basic Stamp are commonly utilized for level shifting in a variety of applications, several modern 3.3V devices need a voltage downshift.

A bidirectional logic level converter that operates from 1.8V to 10V was developed. This converter is compatible with protocols like I2C and TTL, and performs as both a voltage shifter and a communication route. It also enables an Arduino to communicate with two RFID modules. As shown in Figure 3, the microcontroller may share a bus with both RFIDs, ensuring simultaneous scanning and rapid data delivery to the Arduino.

3.3. mm Wave Radar Unit (IWR1642)

Even in low-visibility situations, the mmWave radar system is used to count and track people in a room and then provide a map of their movements. Millimetre-wave radar is based on the Frequency-Modulated Continuous Wave (FMCW) technique which allows the simultaneous measurement of a target's relative radial speed and range.

3.4. Camera and Vision-Based Tracking

The camera module objective is maintaining a full view of the monitored room, performing real-time human detection and tracking, comparing occupant counts with RFID and radar, and recognizing blacklisted individuals at the entrance, triggering alarms for unauthorized access. The system utilized a dual-camera configuration, with an entrance camera for face recognition against a blacklist dataset and a main surveillance camera for indoor monitoring, both equipped with a 1536x1536 resolution fisheye lens for ultra-wide coverage and night vision. This design provides visibility of the entire room and eliminates dead zones. Table 1 lists the specifications of both cameras.

For fulfilling our target to have full view of the room ensuring no dead-regions, our selection for an omnidirectional camera with an ultra-wide viewing angle was a necessity. There are 2 main properties for the camera to be considered wide-angled: focal length and the field of view.

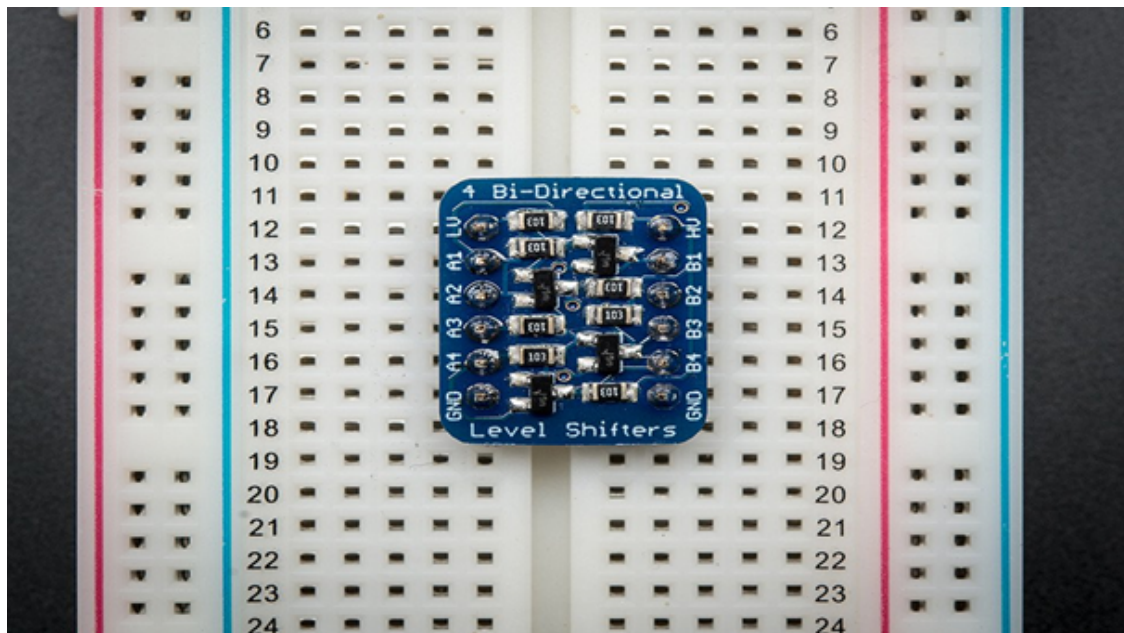


Figure 3. Bidirectional logic level converter.

Table 1. Specifications of Main Camera and Entrance Door Camera

POC	Specification	Main Camera	Entrance Door Camera
Field of View (FOV)	Angle	360° (Fisheye Viewing Lens)	> 100°
Resolution	Pixels	1536 × 1536 – 3.0MP	(Lower resolution camera)
Night Vision Ability	IR Capability	10m range IR sensor	Not required
Operating Power	Consumption	3W (max)	–

Ultra-Wide-angle cameras have a focal length less than 24mm, allowing for a wide field of view (FOV) that exceeds 120 degrees and reaches up to 360 degrees. Wide angle viewing is achieved using fisheye lenses or rectilinear lenses. Rectilinear lenses map the scene on a 2D image without curvilinear distortion, mapping straight lines in the real world to the image generated by the camera. However, this model cannot be achieved for viewing angles more than 180 degrees, leading to the selection of a fisheye camera model with a FOV of 360 degrees.

Camera calibration is done using specific pattern images, e.g., a chessboard pattern image, where the image is printed, and many images (minimum of 15 images) are captured to have multiple views and rotations of the chessboard pattern. In all images, the internal chessboard corners are extracted, and the positions are returned. The typically used chessboard pattern has internal 6x9 corners. The Python OpenCV library offers a toolbox for camera calibration, which is used to generate the camera matrix and transform the image from fisheye view to pinhole view. The flowchart shown in Figure 4 represents the sequence of operation of the calibration method used.

For face recognition, the YOLOv5-Tiny model is trained on a custom dataset with the OIDv4 toolbox. Face embeddings are used to compare against a blacklist. The YOLOv5-Tiny was chosen for human tracking due to its efficient speed and accuracy. To overcome delays, the model is tuned with TensorRT, obtaining 24 FPS on Jetson Nano hardware. The Fisheye distortion is rectified using checkerboard calibration (50-image dataset), and undistorted frames are supplied into the detector.

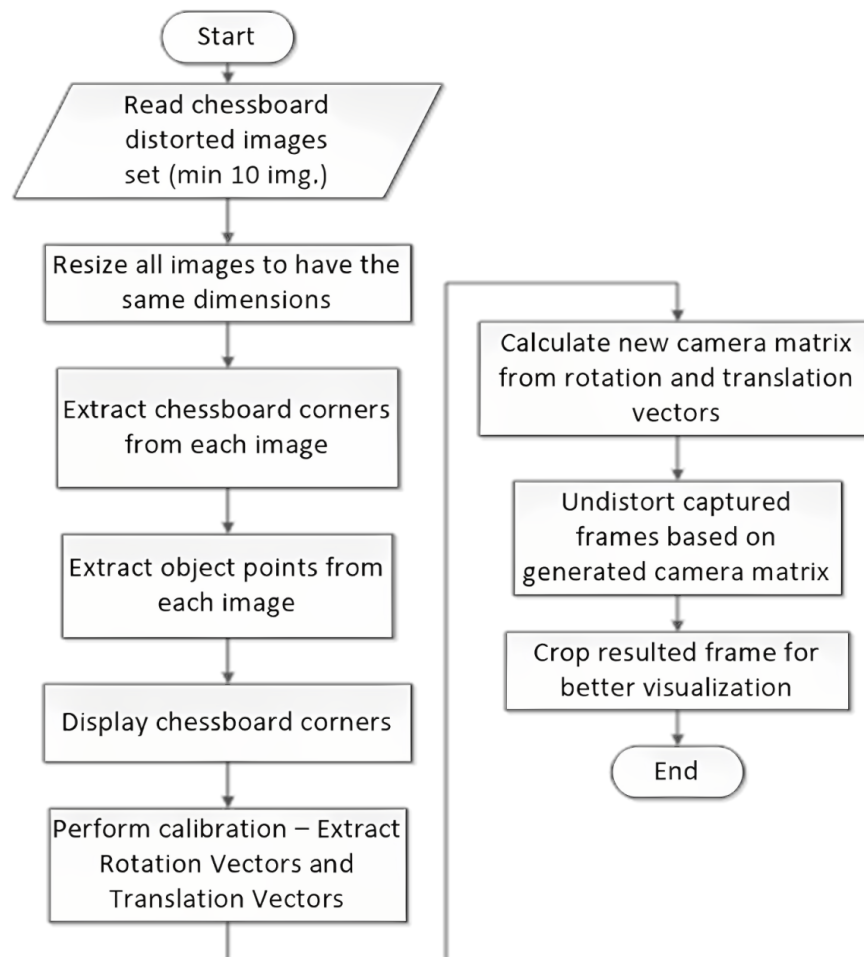


Figure 4. Flowchart of camera calibration algorithm.

