

# Real-time emergency response for indoor fires with integrated UAV and multi-sensor platform

Carlos Quiterio Gómez Muñoz<sup>a,\*</sup>, Isaac Segovia Ramirez<sup>a</sup>,  
Guillermo González de Rivera Peces<sup>a</sup>, Pedro Rafael Fernandez Barbosa<sup>b</sup>,  
Maria Cristina Rodriguez-Sanchez<sup>b</sup>

<sup>a</sup> HCTLab Research Group, Universidad Autónoma de Madrid, Madrid, 28049, Spain

<sup>b</sup> Department of Electronics, Rey Juan Carlos University, Madrid, Spain

## ARTICLE INFO

### Index Terms:

Indoor emergency response  
Unmanned aerial vehicles  
Victim detection  
Visible thermal sensing  
Multi-modal data fusion  
Thermal imaging  
GNSS-Denied navigation

## ABSTRACT

This work presents the design, implementation, and experimental validation of an integrated platform based on unmanned aerial vehicles (UAVs) for indoor emergency response in fire scenarios. The system combines visible and thermal sensing with onboard embedded processing, low-latency Wi-Fi communication, and relative navigation based on optical flow and inertial measurements. A multi-process ground station enables real-time reception, visualization, synchronization, and storage of heterogeneous data streams. The communication architecture is designed to support connections from multiple UAVs, and the experiments were conducted in controlled indoor fire tower environments under clear and smoke-affected conditions with one UAV. A comparison between person detection algorithms was performed, in which YOLOv8n achieved a detection rate of 97.65%, a recall of 0.98, and a miss rate of 2.35%.

The sensor ablation study showed that visible sensing was strongly affected by smoke, whereas thermal sensing maintained more stable performance. The visible and thermal fusion strategy reduced the total person counting error under smoke from 0.35 to 0.23 compared with visible only sensing, corresponding to a 34.3% improvement. These results demonstrate the value of combining complementary sensing modalities with a modular communication and data management architecture to support real-time victim detection and situational awareness in constrained indoor emergency environments.

## 1. Introduction

Emergency response technologies play a critical role in mitigating the impact of hazardous incidents, yet significant gaps persist in their effectiveness and efficiency. Despite recent technological advances, there is still a clear need for systems capable of improving the acquisition, analysis, and dissemination of critical information during crisis situations. In this context, the United Nations 2030 Agenda highlights the importance of developing safer, more resilient urban environments, where effective emergency-response mechanisms are essential [1]. With urban populations already comprising approximately 70% of the global total and projected to reach 85% by 2050, the demand for effective emergency response solutions is both pressing and continuous [2]. In the context of smart cities, providing robust mechanisms for emergency management that support rapid recovery after incidents is a key

objective [3,4]. Advancements in emergency services are particularly critical for addressing fire related incidents in confined indoor spaces such as tunnels, galleries, and buildings. A recent study conducted by the Mapfre Foundation in collaboration with the Professional Association of Fire Technicians (APTb) reported a concerning increase in fire related fatalities in Spain, with 165 deaths recorded in 2019, 158 of which were directly attributed to fires. Notably, 97% of these fatalities were associated with direct exposure to fire hazards, with toxic gas inhalation identified as the primary cause [5,6].

The effective detection and management of fire emergencies in enclosed environments remain challenging for prevention and safety teams. Conventional fire detection systems often require close proximity to the fire source and are significantly affected by adverse conditions such as smoke and poor visibility. In addition, misleading visual cues, including reflections from lighting elements, can further reduce system

\* Corresponding author.

E-mail address: [Carlosq.gomez@uam.es](mailto:Carlosq.gomez@uam.es) (C.Q. Gómez Muñoz).

<https://doi.org/10.1016/j.array.2026.101123>

Received 18 July 2025; Received in revised form 22 July 2026; Accepted 31 July 2026

Available online 2 August 2026

2590-0056/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

reliability and delay timely intervention [7,8].

In recent years, UAVs have gained increasing relevance in public safety and emergency response applications, including search and rescue operations, disaster assessment, and surveillance [9]. Their ability to rapidly access hard to reach or hazardous areas enables enhanced situational awareness for responders operating on the ground. Despite these advantages, the effective deployment of UAVs in emergency scenarios remains challenging. Limitations related to flight endurance, payload capacity, indoor operation, and integration with existing response infrastructures continue to restrict their full potential, particularly in complex indoor environments.

Addressing these challenges requires a comprehensive system level approach. Emergency operations demand real-time acquisition of critical information, including environmental conditions and visual data from hazardous indoor environments, while simultaneously operating under severe communication and positioning constraints, as conventional infrastructures and Global Navigation Satellite System signals are often unavailable indoors [6,10].

Recent studies, such as the work by Roldán-Gómez et al. [11] have highlighted the potential of UAV-based platforms to enhance situational awareness in such environments. However, many existing solutions primarily focus on autonomous navigation or mapping, while less attention is paid to the communication and data management requirements that are crucial for time critical emergency response.

In this context, the objective of this work is to design and experimentally validate an integrated UAV-based platform for indoor emergency response. The proposed system focuses on robust communication, multi-sensor integration, and real-time situational awareness in GNSS denied environments. Rather than targeting full autonomy, the approach prioritizes rapid deployment and reliable information delivery, providing a modular and extensible foundation for future developments.

## 2. State of the art

Mitigating fire related risks in complex environments requires timely and reliable situational awareness, rather than relying solely on long term predictive forecasting. Satellite based solutions, such as the approach proposed by Fukuhara et al. [12], have been explored for detecting fire hot spots at large scales. However, these methods rely on successive scans over wide areas, resulting in significant temporal delays, and are inherently limited in dense urban settings. While effective for monitoring forests or extensive land areas, satellite systems are poorly suited for indoor or subterranean environments, where critical hot spots may remain undetected.

As a result, there is a clear need for complementary approaches capable of providing real-time sensing and decision support in indoor emergency scenarios. In such contexts, rapid data acquisition, low-latency communication, and tight integration between sensing and processing components are essential to support effective intervention under highly dynamic and uncertain conditions. The present work addresses these requirements by focusing on sensor-based data acquisition and system integration tailored to indoor environments.

Smart Cities increasingly rely on data-driven approaches supported by distributed sensing and Internet of Things (IoT) technologies to improve monitoring, prevention, and response processes in complex infrastructures such as electrical substations, large buildings, and industrial facilities [13]. The availability of large volumes of heterogeneous data introduces new challenges related to interpretation and decision-making, making effective visualization techniques essential for hazard assessment, monitoring, and resource optimization [14]. In emergency scenarios characterized by high uncertainty, such as indoor fires that alter spatial layouts or severely limit visibility due to smoke, accurate perception and localization of robotic platforms become critical to support effective human intervention [15]. The work [16] highlights the growing role of sensors, wireless sensor networks, IoT

devices, and UAV-based systems in disaster management and search and rescue operations, emphasizing the need for integrated platforms capable of operating under dynamic and constrained conditions.

### 2.1. UAV applications for monitoring and emergency situations

UAVs have been widely adopted for monitoring and inspection tasks due to their mobility, flexibility, and ability to access hazardous or confined environments. In safety-critical applications, UAVs are commonly used for structural condition monitoring of buildings, bridges, tunnels, and industrial facilities, where high-resolution visual and sensor-based inspections can be performed without exposing human operators to risk [17–20]. UAV-based monitoring has also been explored in large-scale sensing applications using multispectral and thermal imagery for automated anomaly detection and decision support, demonstrating the versatility of aerial sensing platforms across different operational domains [21,22].

UAVs are also extensively employed for visual inspections of hard-to-reach infrastructures such as power lines [23,24], wind turbines [25, 26], and industrial equipment [27–29]. Their maneuverability enables detailed assessment of potential defects, degradation, or abnormal conditions in scenarios where traditional inspection methods are costly, time-consuming, or dangerous.

In the context of emergency response, UAV-based monitoring systems have demonstrated significant potential for supporting situational awareness, damage assessment, and decision-making in dynamic and uncertain environments [30–32]. These applications highlight the suitability of UAV platforms for rapid deployment and real-time data acquisition in safety and emergency-related scenarios, particularly when combined with onboard sensing and ground-based data processing systems.

### 2.2. Emergency situations

UAVs provide rapid access to affected areas and enable real-time data acquisition, making them valuable assets for emergency management and response operations. Their capability to assess hazardous or inaccessible environments, support victim localization, and deliver situational awareness information to first responders has motivated their increasing adoption in emergency scenarios.

Existing applications include time-critical interventions such as the delivery of automated external defibrillators and flotation devices. In addition, UAVs equipped with onboard sensors have been deployed in hazardous contexts, including wildfires and chemical spill incidents, where they support emergency response activities while reducing the exposure of human operators to dangerous conditions [33–36].

In emergency response scenarios, UAV deployments typically support several key functions: (i) fire detection and monitoring, using infrared or thermal imaging to identify hot spots and track fire evolution; (ii) search and rescue, leveraging optical and thermal sensing for victim localization and, when needed, limited payload delivery of essential supplies; (iii) post-disaster assessment, enabling rapid mapping and inspection of damaged structures to support situational planning; (iv) hazard detection, using gas or environmental sensors to detect toxic substances and air contaminants; and (v) communication support, acting as temporary relay nodes or providing connectivity assistance when conventional infrastructure is degraded.

In emergency scenarios, particularly indoor fires, the primary objectives include assessing building conditions, supporting firefighting operations, and rescuing potential victims. In this context, UAVs can contribute by providing timely situational awareness, identifying fire sources and hazardous areas, and supporting the planning of safer intervention routes for emergency teams.

The literature reports a wide range of robotic solutions for firefighting applications, including fire extinguishing robots [37], fire detection sensors [38], and robot-assisted evacuation [39]. These

technologies play a crucial role in acquiring real-time information that supports human decision-making during emergency interventions.