In YOLO algorithm, we are not looking in our image for interesting regions that might contain any object. As shown in Figure 5, we divide our picture into cells, typically a 19x19 grid. Every cell will be responsible for predicting five bounding boxes (if that cell contains more than one object). This will give us a picture of the 1805 bounding boxes. Confidence prediction is essential due to the ubiquity of vacant cells and boxes. Low-confidence boxes are next deleted, and non-max suppression, as shown in Figure 6, is used to eliminate overlapping boxes with a large common area. Table 2 illustrates a YOLO CNN architecture, comprising convolutional and max pooling layers, followed by two fully connected layers. The proposed YOLO model was trained using the COCO (Common Objects in Context) dataset, which is a large-scale object detection, segmentation, and captioning dataset. COCO dataset consists of 200,000 photos from 80 categories and over 500,000 item annotations. Figure 7 depicts the human tracking process, which uses the YOLO algorithm. As seen in Figure 8, YOLO algorithm can recognize and track people with high confidence. However, the first installation was significantly delayed. To address this, we trained a special dataset dedicated just to the "person" class. This reduction in the number of classes reduced processing time and caused insignificant delays.

Based on feedback from security employees, the basic concept has been improved for real-world applications. The revised strategy entails developing a blacklist dataset for facial recognition at the admission gate. This mechanism will prevent banned persons from entering. A second camera within the room will then identify, track, and count persons, merging the results with other subsystems.

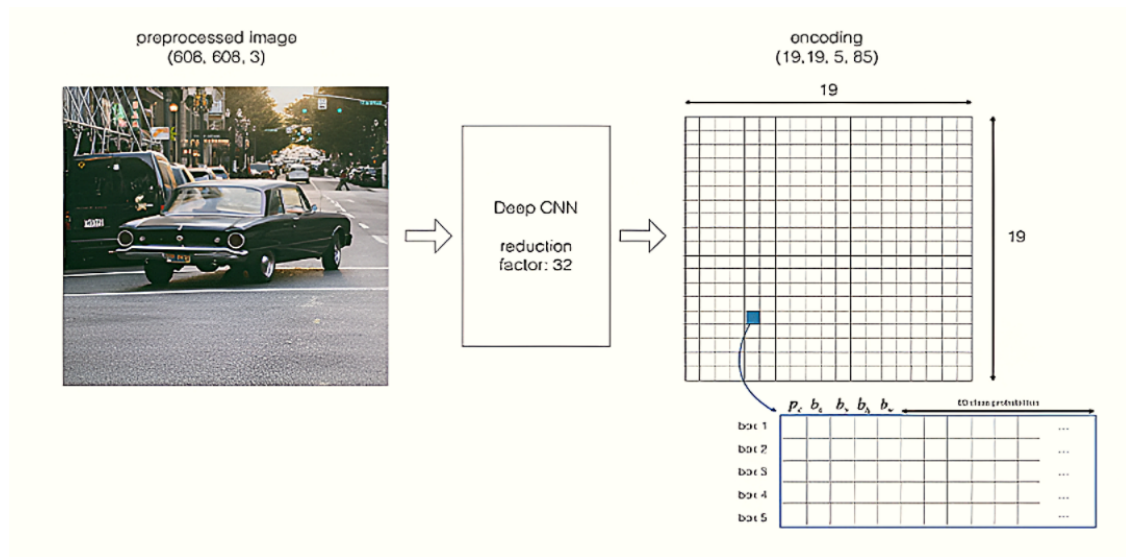


Figure 5. YOLO5 output encoding.

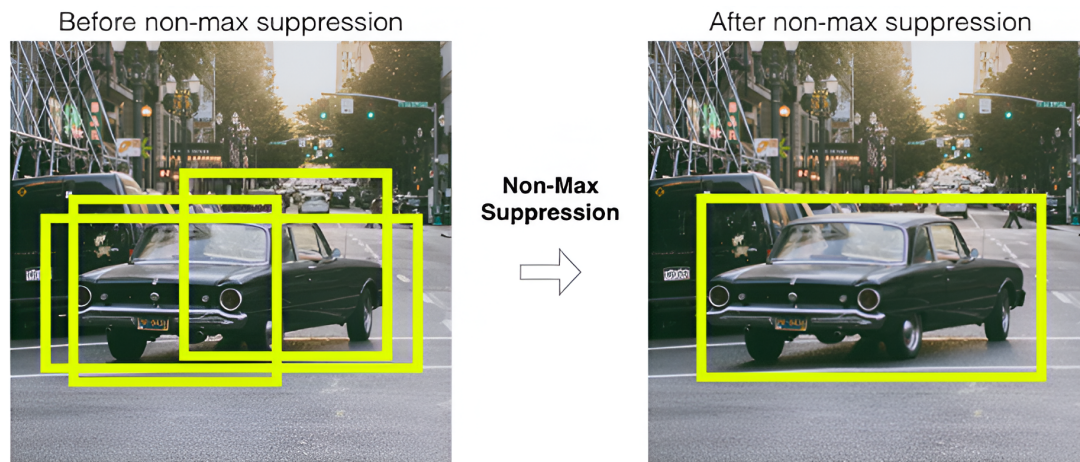


Figure 6. Non-Max Suppression.

Face recognition process recognizes special people, whereas face detection only identifies the existence of a face. At the entrance, the system compares collected facial representations to a list of unauthorized people and denies admittance if a match is detected. Inside, facial detection monitors movement and counts individuals in real time. To be able to detect and track people in different poses and positions, the YOLO algorithm was conducted. Google Images V5 was used as a source of images. Since training requires at least one thousand images, it is not practical to download each individually. There is a tool available on GitHub to download datasets along with annotations called the OIDv4 toolkit [33]. The last trained Y-Tiny model is used to achieve an acceptable frame rate with negligible delay and to have high confidence in detection [34].

3.5. Time synchronization between RFID, radar and camera

Coordinated perception between cameras and radar has become essential for comprehending the surroundings in multi-sensor fusion systems. Radar uses laser or millimeter-wave technology for accurate distance estimations,

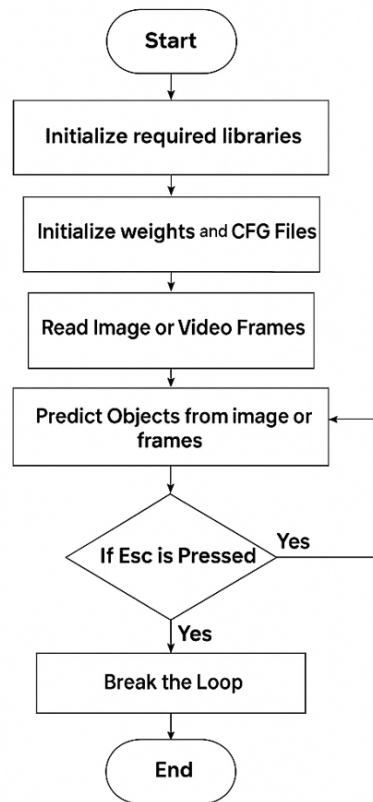


Figure 7. Flowchart of YOLO Algorithm.



Figure 8. YOLO algorithm results for human detection.

Table 2. YOLO Convolutional Neural Network (CNN) Architecture

Layer Name	Filters / Kernel Size (Stride)	Output Dimension
Conv 1	$7 \times 7 \times 64$ (stride = 2)	$224 \times 224 \times 64$
Max Pool 1	2×2 (stride = 2)	$112 \times 112 \times 64$
Conv 2	$3 \times 3 \times 192$	$112 \times 112 \times 192$
Max Pool 2	2×2 (stride = 2)	$56 \times 56 \times 192$
Conv 3	$1 \times 1 \times 128$	$56 \times 56 \times 128$
Conv 4	$3 \times 3 \times 256$	$56 \times 56 \times 256$
Conv 5	$1 \times 1 \times 256$	$56 \times 56 \times 256$
Conv 6	$3 \times 3 \times 512$	$56 \times 56 \times 512$
Max Pool 3	2×2 (stride = 2)	$28 \times 28 \times 512$
Conv 7–15	Mixed 1×1 and 3×3 filters	$28 \times 28 \times 1024$
Max Pool 4	2×2 (stride = 2)	$14 \times 14 \times 1024$
Conv 16–23	3×3 filters (stride = 2 at Conv 17)	$7 \times 7 \times 1024$
FC 1	–	4096
FC 2	–	$7 \times 7 \times 30$ (1470)

while cameras use texture information for object detection. RFID uses radio waves for data transfer between a chip and a reader. These methods address single-sensor limitations by combining data from three sensors. Common time synchronization, such as Network Time Protocol (NTP) which uses network servers to synchronize system time for millisecond-level accuracy, is suitable for low precision needs. While Precision Time Protocol (PTP) offers high-precision time synchronization, typically used in LANs, and compensates for network delays for device clock alignment using a master-slave architecture and bidirectional message exchanges, as illustrated in Figure 9. Both methods are easy to deploy and vulnerable to network delay.

GPS receivers provide PPS (pulse per second) signals that are precisely synchronized with Universal Time Coordination (UTC) second pulses. Devices equipped with hardware interfaces, such as TTL (transistor-transistor logic) circuits, are designed to capture the rising edges of these PPS signals. They then utilize this hardware pulse as a time reference to adjust local clocks or triggers. Unified physical trigger pulses, such as TTL low/high-level signals, are produced by FPGA, MCU, or dedicated triggers. Every device acquires data frames at the same time, and hardware directly generates the timestamps.

Time-aligned fusion techniques combine data from multiple sources, such as sensors or databases, to ensure synchronization with a shared chronology. This is crucial when data sources function independently, with varying sample rates, acquisition durations, or internal clock drifts. The goal is to create a cohesive dataset with properly synchronized matching events or measurements. The procedure involves coarse alignment, bringing data into a suitable time range with one-second accuracy, and fine alignment, using a specialized synchronization signal such as PPS signal. Fusion can involve simple concatenation or more complex processes such as weighted averaging or Kalman filtering.

3.6. Kalman filtering

A recursive technique called a Kalman filter generates precise estimations of a system's unobservable state by using a sequence of noisy measurements taken over time. The Kalman filter minimizes estimation error variance to produce optimal estimates by merging noisy sensor data with a system's dynamics model for state prediction and correction. When direct measurements are incorrect or unavailable, it is frequently employed in statistics and control theory to follow objects and estimate states. Figure 10 shows a Kalman filter structure.

The Kalman filtering algorithm is a recursive state estimate method that integrates predictions and measurements from multiple sensors to obtain the best estimate of the system state from a dynamic model. It offers a statistically ideal estimate in terms of minimum mean square error, assuming a linear system model and Gaussian noise distributions. The algorithm's two primary phases are prediction and updating. The prediction phase propagates the state estimate and its covariance forward in time, while the upgrade phase adjusts the state estimate and covariance

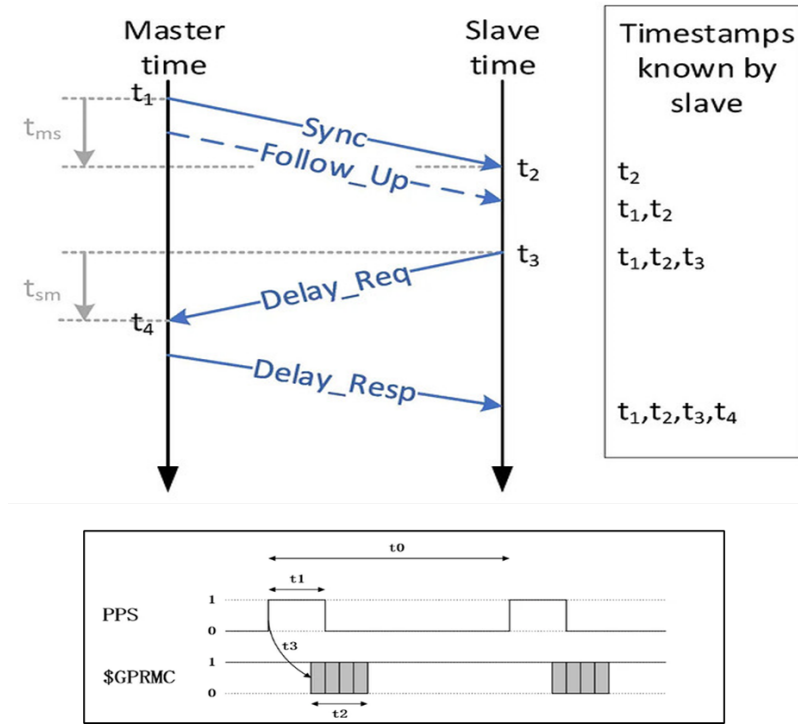


Figure 9. Master-Slave Time Synchronization.

using updated measurements from the sensors. The Kalman filter is widely used in control, tracking, and navigation fields. Its key benefits include managing noise and uncertainty in the system model and sensor data, accounting for each sensor's unique uncertainties and accuracy, and incorporating previous knowledge of system dynamics and measurement procedures.

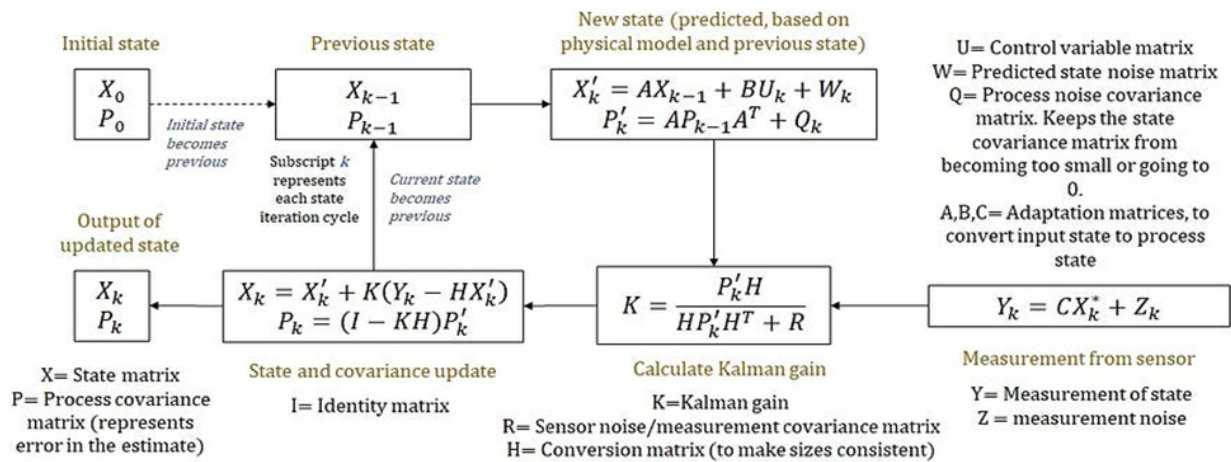


Figure 10. Kalman filter structure.

4. Experimental Results and System Analysis

To enable real-time decision-making, Python scripts (for YOLO and RFID handling) are compiled as callable functions. MATLAB is used for GUI, logic handling, and visualization. A socket-based communication protocol ensures timely data transfer between Python and MATLAB. A time-aligned fusion algorithm matches tracked individuals across radar, camera, and RFID streams. Any discrepancy in the number of detected individuals through the three subsystems triggers a real-time alert to security personnel.

4.1. The RFID Module

The Advanced Design System (ADS) program is used to simulate the magnetic coupling of the RFID system. The average power needed to power up the proximity inductive coupling card (PICC) is five volts (DC) to get an effective connection between the reader and the tag.

After the simulation, the result shows that when Z (distance) is 5.2 cm, there will be an output voltage of five volts, which is acceptable and sufficient to power up the PICC (TAG). Both inductances (L_{READER} and L_{TAG}) ensure the magnetic field generation and allow signal receiving, respectively. The generated magnetic field which is the result of the interaction of both coils, is represented by K (coupling coefficient), in which the simulation shows a significant decrease with an increase in the separating distance as shown in Figure 11. Also, the simulation result shows that the increase in the separating distance makes the output voltage decreases until it will not be enough to power the tag. At a distance above 10.2 cm, the output voltage will be less than 3 volts, which is not able to make the tag able to send or receive signals, and at 20, there will be zero volts, which means no connection as explained in Figure 12. Table 3 illustrates the relationship between the coupling coefficient, the output voltage, and the separating distance.