The survey conducted by Roldan et al. [11] indicates that Spanish firefighters identify health risks and the lack of real-time information on fire evolution as the most critical challenges during firefighting operations, surpassing limitations related to personnel, equipment, or coordination. These difficulties are further exacerbated in indoor environments, where limited ventilation leads to the accumulation of hazardous gases and restricted mobility, significantly increasing the complexity and risk of emergency response operations.

Autonomous robotic systems, including UAVs, play an increasingly important role in search and rescue operations following natural or human-made disasters [40–42] as well as during emergency interventions such as chemical, biological, radiological, and nuclear (CBRN) incidents and wildfires [43,44]. In parallel, the development of realistic simulation environments for robotic platforms and mission scenarios has been shown to support intervention planning and training, particularly when combined with data analysis, artificial intelligence, and predictive modelling techniques [45]. Several studies have specifically addressed the use of UAVs in indoor emergency missions, exploiting their agility to navigate confined spaces while accounting for their structural fragility and limited endurance [46]. In addition, cooperative strategies involving aerial and ground robots have been proposed to improve mission performance and system resilience in complex environments [47]. Recent works have demonstrated the feasibility of UAV-based sensing and perception in indoor fire scenarios. Aláez et al. [48] presented a micro-UAV equipped with an embedded thermal camera and FPV system for real-time visual and thermal situational awareness in structural fires, emphasizing hardware adaptations to improve thermal robustness and endurance. Do et al. [49] proposed an autonomous indoor UAV system combining obstacle avoidance and a You Only Look Once (YOLO) based people detection pipeline, achieving high navigation accuracy and effective victim localization in GPS-denied environments. Boiteau et al. [50] presented an autonomous UAV for target detection in GNSS-denied environments. This system incorporated advanced positioning system based on LiDAR and Inertial Measurement Unit (IMU), although the accuracy of fault detection presented several limitations.

While these approaches focus primarily on autonomous navigation and onboard perception, the present work emphasizes the design of a robust communication and data management platform to support real-time situational awareness in indoor emergency scenarios. Table 1

summarizes a comparison between the proposed approach and representative UAV-based indoor systems reported in the literature.

Based on the comparative analysis presented in Table 1 and the discussion of main limitations and advantages of this approach, several system-level limitations can be identified in existing UAV-based approaches for indoor emergency scenarios. Table 2 summarizes the main gaps observed in the literature to provide a clearer interpretation of the contribution of the proposed system.

This proposal is a comprehensive approach to fire detection, developing a user-centered platform programmed for indoor scenarios. The main novelties are summarized as follows:

- This platform addresses indoor fire scenarios through a UAV based approach whose design, development, and testing were conducted with direct input from professional firefighters, going beyond studies that address only partial aspects of the emergency-response problem.
- The design of a modular ground-station architecture capable of managing concurrent UAV connections and multi-stream data in real-time. The practical integration of visual, thermal, and environmental sensing into a unified acquisition, visualization and logging pipeline, enables real-time awareness.
- The experimental validation of this architecture in an indoor fire-related scenario in a controlled fire-tower environment,

**Table 2**

Analysis of system-level limitations and contributions of the proposed architecture.

| Limitations identified in previous research                        | Improvements introduced in this work                                                            |
|--------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Limited use of thermal sensing in indoor UAV systems               | Integration of visible and thermal perception for operation under reduced visibility            |
| Limited discussion of communication architecture                   | Modular ground-station architecture with decoupled data reception                               |
| Limited system-level validation in fire-related scenarios          | Platform validated in an indoor fire-emergency scenario                                         |
| Complex protocols that introduce latency                           | Lightweight communication design optimized for low-latency real-time data transmission          |
| Limited positioning capabilities in GPS-denied indoor environments | Integration of optical flow-based relative positioning for improved indoor navigation stability |

**Table 1**

Comparison with UAV-based indoor emergency-response studies.

|                                 | Do et al. [49]         | Boiteau [50]                 | Kucukayan and Karacan [51]              | This work                                                                 |
|---------------------------------|------------------------|------------------------------|-----------------------------------------|---------------------------------------------------------------------------|
| Thermal imaging                 | ×                      | ✓                            | ×                                       | ✓                                                                         |
| Environmental sensor            | ✓                      | ✓                            | ×                                       | ✓                                                                         |
| Indoor operation for GPS-denied | ✓                      | ✓                            | ✓                                       | ✓                                                                         |
| Real-time data                  | ✓                      | POMPD and onboard processing | ✓                                       | ✓                                                                         |
| Ground station architecture     | Defined                | Telemetry radio              | ✓                                       | ✓                                                                         |
| Hazard detection latency        | 15.40–16.86 ms         | Not defined                  | 37 ms/frame                             | Operates under low visibility                                             |
| Data transmission               | Not defined            | MAVLink CRC-16               | RTSP over UDP (Non-acknowledged socket) | Non-acknowledged socket transmission for real-time MJPEG video streaming. |
| Victim detection                | Visual YOLOv5n/8n      | Thermal YOLO v5              | Visual YOLO IHD                         | Visual + thermal                                                          |
| Detection accuracy              | 93%                    | Precision: 89%               | Not defined                             | 98%                                                                       |
| FPS/framerate                   | 24 FPS                 | Not defined                  | 7                                       | Variable                                                                  |
| End-user-driven system          | ×                      | ×                            | ×                                       | ✓                                                                         |
| Experiments                     | Real                   | Simulation + real            | Real                                    | Indoor fire-emergency scenario                                            |
| Inference latency               | 31,9 ms per prediction | Not reported                 | 2–5 ms                                  | Low latency                                                               |
| Positioning method              | Real-Time Kinematic    | LiDAR + IMU                  | Not defined                             | Relative optical flow sensor                                              |
| Navigation accuracy             | 91%                    | 96%                          | Not evaluated                           | 94%                                                                       |
| Navigation-related error        | 4.3% displacement      | 0,1 m                        | Not defined                             | 0,15 m, 6% displacement                                                   |

demonstrating real-time operation and data synchronization, under reduced visibility conditions.

### 3. Methodology

#### 3.1. System architecture

The proposed system is designed as a distributed architecture for UAV-based indoor emergency response, integrating aerial sensing platforms with a ground-based computational station. The objective is to enable reliable real-time acquisition, transmission, and processing of heterogeneous data under constrained indoor conditions.

As illustrated in Fig. 1, the system is structured into two main subsystems: (i) the aerial subsystem, responsible for sensing and onboard data acquisition, and (ii) the ground subsystem, responsible for communication management, data processing, visualization, and storage.

The aerial subsystem incorporates sensing devices and embedded processing units that collect multi-modal data and transmit them to the ground station. The ground subsystem acts as a central coordination node, managing multiple UAV connections and enabling real-time situational awareness for emergency operators.

This separation of responsibilities ensures modularity and scalability, allowing the system to support different operational configurations, including single- and multi-UAV deployments. The overall architecture is designed to operate under indoor constraints, where communication

reliability, limited payload capacity, and the absence of GNSS signals must be carefully addressed.

#### 3.2. Aerial subsystem

##### 3.2.1. UAV design and development

The aerial subsystem is built around a custom-designed UAV platform tailored for indoor emergency operations. The system adopts a quadrotor configuration due to its stability, maneuverability, and suitability for confined environments. As shown in Fig. 2, the UAV integrates the main structural, sensing, and control components required for safe operation.

The platform operates in a semiautonomous mode, where flight control can be performed manually by an operator or through pre-defined trajectories, while data acquisition and transmission are fully automated.

To ensure safe operation in indoor environments, the airframe incorporates a reinforced structure with propeller protection, reducing the risk of damage during collisions. Given the absence of GNSS signals indoors, the UAV is equipped with onboard sensors that support relative positioning and altitude estimation.

The main characteristics of the UAV platform include:

- Maximum takeoff weight: 1860 g.
- Maximum payload capacity: 250 g.
- Quadrotor configuration for indoor operation.