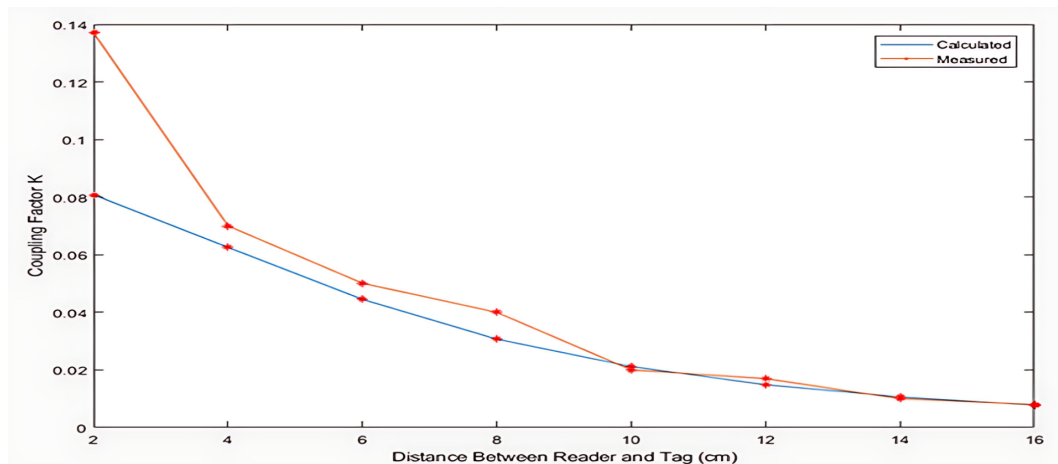


Figure 11. Designed Coupling Coefficient vs. Distance.

To achieve the proposed design, two RFIDs are used, one on the entrance door and the other one on the exit door. There are two important aspects, the first one is how to synchronize the two RFIDs on the same MCU, while the second is the PCB design. Reader module Mifare RC522 with its tags is chosen. This reader is a highly coordinated RFID card reader that manages non-contact correspondence of 13.56 MHz and is configured by the fifth-largest non-memory semiconductor supplier (NXP) as a low-power consumption, minimal-effort, and conservatively sized reader and write chip. The need for data acquisition is to transfer the data received from the RFID to an Excel sheet on a PC, so Parallax microcontroller data acquisition (PLX-DAQ) was used. Figure 13 illustrates the manufacturing of the proposed design PCBs.

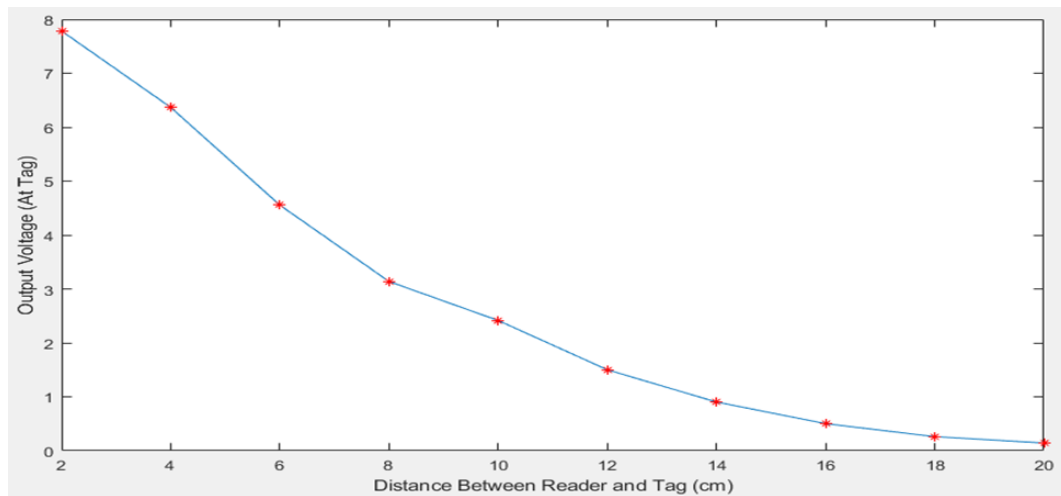


Figure 12. Modeled Output Voltage vs. Distance.

Table 3. Coupling coefficient and Output Voltage vs Distance

Separating Distance (z) [cm]	K Calculated	K Measured	% Difference	Output Voltage (V)
2.2	0.08078081	0.137	41.04	7.783
4.2	0.06254823	0.070	10.64	6.367
6.2	0.04450053	0.0501	11.17	4.554
8.2	0.03068811	0.040	23.27	3.138
10.2	0.02118119	0.020	5.90	2.415
12.2	0.01485900	0.017	12.59	1.501
14.2	0.01065686	0.010	6.56	0.903
16.2	0.00777978	0.008	2.75	0.502

There are some challenges faced with the proposed structure:

- Making a unique detection for one RFID tag and reaching a constant implementation part for adding any wanted number of cards.
- Delay in the identification time and the login time in the excel sheet.
- Delay in the messages that appear in the LCD that clarify what to do.
- Making two RFIDs read at the same time.

The first and third issues were resolved by performing some modifications on the microcontroller source code. The second issue was resolved by increasing the microcontroller's processing speed by optimizing the code and transferring data to the Excel sheet during initialization, so accelerating the fetch, decode, and execute cycles. The third issue was also resolved by modifying the program code. The fourth issue was resolved by connecting the two RFIDs to a bidirectional module, which enabled them to share the microcontroller's communication bus and ensured they were ready to read data from system startup.

4.2. The mmW radar

The radar framework utilizes IWR1642 evaluation module, a programming FMCW Radar, from Texas Instruments (TI). The FMCW radar will perform range, speed and angle estimation at that point transmit every single investigated result through UART serial communication to the PC. The PC will get the information, then do the visualization on MATLAB.

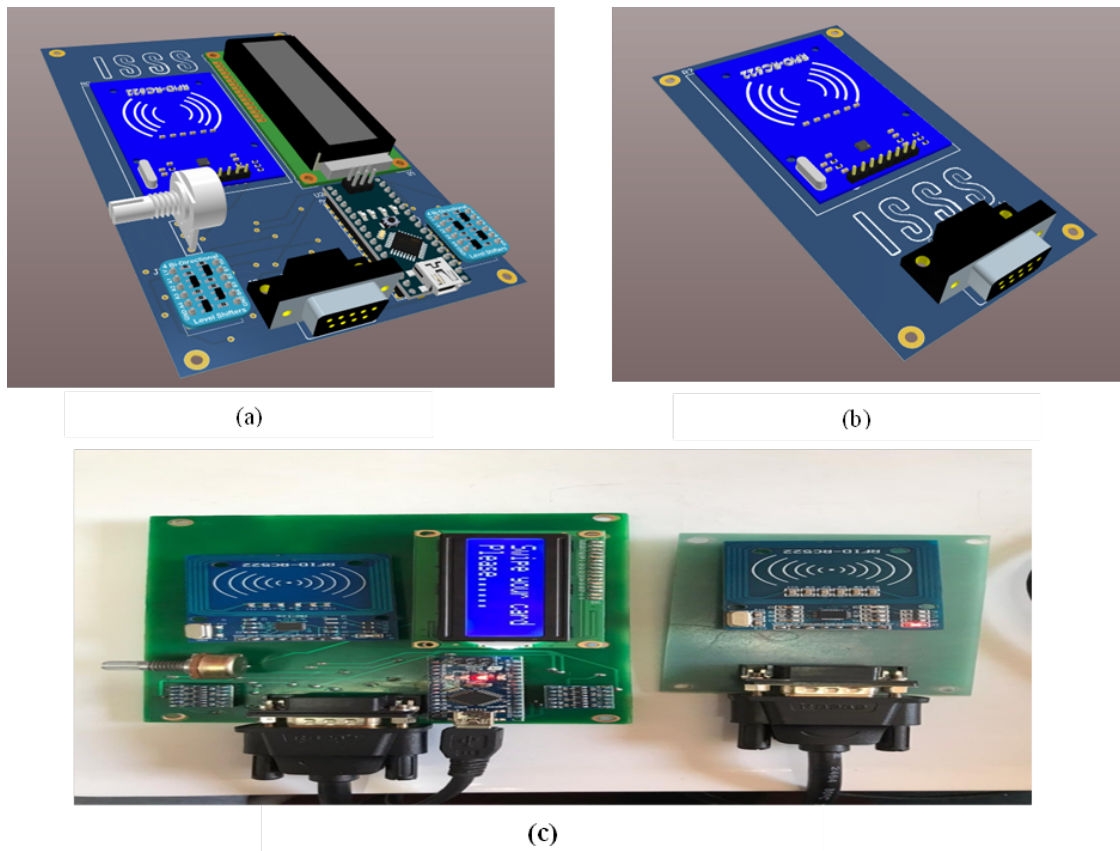


Figure 13. a) Entrance door PCB. b) Exit door PCB. c) Manufacturing the proposed design PCBs.