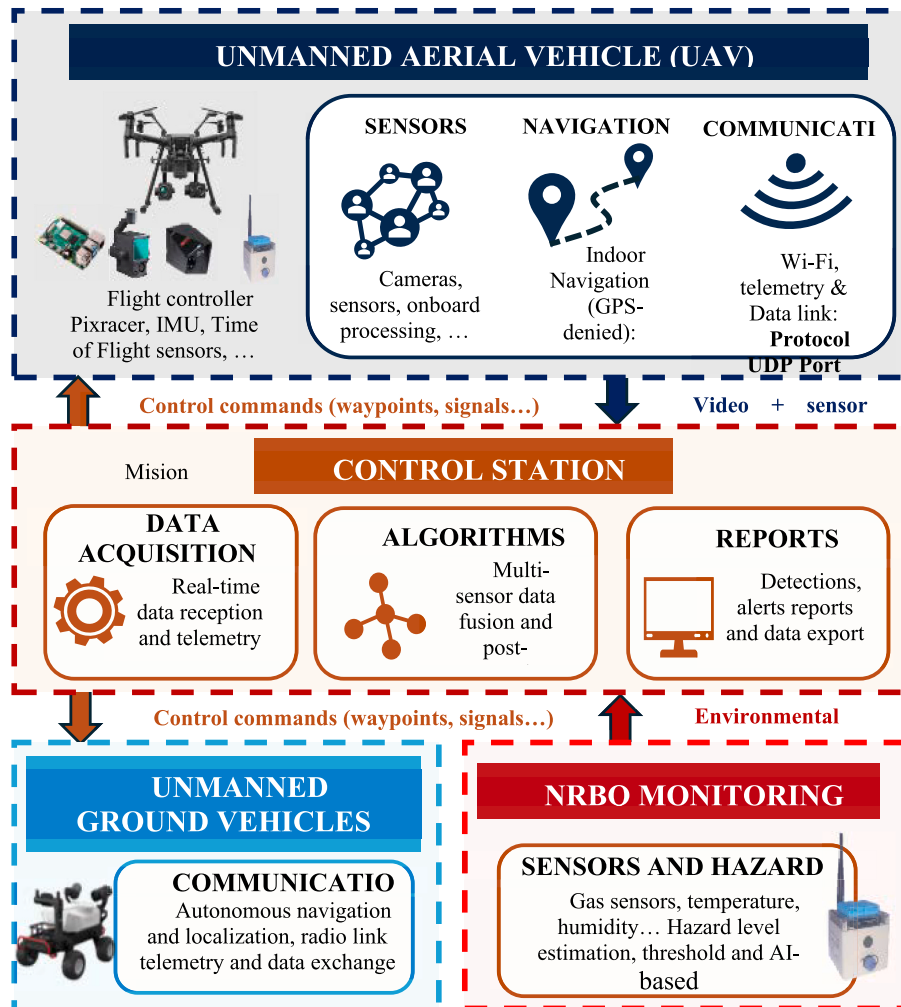
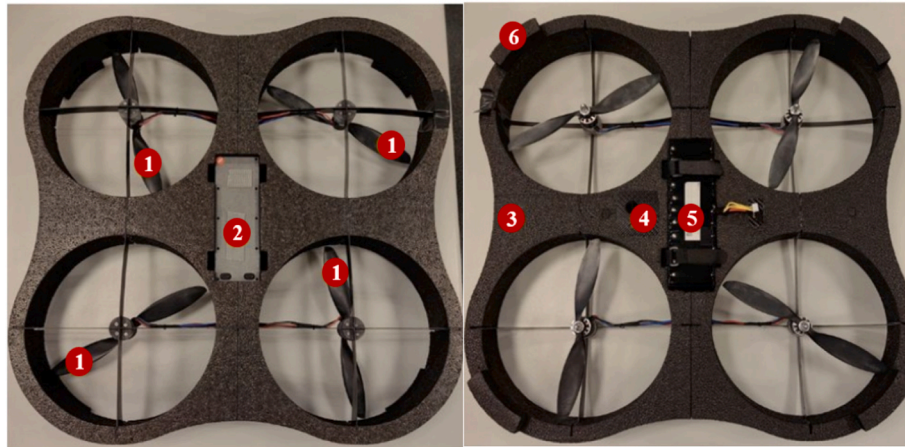


Fig. 1. Distributed architecture of the UAV based indoor emergency response system, showing the aerial sensing subsystem and the ground station communication, processing, visualization, and storage components.



**Fig. 2.** UAV scheme. 1) Propellers; 2) Autopilot; 3) Hull; 4) Optical flow sensor and infrared altitude sensor; 5) Battery; 6) Landing gear.

- Optical flow sensor for relative motion estimation and infrared-based altitude sensing.
- Ultrasonic transducer.
- Lithium-ion battery pack (14.8 V, 10 Ah).
- Wireless telemetry and data transmission via a dedicated Wi-Fi link.
- Onboard visual camera for real-time video streaming.
- Reinforced structure with propeller protection for indoor safety.

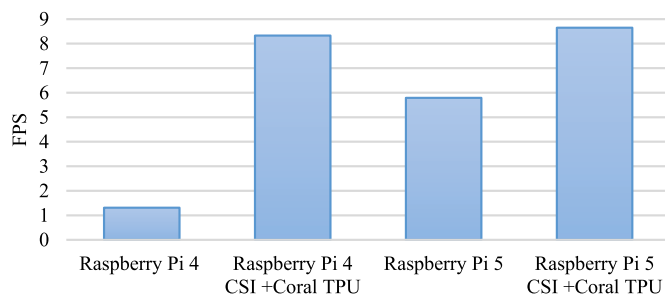
### 3.2.2. Embedded processing architecture

Following the platform design, the onboard processing architecture adopts a heterogeneous embedded computing approach, where multiple devices are assigned specific tasks according to their computational requirements and hardware constraints.

Three main platforms were considered: Raspberry Pi 4, Raspberry Pi 5, and Raspberry Pi Zero 2W. The Raspberry Pi 4 and Raspberry Pi 5 were evaluated for computationally demanding tasks, such as real-time image processing and AI-based inference. Among them, the Raspberry Pi 5 demonstrated superior performance, particularly when combined with hardware acceleration using a Google Coral TPU, as shown in Fig. 3.

The Raspberry Pi Zero 2W is employed for less demanding tasks, including auxiliary image acquisition (e.g., low-light cameras) and basic data forwarding. Its reduced size and power consumption make it suitable for minimizing payload impact, although it introduces bandwidth constraints due to its limited communication capabilities.

This heterogeneous architecture enables an effective balance between computational performance and payload limitations. High demand processing tasks are assigned to more capable embedded units, while lightweight nodes handle secondary sensing and data transmission. This design enhances system flexibility and supports scalable configurations, including the distribution of sensing modules across onboard devices or multiple UAVs.



**Fig. 3.** Frame-rate comparison of Raspberry Pi 4 and Raspberry Pi 5 configurations with and without Google Coral TPU acceleration.

### 3.2.3. Multi-modal data acquisition

The aerial subsystem is designed to acquire and transmit heterogeneous data streams that provide a comprehensive representation of indoor emergency environments. As shown in Fig. 4, the UAV integrates multiple sensing devices and onboard processing units, enabling the concurrent acquisition of different data modalities.

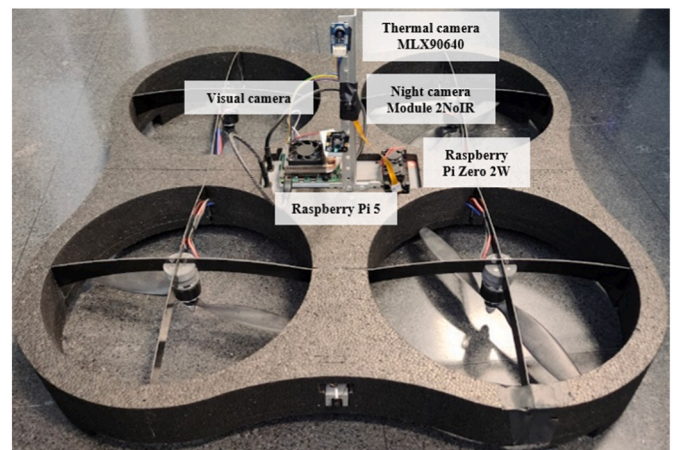
The acquired data are organized into three main categories:

- Visual data, including RGB and low-light imagery
- Thermal data, providing temperature distribution information
- Environmental sensor data, capturing variables such as gas presence or ambient conditions

This multi-modal configuration enables complementary perception of the environment, supporting the detection of visibility conditions, heat sources, and potential hazards in real-time.

In line with the heterogeneous embedded architecture, sensing tasks are distributed across onboard computing units according to their computational requirements. High-resolution image acquisition is assigned to more capable platforms, while lightweight devices handle auxiliary sensing tasks.

To support real-time operation under bandwidth constraints, visual data streams are encoded using the Motion JPEG (MJPEG) format. In the implemented configuration, a resolution of 1280×720 pixels at 30 frames per second was selected as a tradeoff between image quality and



**Fig. 4.** UAV prototype integrating visible and thermal cameras with Raspberry Pi 5 and Raspberry Pi Zero 2W onboard processing units for multi-modal data acquisition.

transmission efficiency.

### 3.2.4. Indoor navigation

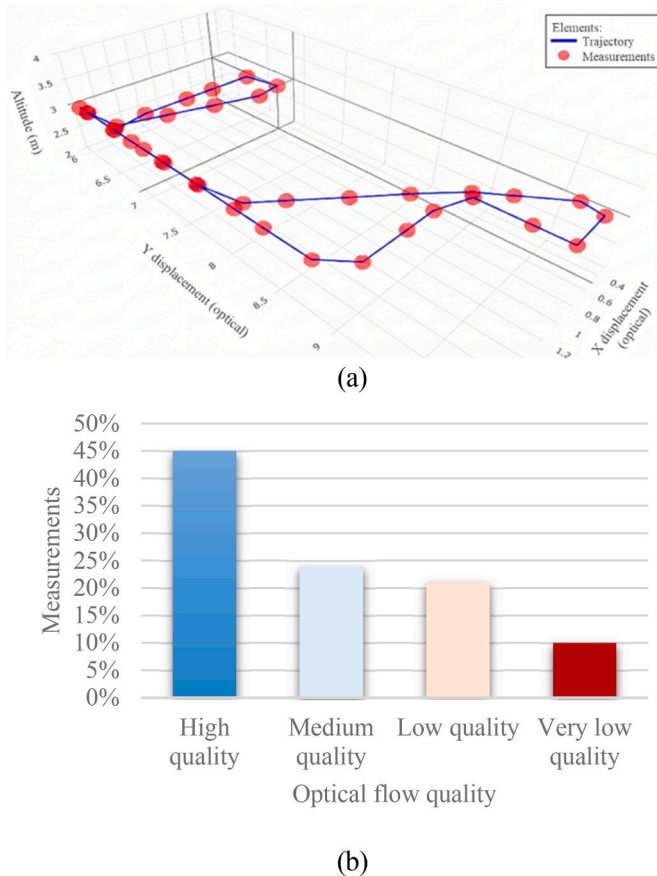
The proposed navigation strategy is based on relative positioning, enabling UAV operation in GNSS-denied indoor environments without relying on full mapping or localization frameworks such as SLAM.

The system integrates an IMU and a downward facing optical flow sensor to estimate relative motion in the horizontal plane. The optical flow sensor tracks ground texture variations to compute incremental displacements, while a quality metric is used to assess the reliability of each measurement. Altitude estimation is obtained through a combination of barometric and range based sensing, improving robustness in indoor conditions.

At the beginning of each mission, a reference position is defined, and all subsequent movements are computed relative to this origin. This approach enables trajectory recording during manual flight and subsequent autonomous replay using stored relative displacements. Fig. 5(a) represents the performance of the optical flow sensor along the path.