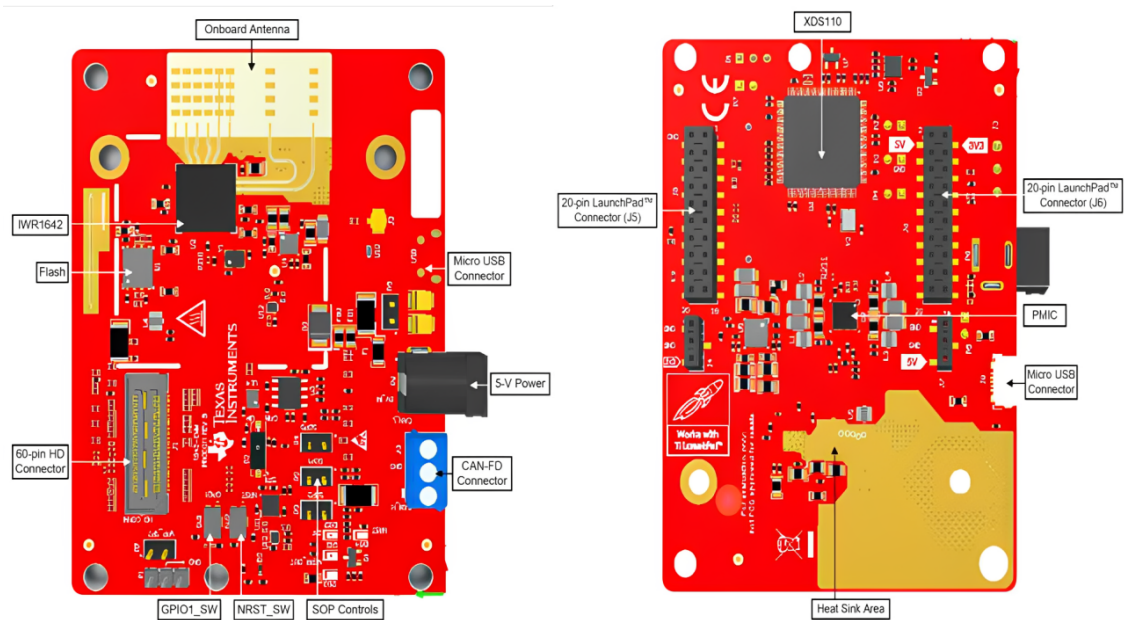


Figure 14. The IWR1642 booster pack front view and back view.

The BoosterPack contains everything required to begin creating programming for on-chip C67x DSP center and low-power ARM R4F controllers, including installed emulators for programming and troubleshooting as well as locally available catches and LEDs for a snappy combination of a basic UI. The standard 20-pin Booster Pack headers make the gadget perfect with a wide assortment of TI MCU Launch Pads and empowers simple prototyping. The IWR1642 Booster Pack, shown in front and back views, is as shown in Figure 14 and the chirp configuration for the IWR1642 Booster Pack is shown in Table 4. The IWR1642 Booster was chosen due to its features, such as,

- Two 20-pin Launch Pad connectors that leverages the ecosystem of the TI Launch Pad.
- XDS110-based JTAG emulation with a serial port for onboard QSPI flash programming.
- Back-channel UART through USB-to-PC for logging purposes.
- Onboard antenna.
- 60-pin, high-density (HD) connector for raw analog-to-digital converter (ADC) data over LVDS and trace-data capability.
- Onboard CAN transceiver.
- One button and two LEDs for basic user interface.
- 5-V power jack to power the board.

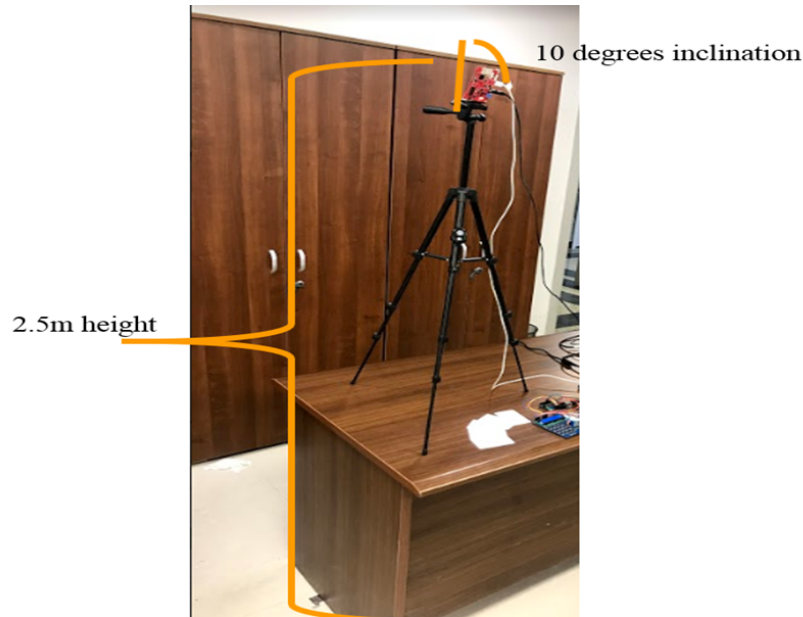


Figure 15. IWR1642 Booster Pack real-life physical setup.

The physical setup of the IWR1642 Booster Pack is illustrated in Figure 15. The result of the radar on the visualization tool is shown in Figure 16. The figure also shows the track of the targets and the count with a camera view to compare the radar readings to the camera.

4.3. Camera module

In the proposed framework, camera calibration is performed using an image set of 50 7x7 corner chessboard images at different views and angles. The results before and after fisheye distortion removal are shown in Figure 17. The stream is perfectly undistorted, with a very small dead region resulting from the transformations. This region could be simply cropped without affecting the required result. The camera matrix resulting from the calibration process, which resembles the mathematical model of the camera matrix discussed previously, is presented in Figure 18.

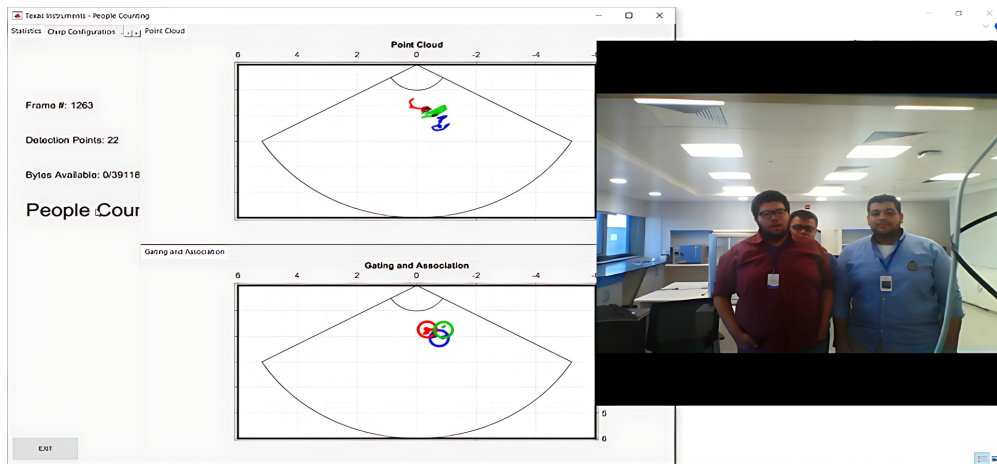


Figure 16. Radar output on the visualization tool.

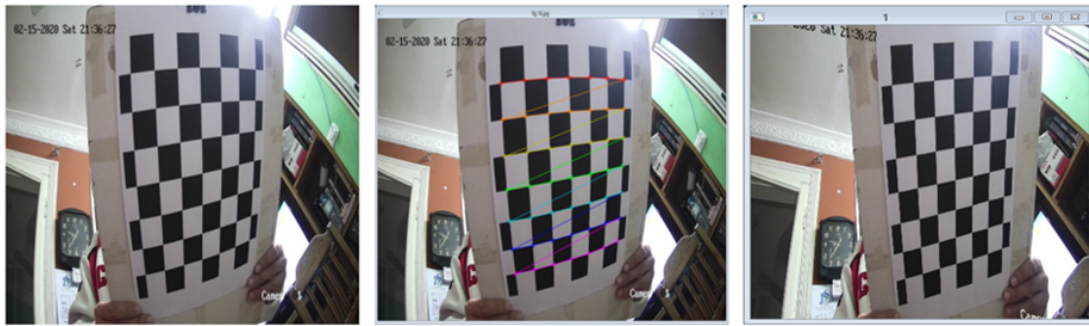
Table 4. IWR1642 Chirp Configuration

Chirp Parameter (Units)	Value
Start Frequency (GHz)	77
Slope (MHz/ μ s)	60
Samples per Chirp	128
Chirps per Frame	256
Frame Duration (ms)	50
Sampling Rate (MSPS)	2.5000
Bandwidth (GHz)	3.0720
Range Resolution (m)	0.0488
Max Unambiguous Range (m)	5
Max Radial Velocity (m/s)	5.2936
Velocity Resolution (m/s)	0.0827
Azimuth Resolution (deg)	14.5
Number of Rx	4
Number of Tx	2

As shown in Figure 19, performing the detection before undistorting the video stream results in exceptionally low accuracy at the edges due to the curvature generated by the fisheye effect. After regenerating the camera matrix and applying it to the stream, human detection is performed with less error.

Google Images V5 is used for image training, supporting 2.8 million object instance segmentation masks in 350 categories. These masks mark the outline of objects, characterizing their spatial scale to a higher level of detail. To download at least 1000 images for training, GITHUB offers the OIDv4 toolkit, which downloads the dataset and annotations.

The proposed model is trained using Keras, a Python-written high-level neural network API that runs on top of TensorFlow. To achieve reliable identification and tracking, the YOLOv5-Tiny model was trained in two steps. First, we utilized the COCO dataset, which comprises around 200,000 images over 80 categories and over 500,000 item annotations. Second, we fine-tuned the network using a custom dataset produced using the OIDv4 toolkit and Google Image V5, constrained to the single class “person.” This decreased model complexity, shortened inference latency, and enhanced accuracy in the desired scenario. The dataset was divided into 80% for training and 20% for testing. Training was performed on GPU, with inference evaluated on Jetson Nano hardware utilizing TensorRT



5a) Distorted image

5b) Corners detected

5c) Undistorted result

Figure 17. Undistorting the camera stream using camera calibration results.

```

new camera matrix:
[[6.14231812e+02 0.00000000e+00 2.71122246e+02]
 [0.00000000e+00 1.37705566e+03 3.29740216e+02]
 [0.00000000e+00 0.00000000e+00 1.00000000e+00]]

distribution parameters:
[[-3.15329016e+00 1.69983678e+01 -2.17064056e-02 -3.51979585e-04
 -4.56751440e+01]]
>>> |

```

Figure 18. The new camera matrix was generated by camera calibration and fisheye distortion.



7a) Human Detection results before stream undistorting



7b) Human detection results after stream undistorting

Figure 19. Effect of Fisheye Distortion Removal on Detection Accuracy.

acceleration. The training of YOLOv5-Tiny model was performed for 100 epochs on the COCO dataset and 50 epochs on the custom dataset.

The customized YOLO-Tiny algorithm results shown in Figure 20 ensure accurate human recognition and tracking. However, the first installation was significantly delayed. To address this, the model was trained purely on the person class, resulting in person detection accuracy at an acceptable frame rate of 24 FPS. To demonstrate the advantages of multi-sensor fusion, we compared the proposed hybrid system to individual modules that operated independently (RFID-only, Radar-only, Camera-only). Performance was evaluated in terms of detection accuracy, latency, false positive rate (FPR), and robustness to occlusion. The results are summarized in Table 5. The hybrid system attained an average accuracy of around 89% in person detection, as evidenced by our experimental configuration in Figure 21. The baseline FPS on the Jetson Nano was around 12 FPS, however optimisation with TensorRT increased the throughput to almost 24 FPS, achieving real-time performance. In comparison to single-sensor baselines, the hybrid system not only decreased false positives (8% versus 12–15%) but also preserved resilience under occlusion due to the incorporation of mmWave radar.

Table 5. Performance Comparison of Different Systems

System	Accuracy (%)	Latency (ms)	FPS	False Positive Rate (%)	Occlusion Robustness
RFID-only	78	10	–	12	Low
Radar-only	82	15	–	10	High
Camera-only	85	40	12	15	Low
Hybrid with TensorRT acceleration (Proposed)	89	20	24	8	High

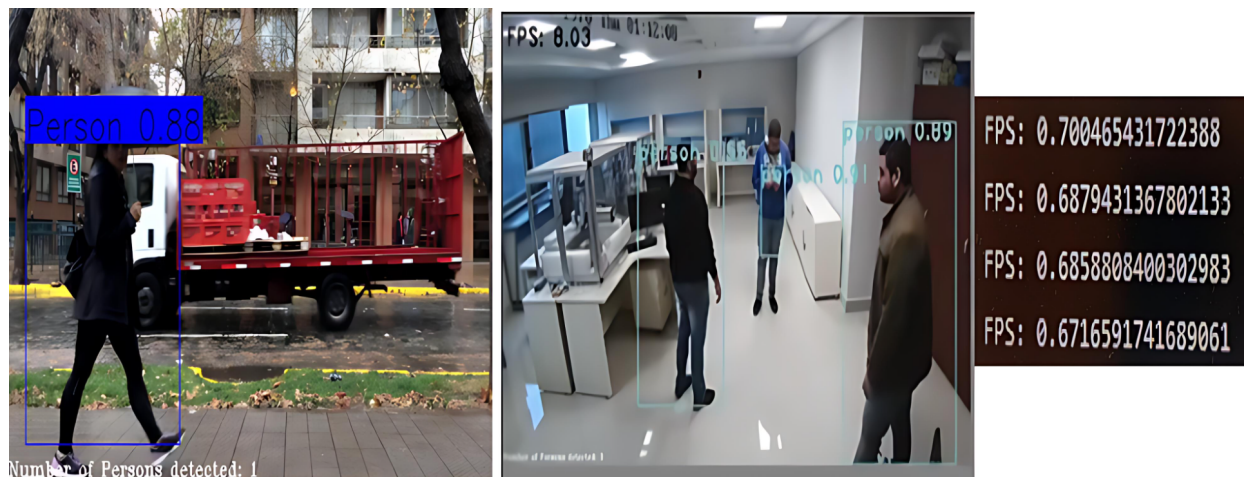


Figure 20. Final customized YOLO-Tiny algorithm results.

4.4. The System Integration

The three systems were merged into a single PC software using MATLAB and compiled into an .exe file that just requires the MCR MATLAB Runtime to be installed. The MATLAB-based GUI shown in Figure 21 validates people counting by analyzing data from RADAR, RFID, and camera sensors. To allow communication between Python and MATLAB, the Python script was transformed into callable functions. The MATLAB version must be compatible with the Python version, and all necessary libraries must be installed to ensure efficient configuration and execution within MATLAB, taking into account library version compatibility with both platforms. The first integration approach was used. Figure 22 depicts the combined result of a proof-of-concept scenario for the interior space.