The optical flow quality parameter shown in Fig. 5 (b) reflects the reliability of image-based texture tracking. High-quality measurements are obtained when the sensor observes well defined visual features, while low quality values are recorded in the case of low contrast, changes in lighting, reflections or smoke, among others. It is demonstrated that the optical sensor has proper reliability, because only 10% of its measurements could cause issues along the path. The estimated error in these cases is around 0.15 m in the horizontal plane, corresponding to an error of approximately 6%.

This approach enables a two-stage operation: first, a manually



**Fig. 5.** Indoor trajectory reconstruction and optical-flow measurement quality: (a) recorded three-dimensional path based on relative optical-flow displacement and altitude estimation; (b) distribution of optical-flow measurements according to quality level.

piloted flight during which a route is recorded, and second, an autonomous replay of the same trajectory using the stored relative displacements. This strategy is suitable for controlled or semi-structured scenarios, because the current system does not incorporate obstacle detection, avoidance or dynamic path planning. The system is designed to operate under human supervision, allowing real-time intervention and control adjustments, although its robustness may be affected under challenging environmental conditions. The performance of the optical flow sensor depends on the presence of sufficient visual features and adequate visibility. In environments with dense smoke, low texture, or poor lighting conditions, tracking performance may degrade, potentially affecting horizontal positioning accuracy. In such cases, the system continues to provide partial situational awareness through onboard sensing and real-time data transmission, even if navigation performance is reduced.

Despite these limitations, the proposed platform prioritizes reliable communication, multi-sensor data acquisition, and real-time monitoring, which are essential for supporting decision-making in emergency scenarios.

### 3.3. Communication framework

The proposed system relies on an integrated communication framework that enables the acquisition, transmission, and processing of heterogeneous data streams between the UAV and the ground station in real-time. This framework encompasses the onboard data generation processes, the communication layer, and the ground-side data handling mechanisms.

#### 3.3.1. Onboard data acquisition and transmission

Once the onboard sensing and processing components are defined, the UAV-side software manages data acquisition and transmission from the embedded computing units. The system follows a modular multi-process architecture, where independent Python processes are executed in parallel, each assigned to a specific data stream, such as visual, thermal, or environmental sensor data.

This design provides fault isolation and operational flexibility, allowing selective activation of sensing modules depending on mission requirements. Although all processes share a common execution logic, they differ in their acquisition interfaces, data formatting procedures, and transmission routines.

As illustrated in Fig. 6, each onboard process follows a sequential workflow: communication is first established with the ground station, sensor or image data are acquired, the corresponding data packet is transmitted, and the process continues while new data are available.

To support real-time video transmission, an optimized acquisition pipeline is implemented. An optimized buffering strategy is used to directly stream captured frames as byte sequences, avoiding intermediate storage and reducing latency. The video acquisition process allows configuration of both spatial resolution and frame rate. In the implemented setup, a resolution of 1280×720 pixels at 30 frames per second was selected as a trade-off between image quality and bandwidth constraints. To further improve transmission efficiency, the video stream is encoded using the Motion JPEG (MJPEG) format. This approach reduces packet size while maintaining sufficient image quality for real-time visualization and downstream processing.

The system relies on continuous high-frequency streaming, with image data transmitted at up to 30 FPS, providing updated information approximately every 33 ms. The loss of individual packets does not affect subsequent frames, allowing the system to maintain continuous visualization despite occasional transmission errors.

The experimental results indicate that packet loss is negligible under stable link conditions but becomes variable as the communication link approaches its degradation threshold, reaching a maximum observed value of 25% when several obstacles separated the transmitter and receiver, and reached up to 45% in highly obstructed conditions close to

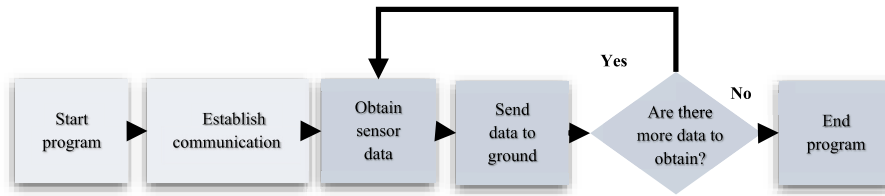


Fig. 6. General flowchart of drone microcomputer processes.

the link degradation threshold. This reduction in reliability can affect the continuity and timeliness of the information received at the ground station. Permanent loss of recorded information is avoided through loop transmission using local storage and store-and-forward transmission after reconnection. The communication tests were conducted under controlled laboratory conditions, where basic tests involving smoke exposure did not reveal a noticeable impact on data transmission. Nevertheless, future work will expand these tests to denser smoke conditions and stronger interference scenarios.

### 3.3.2. Ground-side data management

The ground subsystem acts as the central processing and coordination unit, responsible for receiving, processing, visualizing, and storing the data transmitted by the UAVs.

The ground software follows a multi-process architecture in which a dedicated process is instantiated for each connected UAV. Each process is dynamically instantiated upon connection establishment and configured according to the data streams advertised by the corresponding UAV. This enables parallel and independent handling of multiple data streams, ensuring scalability and robustness in multi-UAV scenarios.

Each process manages the full data pipeline associated with a specific UAV, including data reception, processing, storage, and visualization. This approach ensures fault isolation and prevents failures in one connection from affecting the rest of the system. As illustrated in Fig. 7, the ground system follows a structured workflow in which incoming data are received, processed, stored, and displayed in real-time.

### 3.3.3. Communication protocol and interaction

Communication between the UAV and the ground station is established through a direct Wi-Fi link. A lightweight discovery-based mechanism is used to dynamically detect available UAVs and establish communication channels.

During the discovery phase, the ground station broadcasts connection parameters such as IP address, communication port, and streaming configuration, to which UAV nodes respond by advertising their available data streams.

The system avoids conventional handshake and acknowledgment-based protocols to meet real-time constraints, prioritizing low latency over guaranteed delivery. Once a connection is established, communication proceeds through continuous data streaming.

Each UAV is handled through an independent communication process at the ground station, enabling simultaneous and decoupled data reception.

The interaction between the UAV onboard systems and the ground station during connection establishment and data transmission is

illustrated in Fig. 8, which presents a sequence diagram of the communication workflow.

### 3.4. Data processing and management

#### 3.4.1. Data reception and pre-processing

After data transmission has been established through the communication framework, the ground station processes incoming data streams from each UAV in real-time. A dedicated process is assigned to each active connection, continuously receiving and handling incoming data.

The communication design is defined by a continuous data streaming architecture, where sensor measurements and image data are transmitted at high frequency. This approach introduces temporal redundancy, ensuring that eventual packet loss does not affect system performance, as updated information is rapidly received and replaces missing data. Upon reception, raw byte streams are parsed and converted into structured representations suitable for further processing. Image data are reconstructed into numerical arrays, while sensor measurements are decoded and formatted into JSON structures.

As illustrated in Fig. 9, the processing workflow follows dual-path architecture. First, the processed data are pushed into a consumption queue for downstream modules. In parallel, the same data are used to update the graphical interface, enabling continuous real-time visualization. This pipeline operates concurrently across all active UAV connections, ensuring synchronized and efficient handling of multi-source data streams.

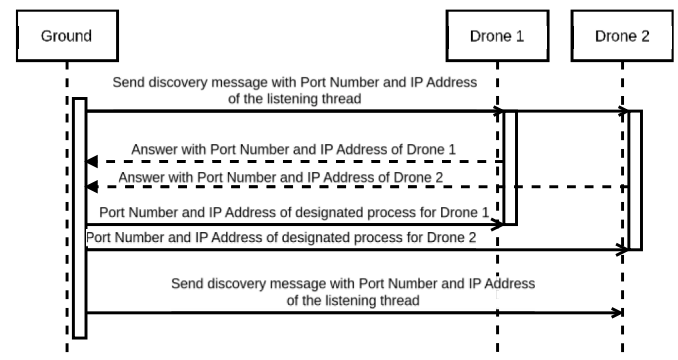


Fig. 8. Sequence diagram of communication between ground and embedded microcomputers in drones.

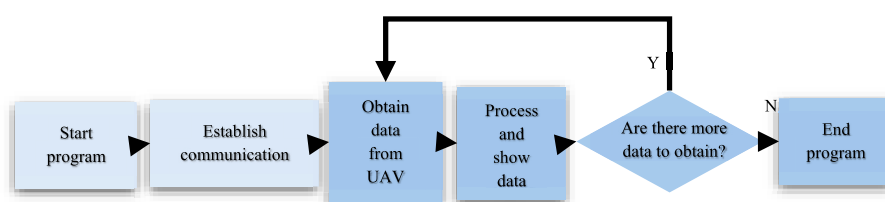


Fig. 7. General Flowchart of ground-based computer Processes.

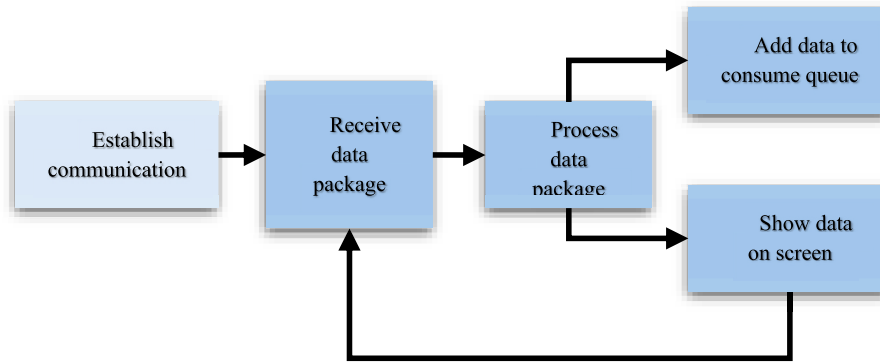


Fig. 9. Flowchart of data reception and processing from UAVs.