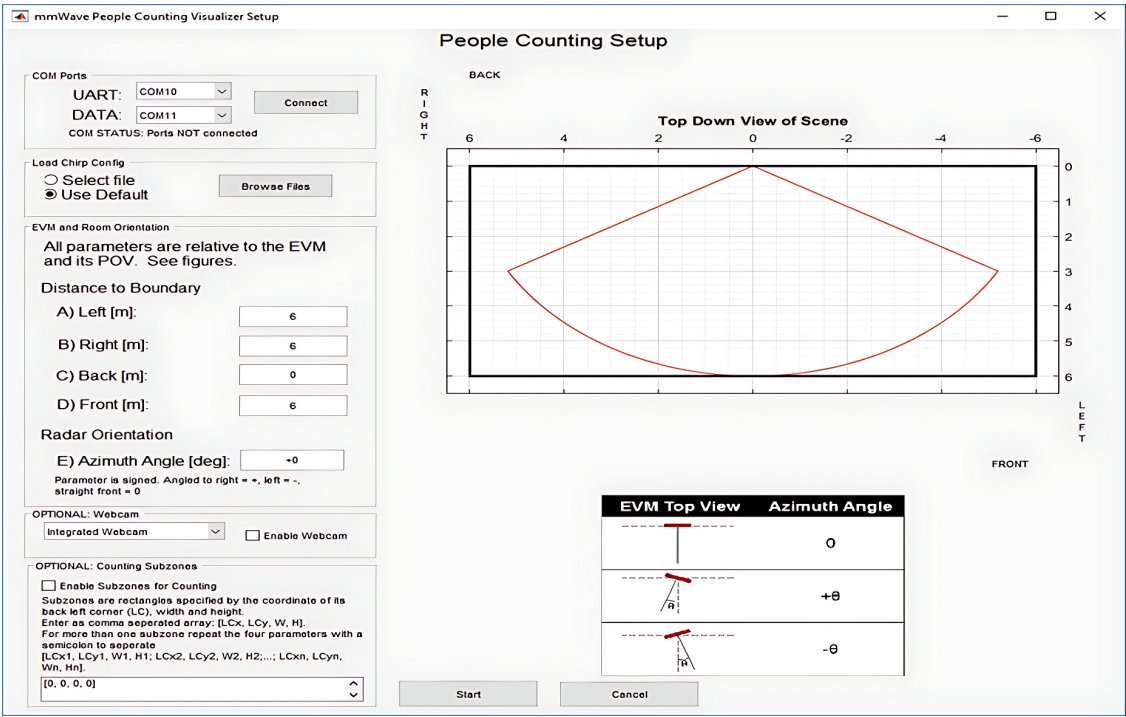


Figure 21. MATLAB GUI for Visualization.

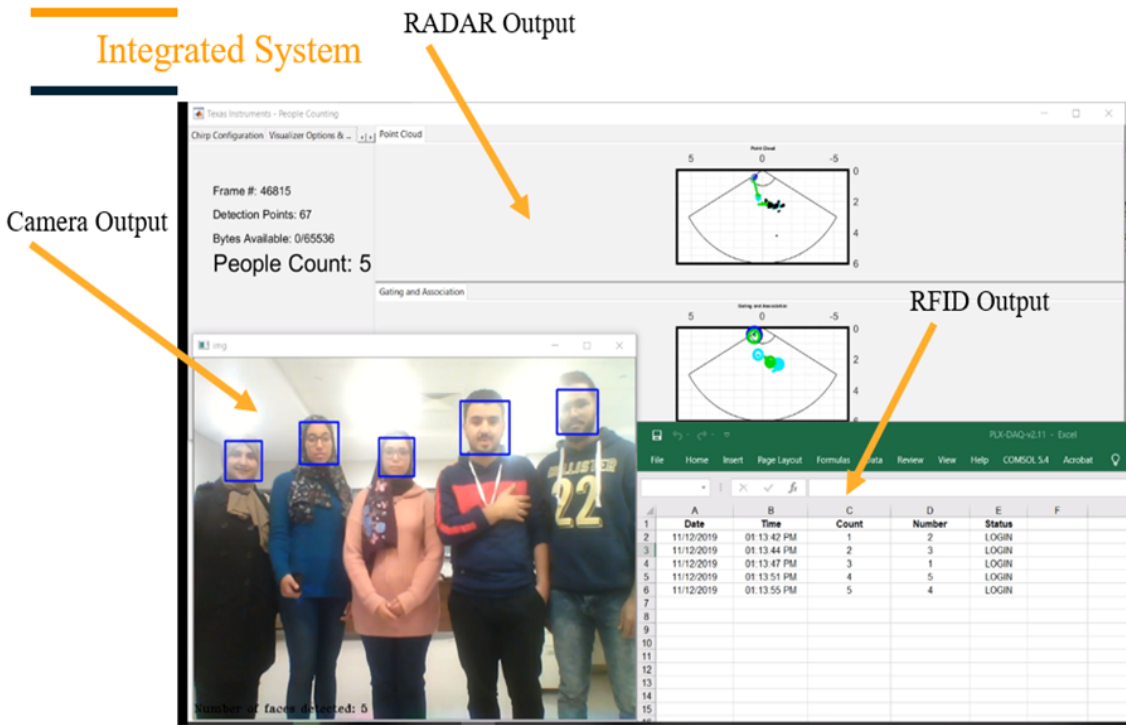


Figure 22. The result of the proposed integrated system.

5. Limitations

The proposed structure faced challenges such as unique detection for one RFID tag, delayed identification and login times in Excel sheets, delayed LCD messages, and simultaneous reading of two RFIDs. These issues were resolved by counting RFID codes and connecting the RFIDs to a bidirectional module. This shared communication bus enables the microcontroller to read the RFIDs continuously, ensuring they are ready for use when the system is turned on.

6. Conclusion

This study introduces a powerful multi-sensor indoor security system that combines UHF RFID, millimeter-wave radar (IWR1642), and a YOLOv5-Tiny-based vision subsystem, enabling continuous surveillance, recognition, and anomaly detection in confined areas such as museums. Unlike typical single-sensor approaches, the proposed framework takes advantage of the complementing characteristics of each modality. RFID enables identification verification and entry/exit auditing, radar allows for continuous motion monitoring even in poor visibility or occlusions, and computer vision gives accurate facial/person recognition. A Kalman Filter-based fusion system synchronizes asynchronous data streams, while Edge-based inference acceleration with TensorRT enables real-time vision inference. Experimental results show that the system is successful in dynamic, low-visibility, and congested situations, with higher detection reliability and fewer false positives than standard techniques. The suggested system addresses major restrictions such as camera occlusion, RFID range constraints, and poor synchronization in previous efforts, resulting in a feasible, scalable solution for current interior security concerns.

While the proposed system was verified in a single room setting, extensive implementations like museums, airports, or multi-room display spaces need meticulous evaluation of computing and communication resources. Scaling the system increases GPU demand for multi-camera systems, necessitating either high-end GPUs or distributed Jetson-class devices for concurrent vision inference. Similarly, multi-radar installations offer larger data rates that must be handled by efficient multi-threaded pipelines, whilst video streaming can surpass 500 Mbps unless compressed or pre-processed at the edge. To address these issues, a modular design is presented in which each monitored room functions as its own node, incorporating local sensors and executing a local fusion pipeline. Only metadata such as occupancy counts, alarms, and anomalies are sent to a central control server, minimising bandwidth and enabling smooth addition or removal of nodes. This approach enhances scalability, fault tolerance, and deployment simplicity across diverse contexts.

The suggested system incorporates privacy-preserving mechanisms to ensure compliance with international legislation, including the General Data Protection Regulation (GDPR), due to the sensitive nature of biometric data. Facial recognition is done on-device, with just anonymised embeddings retained momentarily to prevent central storage of raw face photos. RFID records are connected to pseudonymized IDs rather than personal information.

7. Future Work

In future work, we will work on transforming the model to a fully edge computing model deployment to enable PC-free mobility. Additionally, we will integrate adaptive learning fusion algorithms to adjust the weighting between sensors in response to environmental changes. In addition, we aim to preserve visitors facial recognition privacy with on-device encryption and federated learning. We will test the model's multi-room scalability in big deployments. Future deployments might include federated learning for facial recognition models, which ensures that data never leaves local devices, and encrypted connections between nodes to protect transferred metadata. Furthermore, user permission methods and stringent data retention regulations are required for real-world implementation. These improvements will transform the system into a cutting-edge smart surveillance platform for safety, asset protection, and building management.

Data availability

The data used to support the findings of this study can be freely accessed at <https://drive.google.com/file/d/1ov2XhsYIiWan3oucSrQQJl0PXwMXv8wc/view?usp=sharing>.