### 3.4.2. Multi-modal synchronization

Sensor data are acquired by multiple onboard processing units, each assigning a timestamp at capture time to every frame or measurement. This timestamping mechanism provides a common temporal reference across heterogeneous sensors, including visual, thermal, and environmental sources. A time-window-based synchronization strategy is applied to compensate for differences in sampling rates and transmission latencies, enabling coherent alignment of multi-modal data streams.

To address the heterogeneous nature of the sensors and the real-time operational constraints of indoor emergency scenarios, the system adopts a decision-level integration approach. In this framework, visual observations, thermal patterns, and environmental measurements are interpreted independently and subsequently combined to support situational awareness and alert generation, rather than relying on early-stage data fusion.

### 3.4.3. Visualization

Data visualization constitutes the final stage of the processing pipeline associated with each UAV connection at the ground station, enabling operators to monitor system status and interpret incoming information in real-time.

For visual data streams, received image frames are decoded and converted into numerical representations suitable for processing and display. As illustrated in Fig. 10, this stage supports both raw image visualization and the overlay of detection outputs when image recognition is enabled, allowing operators to simultaneously assess scene context and extracted semantic information.

In parallel, sensor data streams follow a lightweight processing pipeline, as measurements are already structured in a standardized JSON format upon reception. These data are parsed and directly forwarded to the visualization module, where dynamic plots represent the temporal evolution of sensor values.

This dual visualization approach enables simultaneous monitoring of visual and environmental conditions, providing a comprehensive and coherent representation of the operational scenario and supporting rapid decision-making in emergency contexts.

### 3.4.4. Data storage

The system incorporates a persistent storage mechanism to ensure that all acquired data can be recorded and subsequently analyzed. Data storage operates concurrently with the processing pipeline, enabling continuous recording without interfering with real-time system performance.

For visual data, captured image frames are stored in compressed format and later aggregated into video sequences associated with each UAV connection. This organization facilitates efficient retrieval and post-mission analysis while preserving the temporal structure of the acquired data.

Sensor and thermal data are handled through a structured storage approach based on a lightweight relational database. Each incoming measurement is recorded along with its corresponding timestamp and data type, enabling efficient querying and categorization of heterogeneous data streams.

SQLite was selected as the database management system due to its low resource requirements, portability, and ease of integration within embedded and portable computing environments. The database scheme is designed to support clear separation between different data modalities, including environmental sensors, thermal measurements, and visual data references. This storage architecture enables reliable data persistence while maintaining compatibility with real-time system constraints, supporting both live operation and offline analysis of emergency scenarios.

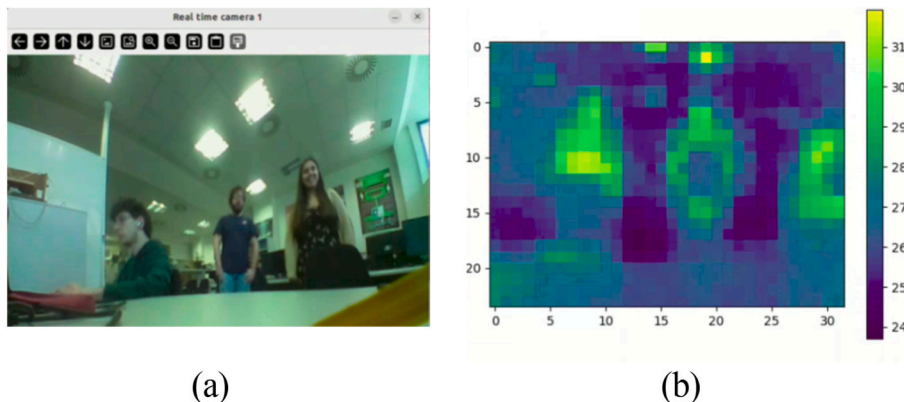


Fig. 10. Preliminary tests with streaming cameras (a) Visible Light Camera and (b) Thermal Sensor.

### 3.4.5. Detection algorithms

To enable real-time person detection within the proposed system, different image recognition strategies were evaluated, considering the strict latency constraints associated with emergency response scenarios. In this context, achieving a balance between detection accuracy and computational efficiency is essential.

Two primary approaches were considered: a deep learning-based detector and a classical feature-based method. For the deep learning approach, the YOLO architecture was selected due to its proven performance in real-time object detection tasks [52]. In particular, YOLOv8n was adopted, as it provides a favorable trade-off between model size, inference speed, and detection accuracy, making it suitable for deployment on embedded and resource-constrained platforms.

The model was fine-tuned using a lightweight training configuration designed to ensure stable convergence and adequate generalization while maintaining computational efficiency. This setup allows the model to perform robust person detection under varying environmental conditions, including reduced visibility scenarios.

As a baseline, the Histogram of Oriented Gradients (HOG) method combined with a Support Vector Machine (SVM) classifier was also implemented. HOG represents objects based on local gradient orientation distributions, effectively capturing structural and shape-related features. Although less flexible than deep learning approaches, it remains computationally efficient and interpretable, making it a suitable reference for comparison.

For person detection using HOG, a pre-trained pedestrian detector was employed. Incoming frames are first converted to grayscale to reduce computational complexity, after which a multi-scale sliding-window approach is applied to identify candidate regions. Detected instances are enclosed by bounding boxes and forwarded to the visualization module. This dual-approach strategy enables comparative evaluation between classical and deep learning methods, while ensuring that real-time operational constraints are satisfied.

The proposed methodology integrates distributed onboard sensing, low-latency communication, and real-time ground-based processing into a unified framework for indoor emergency response scenarios.

## 4. Results

A series of experiments were conducted under controlled indoor conditions designed to emulate emergency scenarios and to evaluate the performance of the proposed system.

The experimental setup used one UAV equipped with visible and thermal sensing capabilities, together with onboard processing and communication modules. The acquired data streams were transmitted through a dedicated Wi-Fi network to a centralized ground station for real-time processing, visualization, and storage. Although the communication architecture is designed to support independent connections from multiple UAVs, the experimental results reported in this work were obtained using data acquired from a single UAV.

Different operational conditions were considered during evaluation, including both clear visibility and degraded environments, in order to assess system robustness under realistic emergency scenarios. The evaluation focuses on the system's ability to maintain stable communication, support real-time data processing, and provide meaningful situational awareness through integrated multi-modal sensing.

However, several technical limitations have been identified, particularly concerning the payload capacity of the drone. To address these constraints, different strategies were explored, including the use of lightweight imaging devices and the integration of Raspberry Pi Zero boards to reduce overall system weight, as well as adjustments to the power supply to support these components.

Despite these optimizations, communication performance remained constrained by bandwidth limitations. In particular, the use of Raspberry Pi Zero devices, which only support 2.4 GHz Wi-Fi networks, restricted the achievable data rate and affected the stability of real-time

video transmission. To overcome these limitations, Raspberry Pi 5 devices were selected for video transmission, as they support dual-band 2.4/5 GHz Wi-Fi and provide improved processing capability. This enables higher transmission speeds and more stable communication, which are critical for real-time streaming in emergency-response scenarios.

The selection of Python as the programming language introduced additional limitations due to the Global Interpreter Lock. This affected the parallel execution of threads and, consequently, the computational efficiency of the system. This was partially addressed through the multiprocessing library, which albeit allowed for concurrent function execution, imposing its own set of limitations on process parallelization.

Additionally, a test was conducted in two different fire towers using a mannequin to simulate a non-heat-emitting person, while different equipment setups are considered to represent typical items carried by firefighters or rescue personnel. Table 3 presents a quantitative comparison between the classical HOG-SVM detector and the YOLOv8n model for person detection in UAV-based indoor emergency scenarios.

While HOG-SVM achieves higher processing speeds ( $12.90 \pm 0.85$  FPS), its detection reliability is significantly lower, characterized by a 37.92% miss rate and frequent detection failures in over 30% of the frames. In contrast, YOLOv8n demonstrates superior robustness with a detection rate of 97.65%, a  $\text{mAP@0.5}$  of 0.95, and higher precision and recall values (0.94 and 0.98, respectively). Although YOLOv8n operates at a lower frame rate, its temporal stability and accuracy make it more suitable for critical emergency response. Consequently, YOLOv8n was selected for subsequent sensor-level analysis, while HOG-SVM was retained as a high-speed, low-robustness baseline.

### 4.1. Environmental impact and sensor robustness

YOLOv8n was evaluated under two indoor conditions: clear visibility and smoke-degraded visibility. Clear scenarios represent favorable conditions for both visible and thermal modalities. Smoke scenarios represent a degraded environment where visible-light information is significantly attenuated, while thermal sensing is expected to preserve heat-signature data.

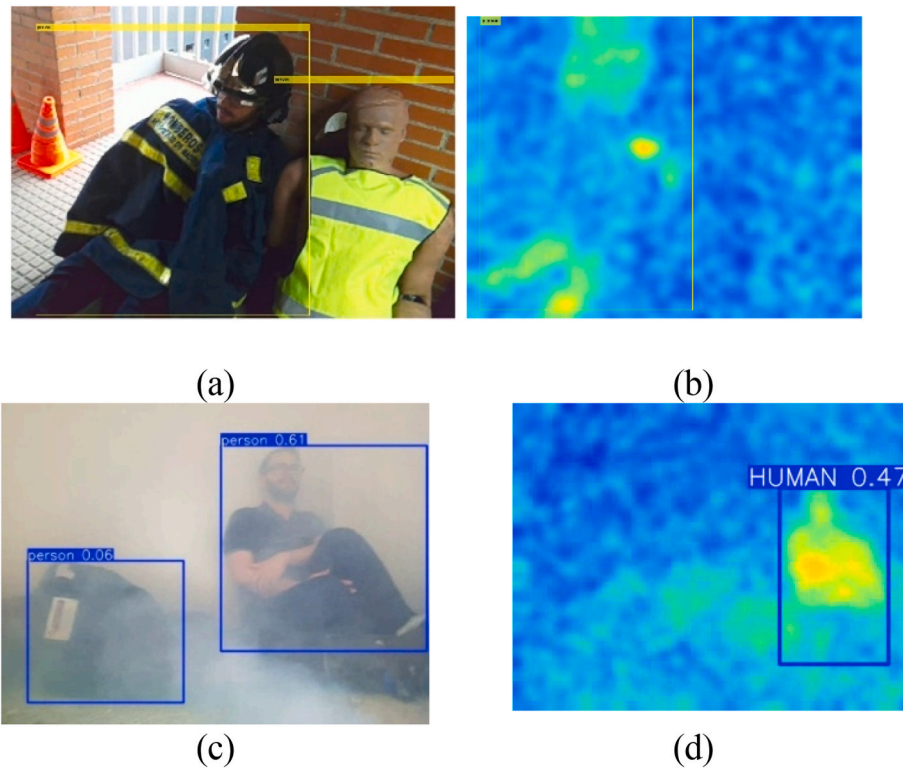
Fig. 11 illustrates representative detections across these four cases. In clear visible-light images, all targets are detected; however, in clear thermal images, only heat-emitting individuals produce distinguishable signatures. Under smoke conditions, visible-light detection becomes unreliable, whereas thermal sensing maintains detection capability despite a reduction in spatial detail.

### 4.2. Sensor ablation study

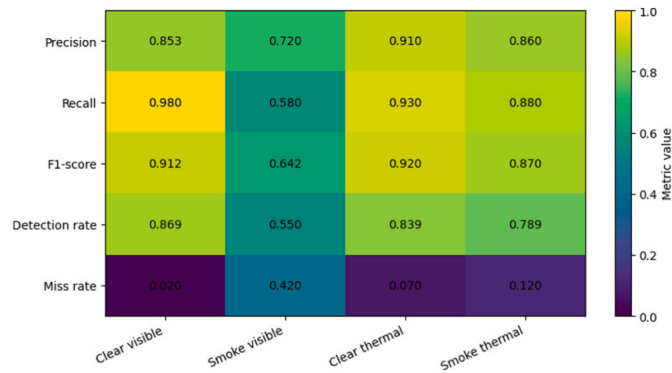
A sensor-level ablation study was conducted under clear and smoke conditions to quantify the contribution of each sensing modality. Fig. 12 reports the performance of visible and thermal sensing independently under both visibility conditions.

**Table 3**  
Comparison with UAV-based indoor emergency-response results.

| Metric                         | HOG - SVM        | YOLOv8n         |
|--------------------------------|------------------|-----------------|
| FPS                            | $12.90 \pm 0.85$ | $8.40 \pm 0.60$ |
| ms/frame                       | $77.5 \pm 5.1$   | $119.0 \pm 9.5$ |
| Detections per frame           | $0.70 \pm 0.35$  | $1.02 \pm 0.18$ |
| Detection rate (%)             | 62.08            | 97.65           |
| Miss rate (%)                  | 37.92            | 2.35            |
| Duplicate rate (%)             | $12.0 \pm 4.0$   | $4.5 \pm 2.0$   |
| Precision                      | $0.72 \pm 0.08$  | $0.94 \pm 0.03$ |
| Recall                         | $0.62 \pm 0.10$  | $0.98 \pm 0.02$ |
| $\text{mAP@0.5}$               | 0.58             | 0.95            |
| $\text{mAP@0.5:0.95}$          | 0.31             | 0.66            |
| Average IoU                    | $0.61 \pm 0.09$  | $0.78 \pm 0.05$ |
| Frames with $\geq 1$ detection | 598              | 941             |
| Frames without detection       | 365              | 22              |



**Fig. 11.** Representative YOLOv8n detection examples under clear and smoke conditions using visible and thermal sensing. (a) Clear visible image showing two detected people. (b) Clear thermal image showing the distinguishable signature of the heat-emitting person. (c) Smoke visible image showing degraded detection under reduced visibility. (d) Smoke thermal image showing preserved detection capability under smoke, although with reduced spatial detail.



**Fig. 12.** Performance matrix of sensor ablation study under clear and smoke conditions.

The results indicate that visible sensing is highly sensitive to smoke. Its recall decreases from 0.980 under clear conditions to 0.580 under smoke, while the miss rate increases from 0.020 to 0.420. Thermal sensing shows a more stable behavior, with recall decreasing only from 0.930 to 0.880 and a comparatively low miss rate under smoke. These results confirm the relevance of thermal sensing in smoke-filled environments.

#### 4.3. Visible-thermal fusion performance

Since thermal sensing alone may be insufficient for targets with weak or occluded signatures, a fused visible-thermal strategy was evaluated. Table 4 reports the performance of the fused approach for victim counting and thermal-state estimation, together with its improvement over individual sensing modalities in terms of total-counting error.

The fused approach provides a more complete estimation of the

**Table 4**

Fused visible-thermal performance and gain over individual sensing modalities.

| Metric                                    | Clear  | Smoke  | Interpretation                                                               |
|-------------------------------------------|--------|--------|------------------------------------------------------------------------------|
| Total persons MAE (Visible only)          | 0.12   | 0.35   | Visible counting degrades under smoke                                        |
| Total persons MAE (Thermal only)          | 0.31   | 0.41   | Thermal alone is limited for total counting                                  |
| Total persons MAE (Fusion)                | 0.12   | 0.23   | Fusion maintains lower counting error                                        |
| Hot persons MAE (Fusion)                  | 0.08   | 0.13   | Thermal information supports heat-emitting person estimation                 |
| Cold persons MAE (Fusion)                 | 0.10   | 0.20   | Visible-thermal complementarity supports non-heat-emitting person estimation |
| State accuracy (Fusion)                   | 0.94   | 0.86   | Thermal-state classification remains effective under smoke                   |
| Hot F1-score (Fusion)                     | 0.95   | 0.89   | Heat-emitting persons are reliably identified                                |
| Cold F1-score (Fusion)                    | 0.91   | 0.81   | Non-heat-emitting persons are more affected by smoke                         |
| Improvement in total-count MAE vs visible | 0%     | +34.3% | Fusion reduces smoke-induced degradation                                     |
| Improvement in total-count MAE vs thermal | +61.3% | +43.9% | Fusion outperforms thermal-only total counting                               |

scene by combining total person detection with thermal-state classification. In clear conditions, visible sensing already provides an accurate total count, so fusion does not further reduce the total-counting error with respect to visible-only detection. However, under smoke, the fused system reduces the total-count MAE from 0.35 to 0.23 compared with visible-only detection, corresponding to a 34.3% improvement. Compared with thermal-only detection, fusion reduces the error by 61.3% in clear conditions and 43.9% in presence of smoke. These results show that the fused visible-thermal configuration exploits the

complementary behavior of both sensors. Visible sensing contributes to estimating the total number of people, while thermal sensing improves robustness under smoke and supports thermal-state classification. This complementarity is particularly relevant in indoor fire-response scenarios, where visibility and thermal signatures may vary over time.

## 5. Discussion

This section discusses the main operational constraints of the proposed system, with the aim of clearly defining its intended scope and realistic deployment conditions in indoor fire-response scenarios.

While the prototype confirms the viability of the proposed approach, some aspects naturally remain at an early development stage. Sensor data streams are currently processed mostly in parallel, and deeper multi-modal fusion, including tighter time synchronization and uncertainty-aware integration, would further enhance robustness and decision consistency. Similarly, the reliance on a Wi-Fi communication link, which may operate in the 2.4 GHz band depending on onboard hardware, can limit throughput and resilience in interference-prone indoor environments, although it proved sufficient for validating the system architecture and real-time operation. Additional factors such as the use of lightweight perception algorithms, the limited resolution of the thermal sensor, and the restricted scale of experimental validation reflect practical design trade-offs made to ensure real-time performance and system stability. These aspects open clear opportunities for future improvements rather than fundamental limitations.

Although the ground-station architecture supports independent communication processes for multiple UAV connections, the present experiments were conducted with a single UAV. Consequently, fleet coordination, inter-UAV collision avoidance, cooperative mission planning, and multi-UAV scalability were not experimentally evaluated.

The platform relies on a Wi-Fi communication link to transmit sensor data and video streams to the ground station. In complex indoor environments, particularly in the presence of dense smoke, metallic structures, or strong multipath effects, wireless performance may be degraded, leading to increased latency or temporary loss of data. Despite these limitations, the communication setup proved adequate for real-time situational awareness during the experimental trials conducted in this work. The system prioritizes low-latency data delivery over other mechanisms, which is a deliberate design choice for real-time emergency-response scenarios. The system mitigates potential data loss through continuous high-frequency streaming and temporal redundancy avoiding the implementation of complex procedures, e.g., handshake or retransmission mechanisms. Image data are transmitted using MJPEG encoding, ensuring that the loss of individual packets does not propagate across the stream. This design reduces the impact of occasional packet loss, as new data are continuously received and quickly replace missing information.

Operational endurance is constrained by onboard battery capacity, especially during hover-dominated missions and continuous operation of multiple sensors. Consequently, the system is best suited for short-duration reconnaissance and assessment tasks rather than prolonged autonomous missions in extreme conditions. The estimated flight endurance decreases from approximately 29.6 min without payload to 24.3 min at maximum payload of 250 g according to the manufacturer's reference data, representing a reduction of about 18%. The flight endurance is considered sufficient for short-duration indoor monitoring, as the integrated monitoring system operates close to the UAV's maximum payload capacity while remaining within specified limits.