REFERENCES

1. A. Sirajdin Olagoke, H. Ibrahim, and S. S. Teoh, "Literature survey on multi-camera system and its application," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3024568.
2. T. I. Amosa, P. Sebastian, L. I. Izhar, O. Ibrahim, L. S. Ayinla, A. A. Bahashwan, A. Bala, and Y. A. Samaila, "Multi-camera multi-object tracking: A review of current trends and future advances," *Neurocomputing*, vol. 552, 2023, doi: 10.1016/j.neucom.2023.126558.
3. D. Khan, Z. Cheng, H. Uchiyama, S. Ali, M. Asshad, and K. Kiyokawa, "Recent advances in vision-based indoor navigation: A systematic literature review," *Computers and Graphics (Pergamon)*, vol. 104, 2022, doi: 10.1016/j.cag.2022.03.005.
4. D. Naukarkar and P. Saktel, "Survey on Object Location Tracking Using RF-Based Tag Detection for Indoor Environment," *IJSR*, vol. 3, no. 11, 2014.
5. K. S. Arikumar, A. Deepak Kumar, T. R. Gadekallu, S. B. Prathiba, and K. Tamarasi, "Real-Time 3D Object Detection and Classification in Autonomous Driving Environment Using 3D LiDAR and Camera Sensors," *Electronics (Switzerland)*, vol. 11, no. 24, 2022, doi: 10.3390/electronics1124203.
6. P. Zhao, C. X. Lu, J. Wang, C. Chen, W. Wang, N. Trigoni, and A. Markham, "Human tracking and identification through a millimeter wave radar," *Ad Hoc Networks*, vol. 116, 2021, doi: 10.1016/j.adhoc.2021.102475.
7. C. Kammel, T. Kogel, M. Gareis, and M. Vossiek, "A Cost-Efficient Hybrid UHF RFID and Odometry-Based Mobile Robot Self-Localization Technique with Centimeter Precision," *IEEE Journal of Radio Frequency Identification*, vol. 6, 2022, doi: 10.1109/JRFID.2022.3186852.
8. J. S. Murthy, G. M. Siddesh, W. C. Lai, B. D. Parameshachari, S. N. Patil, and K. L. Hemalatha, "ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5," *Wireless Communications and Mobile Computing*, vol. 2022, 2022, doi: 10.1155/2022/9444360.
9. S. A. Kashinath, S. A. Mostafa, A. Mustapha, H. Mahdin, D. Lim, M. A. Mahmoud, M. A. Mohammed, B. A. S. Al-Rimy, M. F. M. Fudzee, and T. J. Yang, "Review of data fusion methods for real-time and multi-sensor traffic flow analysis," *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3069770.
10. T. M. A. M. Fadzir, H. Mansor, T. S. Gunawan, and Z. Janin, "Development of School Bus Security System Based on RFID and GSM Technologies for Klang Valley Area," in *Proc. IEEE 5th Int. Conf. Smart Instrumentation, Measurement and Application (ICSIMA)*, 2019, doi: 10.1109/ICSIMA.2018.8688783.
11. B. Dias, A. Mohammad, H. Xu, and P. Tan, "Intelligent Student Attendance Management System Based on RFID Technology," in *Advances in Intelligent Systems and Computing*, 2020, doi: 10.1007/978-3-030-22354-0_51.
12. H. Ouyang, X. Wang, W. Zhu, and Y. Li, "Design of auto-guard system based on RFID and network," in *Proc. Int. Conf. Electric Information and Control Engineering (ICEICE)*, 2011, doi: 10.1109/ICEICE.2011.5777475.
13. M. Khalid, U. Mujahid, and N. U. I. Muhammad, "Ultralightweight RFID Authentication Protocols for Low-Cost Passive RFID Tags," *Security and Communication Networks*, vol. 2019, 2019, doi: 10.1155/2019/3295616.
14. J. C. Yang, C. L. Lai, H. T. Sheu, and J. J. Chen, "An intelligent automated door control system based on a smart camera," *Sensors (Switzerland)*, vol. 13, no. 5, 2013, doi: 10.3390/s130505923.
15. B. Liu, H. Lai, S. Kan, and C. Chan, "Camera-Based Smart Parking System Using Perspective Transformation," *Smart Cities*, vol. 6, no. 2, 2023, doi: 10.3390/smartcities6020056.
16. S. Li, B. Yu, Y. Jin, L. Huang, H. Zhang, and X. Liang, "Image-Based Indoor Localization Using Smartphone Camera," *Wireless Communications and Mobile Computing*, vol. 2021, 2021, doi: 10.1155/2021/3279059.
17. L. Li, J. Yuan, H. Sun, and Y. Song, "An SLAM Algorithm Based on Laser Radar and Vision Fusion with Loop Detection Optimization," *Journal of Physics: Conference Series*, 2022, doi: 10.1088/1742-6596/2203/1/012029.
18. H. Deng, Q. Fu, Q. Quan, K. Yang, and K. Y. Cai, "Indoor Multi-Camera-Based Testbed for 3-D Tracking and Control of UAVs," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, 2020, doi: 10.1109/TIM.2019.2928615.
19. S. Nowok, S. Kueppers, H. Cetinkaya, M. Schroeder, and R. Herschel, "Millimeter wave radar for high resolution 3D near field imaging for robotics and security scans," in *Proc. Int. Radar Symp.*, 2017, doi: 10.23919/IRS.2017.8008132.
20. A. Kanno, N. Yamamoto, R. Takaoka, S. Otani, and H. Sotobayashi, "IMU-Enabled Nondestructive Imaging System Based on Millimeter-Wave Radar," in *Proc. IEEE Int. Conf. Antenna Measurements and Applications (CAMA)*, 2019, doi: 10.1109/CAMA47423.2019.8959578.
21. J. Guan, S. Madani, S. Jog, S. Gupta, and H. Hassanieh, "Through Fog High-Resolution Imaging Using Millimeter Wave Radar," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition (CVPR)*, 2020, doi: 10.1109/CVPR42600.2020.01148.
22. N. Long, K. Wang, R. Cheng, W. Hu, and K. Yang, "Unifying obstacle detection, recognition, and fusion based on millimeter wave radar and RGB-depth sensors for the visually impaired," *Review of Scientific Instruments*, vol. 90, no. 4, 2019, doi: 10.1063/1.5093279.
23. S. Song, J. Wu, S. Zhang, Y. Liu, and S. Yang, "Research on Target Tracking Algorithm Using Millimeter-Wave Radar on Curved Road," *Mathematical Problems in Engineering*, vol. 2020, 2020, doi: 10.1155/2020/3749759.
24. H. Cui and N. Dahnoun, "High precision human detection and tracking using millimeter-wave radars," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 1, 2021, doi: 10.1109/MAES.2020.3021322.

25. X. Huang, H. Cheena, A. Thomas, and J. K. P. Tsoi, "Indoor Detection and Tracking of People Using mmWave Sensor," *Journal of Sensors*, vol. 2021, 2021, doi: 10.1155/2021/6657709.
26. S. Li and R. Hishiyama, "An Indoor People Counting and Tracking System using mmWave sensor and sub-sensors," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 7096–7101, Jan. 2023, doi: 10.1016/j.ifacol.2023.10.577.
27. C. C. Wong, H. M. Feng, and K. L. Kuo, "Multi-Sensor Fusion Simultaneous Localization Mapping Based on Deep Reinforcement Learning and Multi-Model Adaptive Estimation," *Sensors*, vol. 24, no. 1, 2024, doi: 10.3390/s24010048.
28. S. Abba, A. M. Bizi, J. A. Lee, S. Bakouri, and M. L. Crespo, "Real-time object detection, tracking, and monitoring framework for security surveillance systems," *Heliyon*, vol. 10, no. 15, p. e34922, Aug. 2024, doi: 10.1016/j.heliyon.2024.e34922.
29. Y. A. Nasser, M. A. Bazzoun, H. Hijazi, and A. Al Ghouwayel, "Modeling and design of NFC/RFID backbone using a Computer Aided Design tool," in *Proc. 6th Int. Conf. Digital Information Processing and Communications (ICDIPC)*, 2016, doi: 10.1109/ICDIPC.2016.7470811.
30. K. Finkenzeller, *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication*, 3rd ed. Wiley, 2010, doi: 10.1002/9780470665121.
31. A. Ziobro, P. Jankowski-Mihulowicz, M. Weglarski, and P. Pyt, "Investigation of Factors Affecting the Performance of Textronic UHF RFID Transponders," *Sensors*, vol. 23, no. 24, 2023, doi: 10.3390/s23249703.
32. K. A. Kalwar, M. Aamir, and S. Mekhilef, "Inductively coupled power transfer (ICPT) for electric vehicle charging – A review," *Renewable and Sustainable Energy Reviews*, vol. 47, 2015, doi: 10.1016/j.rser.2015.03.040.
33. R. Huang, J. Pedoeem, and C. Chen, "YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers," in *Proc. IEEE Int. Conf. Big Data*, 2018, doi: 10.1109/BigData.2018.8621865.
34. L. Fu, Y. Feng, J. Wu, Z. Liu, F. Gao, Y. Majeed, A. Al-Mallahi, Q. Zhang, R. Li, and Y. Cui, "Fast and accurate detection of kiwifruit in orchard using improved YOLOv3-tiny model," *Precision Agriculture*, vol. 22, no. 3, 2021, doi: 10.1007/s11119-020-09754-y.