In the event of temporary communication dropouts, the system maintains onboard flight stabilization while real-time data transmission is interrupted until connectivity is restored. From a sensing perspective, extreme fire conditions such as very dense smoke or high temperatures may reduce the effectiveness of both visible-spectrum and thermal imaging. Under such circumstances, the system outputs should be interpreted as complementary information to support human decision-

making rather than as fully reliable measurements. Despite the technical challenges encountered, such as the payload capacity of the drones and the limitations of the compatibility of different Raspberry Pi versions with 5 GHz Wi-Fi networks, the implementation remained functional and demonstrated the potential for practical application. The selection of Python as the programming language, while introducing some limitations due to the Global Interpreter Lock, was effectively managed using the multiprocessing library to enable concurrent function execution.

The proposed indoor navigation strategy is based on relative positioning using optical flow, inertial measurements, and altitude estimation, enabling trajectory recording and subsequent autonomous reproduction. The system presents several limitations because it does not incorporate environment perception capabilities such as obstacle detection, avoidance, or dynamic path replanning. The navigation is restricted to the execution of pre-recorded trajectories developed by pilots, and it should be considered as a preliminary solution focused on sensing, communication and situational awareness in GPS-denied environments. Future work will address these limitations through the integration of perception-driven navigation techniques, including obstacle detection, sensor fusion (e.g., LiDAR or depth cameras), and real-time path replanning algorithms.

Overall, these limitations delineate the operational envelope of the proposed platform and highlight directions for future enhancements aimed at improving robustness, endurance, and communication reliability in harsh indoor fire environments.

## 6. Conclusions and future lines

Indoor emergency scenarios, including fires, industrial accidents, and structural collapses, present significant challenges for first responders due to limited visibility, dynamic conditions, and restricted accessibility. UAV-based platforms provide valuable capabilities for real-time situational awareness; however, indoor deployment requires robust solutions to address constraints in communication, perception, and onboard processing.

This paper presents and validates an integrated UAV-based platform for indoor emergency response, designed to manage independent connections from multiple UAVs and supporting real-time acquisition, transmission, visualization, and storage of heterogeneous data streams. The proposed architecture enables reliable operation and coordinates multi-modal sensing in complex indoor environments.

A comparative evaluation demonstrates the clear advantage of deep learning-based detection over classical approaches. While the HOG-based method offers higher processing speed, the YOLOv8n detector significantly improves detection performance, achieving substantially higher detection rates and recall while drastically reducing missed detections. These results highlight the suitability of modern deep learning models for reliable victim detection in emergency scenarios.

Furthermore, the multi-modal fusion strategy proves essential under degraded visibility conditions. By combining visual and thermal data, the system reduces counting errors and maintains robust performance under smoke, while enabling effective classification of heat-emitting and non-heat-emitting individuals. This demonstrates the importance of integrating complementary sensing modalities to ensure reliable perception in challenging environments.

Future work will focus on enhancing payload capacity, optimizing processing efficiency, and exploring advanced multi-modal fusion and autonomous navigation strategies for improved performance in GNSS-denied environments. Future developments will incorporate more advanced perception capabilities, including LiDAR and depth cameras, to improve environmental understanding, obstacle detection, autonomous navigation, and victim detection. They will also analyze coupled fusion strategies, such as pixel-level and feature-level integration, to complement synchronized data visualization with more integrated multi-modal analysis. Further developments may also consider more

advanced synchronization mechanisms, including hardware-triggered acquisition, and probabilistic sensor-fusion frameworks capable of accounting for uncertainty under dynamic fire conditions. The platform will also be evaluated in more complex and realistic indoor fire scenarios, including unpredictable fire propagation, toxic atmospheres, and structural hazards, to further assess its operational robustness.

A standardized evaluation protocol is required for UAV based indoor emergency-response systems. Differences in experimental environments, environmental conditions, sensing hardware, processing platforms, and reported metrics increase the complexity of quantitative comparison with previous studies. Future iterations will address more rigorous comparisons by defining reproducible evaluation procedures based on accuracy, end-to-end latency, communication reliability, and navigation performance under different visibility and interference conditions. The influence of FPS on the detection and measurement process will also be analyzed, considering temporal resolution, computational load, and detection reliability. Hardware acceleration, model pruning, and quantization strategies will also be evaluated to improve frame rates without significantly reducing detection accuracy.

Overall, the proposed system demonstrates the effectiveness of combining a scalable communication architecture, deep-learning-based perception, and multi-modal sensing to support emergency response operations in complex indoor scenarios.

### CRedit authorship contribution statement

**Carlos Quiterio Gómez Muñoz:** Writing – review & editing, Visualization, Validation, Software, Resources, Investigation, Funding acquisition, Data curation, Conceptualization. **Isaac Segovia Ramirez:** Writing – original draft, Visualization, Validation, Software, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Guillermo González de Rivera Peces:** Writing – original draft, Validation, Software, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Pedro Rafael Fernandez Barbosa:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization. **Maria Cristina Rodríguez-Sánchez:** Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

### Ethics declaration

The authors have NOT obtained informed consent from participants or their legal representatives. Not applicable.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. Ethics committee approval was not required under relevant laws and institutional guidelines. Not applicable.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: M. Cristina Rodríguez-Sánchez reports financial support was provided by Rey Juan Carlos University. M. Cristina Rodríguez-Sánchez reports a relationship with Rey Juan Carlos University that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

The authors thank the financial support received from the Young Researchers R&D Project GUIDE2FR (ref. TED2021-131430B-I00), funded by MCIN/AEI/10.13039/501100011033. The authors also acknowledge financial support from RED2022-134355-T, funded by

MCIN/AEI/10.13039/501100011033/FEDER, UE.

Additionally, the project “ICARUS - Inspección y Control Automatizado con Redes neuronales y UAVs en Sistemas Fotovoltaicos” (ref. SI4/PJI/2024-00233) is funded by the Comunidad de Madrid through the direct grant agreement for the promotion of research and technology transfer at the Universidad Autónoma de Madrid.

### Data availability

Data will be made available on request.

### References

- [1] Agbedahin AV. Sustainable development, education for sustainable development, and the 2030 agenda for sustainable development: emergence, efficacy, eminence, and future. *Sustain Dev* 2019;27:669–80. <https://doi.org/10.1002/sd.1931>.
- [2] Statistics E. *Statistics on European cities*. 2014.
- [3] Ekins P. Final report of the high-level panel of the European decarbonisation pathways initiative. 2018.
- [4] Abdulhalim I, Mutch C, González VA, Amor R. Improving post-earthquake evacuation preparedness for deaf and hard of hearing children: a conceptual framework. *Int J Disaster Risk Reduct* 2021;62:102360. <https://doi.org/10.1016/j.ijdr.2021.102360>.
- [5] Mapfre F. Estudio de víctimas de incendios en España en 2019. 2019.
- [6] Barberán García, C.; Vieitez, A. Incendio en túneles y galerías.
- [7] Kong SG, Jin D, Li S, Kim H. Fast fire flame detection in surveillance video using logistic regression and temporal smoothing. *Fire Saf J* 2016;79:37–43. <https://doi.org/10.1016/j.firesaf.2015.11.015>.
- [8] Sandhu HS, Singh KJ, Kapoor DS. Automatic edge detection algorithm and area calculation for flame and fire images. In: *Proceedings of the 2016 6th international conference-cloud system and big data engineering (Confluence)*; 2016. p. 403–7. <https://doi.org/10.1109/CONFLUENCE.2016.7508152>.
- [9] Qin H, Cui JQ, Li J, Bi Y, Lan M, Shan M, Liu W, Wang K, Lin F, Zhang Y. Design and implementation of an unmanned aerial vehicle for autonomous firefighting missions. In: *Proceedings of the 2016 12th IEEE international conference on control and automation (ICCA)*; 2016. p. 62–7. <https://doi.org/10.1109/ICCA.2016.7505253>.
- [10] Khan AA, Nauman MA, Shoaib M, Jahangir R, Alroobaea R, Alsafyani M, Binmahfoudh A, Wechtaisong C. Crowd anomaly detection in video frames using fine-tuned AlexNet model. *Electronics* 2022;11:3105. <https://doi.org/10.3390/electronics11193105>.
- [11] Roldán-Gómez JJ, González-Gironda E, Barrientos A. A survey on robotic technologies for forest firefighting: applying drone swarms to improve firefighters' efficiency and safety. *Appl Sci* 2021;11:363. <https://doi.org/10.3390/app11010363>.
- [12] Fukuhara T, Kouyama T, Kato S, Nakamura R, Takahashi Y, Akiyama H. Detection of small wildfire by thermal infrared camera with the uncooled microbolometer array for 50-kg class satellite. *IEEE Trans Geosci Rem Sens* 2017;55:4314–24. <https://doi.org/10.1109/TGRS.2017.2690996>.
- [13] Gómez CQ, Villegas MA, García FP, Pedregal DJ. Big data and web intelligence for condition monitoring: a case study on wind turbines. In: *Handbook of research on trends and future directions in big data and web intelligence*. IGI global; 2015. p. 149–63.
- [14] Lavallo A, Teruel MA, Maté A, Trujillo J. Improving sustainability of smart cities through visualization techniques for big data from IoT devices. *Sustainability* 2020;12:5595. <https://doi.org/10.3390/su12145595>.
- [15] Cheltha JN, Sharma C, Prashar D, Khan AA, Kadry S. Enhanced human motion detection with hybrid RDA-WOA-based RNN and multiple hypothesis tracking for occlusion handling. *Image Vis Comput* 2024;150:105234.
- [16] AlAli ZT, Alabady SA. A survey of disaster management and SAR operations using sensors and supporting techniques. *Int J Disaster Risk Reduct* 2022;82:103295. <https://doi.org/10.1016/j.ijdr.2022.103295>.
- [17] Bemposta Rosende S, Sánchez-Soriano J, Gómez Muñoz CQ, Fernández Andrés J. Remote management architecture of UAV fleets for maintenance, surveillance, and security tasks in solar power plants. *Energies* 2020;13:5712. <https://doi.org/10.3390/en13215712>.
- [18] Montoya JC, Muñoz CQG, Márquez FPG. Remote condition monitoring for photovoltaic systems. In: *Non-Destructive testing and condition monitoring techniques for renewable energy industrial assets*. Elsevier; 2020. p. 133–42.
- [19] Gómez Muñoz CQ, García Márquez FP. Future maintenance management in renewable energies. In: *Renewable energies: business Outlook 2050*; 2018. p. 149–59.
- [20] Salazar F, Martínez-García MS, de Castro A, Logroño N, Cazorla-Logroño MF, Guamán-Molina J, Gómez C. Optimization of the solar energy storage capacity for a monitoring UAV. *Sustain Futures* 2024;7:100146. <https://doi.org/10.1016/j.sfr.2023.100146>.
- [21] Gómez Muñoz CQ, Paredes Álvarez C, García Márquez FP. Smart farming: intelligent management approach for crop inspection and evaluation employing unmanned aerial vehicles. In: *Proceedings of the international conference on management science and engineering management*; 2020. p. 119–30. [https://doi.org/10.1007/978-3-030-49889-4\\_11](https://doi.org/10.1007/978-3-030-49889-4_11).

- [22] Aliane N, Muñoz CQG, Sánchez-Soriano J. Web and MATLAB-based platform for UAV flight management and multispectral image processing. *Sensors* 2022;22: 4243. <https://doi.org/10.3390/s22114243>.
- [23] Zhang Y, Yuan X, Li W, Chen S. Automatic power line inspection using UAV images. *Remote Sens* 2017;9:824. <https://doi.org/10.3390/rs9080824>.
- [24] Santos T, Moreira M, Almeida J, Dias A, Martins A, Dinis J, Formiga J, Silva E. PLineD: Vision-based power lines detection for unmanned aerial vehicles. In: Proceedings of the 2017 IEEE International conference on autonomous robot systems and competitions (ICARSC); 2017. p. 253–9. <https://doi.org/10.1109/ICARSC.2017.7964084>.
- [25] Stokkeland M, Klausen K, Johansen TA. Autonomous visual navigation of unmanned aerial vehicle for wind turbine inspection. In: Proceedings of the 2015 international conference on unmanned aircraft systems (ICUAS); 2015. p. 998–1007. <https://doi.org/10.1109/ICUAS.2015.7152389>.
- [26] Schäfer BE, Picchi D, Engelhardt T, Abel D. Multicopter unmanned aerial vehicle for automated inspection of wind turbines. In: Proceedings of the 2016 24th mediterranean conference on control and automation (MED); 2016. p. 244–9. <https://doi.org/10.1109/MED.2016.7536055>.
- [27] Nikolic J, Burri M, Rehder J, Leutenegger S, Huerzeler C, Siegwart R. A UAV system for inspection of industrial facilities. In: Proceedings of the 2013 IEEE Aerospace Conference; 2013. p. 1–8. <https://doi.org/10.1109/AERO.2013.6496959>.
- [28] Banić M, Miltenović A, Pavlović M, Ćirić I. Intelligent machine vision based railway infrastructure inspection and monitoring using UAV. *Facta Univ – Ser Mech Eng* 2019;17:357–64. <https://doi.org/10.22190/FUME190507041B>.
- [29] Hallermann N, Morgenthal G. Visual inspection strategies for large bridges using unmanned aerial vehicles (UAV). In: Proceedings of the proc. of 7th IABMAS, international conference on Bridge maintenance, safety and management; 2014. p. 661–7. <https://doi.org/10.1201/b17063-96>.
- [30] Muñoz CQG, García Márquez FP. Wind energy power prospective. *Renew Energies: Bus Outlook* 2018;83–95. [https://doi.org/10.1007/978-3-319-45364-4\\_6](https://doi.org/10.1007/978-3-319-45364-4_6).
- [31] Gómez CQ, García Márquez FP, Arcos A, Cheng L, Kogia M, Papaelias M. Calculus of the defect severity with EMATs by analyzing the attenuation curves of the guided waves. 2016. <https://doi.org/10.12989/sss.2017.19.2.195>.
- [32] Arcos Jiménez A, Gómez CQ, García Márquez FP. Concentrated solar plants management: big data and neural network. In: *Renewable energies*. Springer; 2018. p. 63–81. [https://doi.org/10.1007/978-3-319-45364-4\\_5](https://doi.org/10.1007/978-3-319-45364-4_5).
- [33] Sanjana P, Prathilothamai M. Drone design for first aid kit delivery in emergency situation. In: Proceedings of the 2020 6th international conference on advanced computing and communication systems (ICACCS); 2020. p. 215–20. <https://doi.org/10.1109/ICACCS48705.2020.9074487>.
- [34] Daud SMSM, Yusof MYPM, Heo CC, Khoo LS, Singh MKC, Mahmood MS, Nawawi H. Applications of drone in disaster management: a scoping review. *Sci Justice* 2022;62:30–42. <https://doi.org/10.1016/j.scijus.2021.11.002>.
- [35] Seguin C, Blaquière G, Loundou A, Michelet P, Markarian T. Unmanned aerial vehicles (drones) to prevent drowning. *Resuscitation* 2018;127:63–7. <https://doi.org/10.1016/j.resuscitation.2018.04.005>.
- [36] Segovia Ramírez I, Gómez Muñoz CQ, López Pérez S, Rodríguez Sánchez MC. Real-time detection in indoor scenarios for UAV-based thermal-visual imaging. In: Proceedings of the the nineteenth international conference on management science and engineering management, Singapore, 2025; 2025. p. 377–87. [https://doi.org/10.1007/978-981-95-1595-0\\_27](https://doi.org/10.1007/978-981-95-1595-0_27).
- [37] Liu P, Yu H, Cang S, Vladareanu L. Robot-assisted smart firefighting and interdisciplinary perspectives. In: Proceedings of the 2016 22nd international conference on automation and computing (ICAC); 2016. p. 395–401. <https://doi.org/10.1109/IConAC.2016.7604952>.
- [38] Gaur A, Singh A, Kumar A, Kulkarni KS, Lala S, Kapoor K, Srivastava V, Kumar A, Mukhopadhyay SC. Fire sensing technologies: a review. *IEEE Sens J* 2019;19: 3191–202. <https://doi.org/10.1109/JSEN.2019.2894665>.
- [39] Fonollosa J, Solórzano A, Marco S. Chemical sensor systems and associated algorithms for fire detection: a review. *Sensors* 2018;18:553. <https://doi.org/10.3390/s18020553>.
- [40] Luneckas M, Luneckas T, Kriauciūnas J, Udris D, Plonis D, Damaševičius R, Maskeliūnas R. Hexapod robot gait switching for energy consumption and cost of transport management using heuristic algorithms. *Appl Sci* 2021;11:1339. <https://doi.org/10.3390/app11031339>.
- [41] Murphy RR. A decade of rescue robots. In: Proceedings of the 2012 IEEE/RSJ international conference on intelligent robots and systems; 2012. p. 5448–9. <https://doi.org/10.1109/IROS.2012.6386301>.
- [42] Grogan S, Pellerin R, Gamache M. The use of unmanned aerial vehicles and drones in search and rescue operations—a survey. In: Proceedings of the PROLOG; 2018. p. 1–13.
- [43] Guzman R, Navarro R, Ferre J, Moreno M. Rescuer: development of a modular chemical, biological, radiological, and nuclear robot for intervention, sampling, and situation awareness. *J Field Robot* 2016;33:931–45. <https://doi.org/10.1002/rob.21588>. DOI: 10.1142/9789814291279\_0050.
- [44] Baudoin Y, Doroftei D, De Cubber G, Berrabah S, Colon E, Pinzon C, Maslowski A, Bedkowski J, Penders J. VIEW-FINDER: robotics assistance to fire-fighting services. In: *Mobile robotics: solutions and challenges*. World Scientific; 2010. p. 397–406.
- [45] Jiménez AA, Muñoz CQG, Márquez FPG. Machine learning and neural network for maintenance management. In: Proceedings of the proceedings of the eleventh international conference on management science and engineering management, vol. 11; 2018. p. 1377–88.
- [46] Aydin B, Selvi E, Tao J, Starek MJ. Use of fire-extinguishing balls for a conceptual system of drone-assisted wildfire fighting. *Drones* 2019;3:17. <https://doi.org/10.3390/drones3010017>.
- [47] Akhloufi MA, Couturier A, Castro NA. Unmanned aerial vehicles for wildland fires: sensing, perception, cooperation and assistance. *Drones* 2021;5:15. <https://doi.org/10.3390/drones5010015>.
- [48] Aláez D, Prieto M, Villadangos J, Astrain JJ. Towards a heat-resistant tethered micro-aerial vehicle for structure fire sensing. *Appl Sci* 2025;15:2388. <https://doi.org/10.3390/app15052388>.
- [49] Do TT, Nguyen BD, Nguyen DD, Nguyen DM, Nguyen TT, Nguyen QK. Drone-based autonomous navigation and victim detection using deep learning technique for Non-GPS In-Building environments. In: Proceedings of the advances in computational collective intelligence; 2026. p. 15–36. [https://doi.org/10.1007/978-3-032-10209-6\\_2](https://doi.org/10.1007/978-3-032-10209-6_2). Cham, 2026.
- [50] Boiteau S, Vanegas F, Gonzalez F. Framework for autonomous uav navigation and target detection in global-navigation-satellite-system-denied and visually degraded environments. *Remote Sens* 2024;16:471. <https://doi.org/10.3390/rs16030471>.
- [51] Kucukayan G, Karacan H. Yolo-ihd: improved real-time human detection system for indoor drones. *Sensors* 2024;24:922. <https://doi.org/10.3390/s24030922>.
- [52] Kang S, Hu Z, Liu L, Zhang K, Cao Z. Object detection YOLO algorithms and their industrial applications: overview and comparative analysis. *Electronics* 2025;14: 1104. <https://doi.org/10.3390/electronics14061104>.