

Article

Architecture of an AI-Driven Optoelectronic ISR UAV System with Operator-Supervised Autonomy

Alexandru-Dragoş Adam ^{1,*}, Alina Nirvana Popescu ^{1,*}  and Jair Gonzalez ²

¹ Faculty of Automatic Control and Computer Science, National University of Science and Technology POLITEHNICA Bucharest, 060042 Bucharest, Romania; alexandru.adam2610@stud.acs.upb.ro

² Ansys Inc., 06270 Villeneuve-Loubet, France; jair.gonzalez@ansys.com

* Correspondence: nirvana.popescu@upb.ro

Abstract

This paper presents a proposed architecture for an artificial intelligence-driven unmanned aerial vehicle (UAV) system intended for tactical intelligence, surveillance, and reconnaissance (ISR) missions. The architecture brings together electro-optical imaging, long-wave infrared sensing, two-dimensional light detection and ranging (LiDAR), inertial navigation support, onboard edge computing, and resilient communication links within a unified system-level framework. Unlike many existing approaches that treat perception, autonomy, communication, and safety as loosely coupled functions, the proposed architecture combines multi-modal sensing, operator-supervised autonomy, and a safety-oriented decision validation layer intended for future integration with Ansys SCADE. The system is structured around operational and sensor-performance requirements used to justify the selection and interaction of the main onboard subsystems. At the architectural level, the proposed framework is intended to support target detection, tracking, environment awareness, and mission-level decision support under degraded visibility, constrained communication, and contested operating conditions. The paper therefore contributes a requirement-driven and safety-aware ISR UAV architecture that provides a scalable basis for future implementation, validation, and multi-UAV extension.

Keywords: AI-enabled UAVs; ISR; EO–IR multi-sensor fusion; edge AI and onboard perception; safety-critical autonomy; Ansys SCADE-based verification

1. Introduction

Modern Intelligence, Surveillance, and Reconnaissance (ISR) missions require rapid situational awareness, robust sensing, and timely interpretation of large volumes of heterogeneous data. In tactical environments, traditional ISR assets are often limited by restricted endurance, exposure to hostile conditions, and dependence on human operators for continuous perception and interpretation. Unmanned Aerial Vehicles (UAVs) have therefore become increasingly relevant for reconnaissance and surveillance tasks, as they can provide flexible deployment, persistent observation, and access to areas that are difficult or dangerous to monitor from the ground.

At the same time, the growing use of electro-optical, infrared, LiDAR, and inertial sensing on UAV platforms has significantly increased the amount and diversity of mission data available during flight. This creates a need for onboard processing architectures capable not only of acquiring sensor data, but also of transforming it into actionable information under communication, energy, and environmental constraints. Artificial intelligence is



Academic Editor: Libor Pekař

Received: 21 March 2026

Revised: 24 April 2026

Accepted: 27 April 2026

Published: 29 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

particularly relevant in this context, as it enables automated target detection, tracking, anomaly identification, and multi-sensor interpretation, while supporting operation in degraded visibility or partially GNSS-denied conditions.

Despite these advances, many UAV systems still treat sensing, perception, communication, navigation, and supervisory control as separate functions. However, tactical ISR missions require these capabilities to operate coherently within a unified framework that supports both mission effectiveness and operational safety. This motivates the need for an integrated architecture combining multi-modal sensing, onboard AI-assisted perception, resilient communication, and safety-oriented supervisory logic.

The objective of this paper is to present a coherent UAV ISR architecture that integrates multi-modal perception, onboard AI-assisted processing, resilient communication, and safety-oriented supervisory logic within a single system-level framework. The following sections summarize the relevant literature, describe the proposed platform architecture, and discuss its role as a basis for future implementation and validation.

The paper is organized as follows. Section 2 reviews the related work on AI-enabled UAV reconnaissance systems and multi-sensor perception architectures. Section 3 presents the main system components and hardware–software architecture of the proposed platform. Section 4 describes the system diagram and communication flow between subsystems. Section 5 details the AI perception and decision-support architecture. Section 6 concludes the paper, while Section 7 outlines the main future research directions and contributions of the proposed work.

2. Related Work

This section provides a concise overview of relevant work in AI-enabled UAV ISR systems, focusing on key technological trends and limitations that motivate the proposed architecture.

The integration of artificial intelligence into UAV platforms has significantly advanced airborne perception capabilities. Early approaches relied on classical image processing and geometry-based methods with limited real-time performance [1]. More recent developments in deep learning have enabled robust object detection and tracking, with convolutional and transformer-based models adapted to aerial imagery constraints such as small target size, scale variation, and cluttered backgrounds [2–6].

A major trend in UAV-based ISR systems is the transition from single-sensor perception to multi-sensor and multi-modal fusion. Combining electro-optical (EO), infrared (IR), LiDAR, and inertial measurements has been shown to improve robustness and accuracy in both perception and navigation tasks, particularly in degraded or GPS-denied environments [7–10]. EO–IR fusion, in particular, has become a key enabler for reliable detection under low-visibility or camouflage conditions, while LiDAR integration supports spatial awareness and mapping [11,12].

In parallel, navigation and autonomy have evolved toward AI-assisted and multi-sensor-driven approaches. Visual SLAM, visual–inertial odometry, and learning-based navigation frameworks enable UAV operation in complex and cluttered environments where GNSS signals may be unreliable [4,7,8,13]. These systems increasingly rely on tightly coupled perception and control pipelines to support obstacle avoidance, trajectory planning, and environment understanding.

Beyond detection and navigation, anomaly detection and behavior analysis techniques are gaining relevance in ISR applications. Methods based on autoencoders, clustering, and statistical learning allow the identification of unusual patterns without extensive labeled datasets, supporting surveillance tasks such as intrusion detection or activity monitoring [14,15].

Communication and coordination frameworks have also evolved to support distributed UAV operations. Encrypted telemetry links, frequency-hopping communication, and mesh networking architectures improve resilience in contested environments, while concepts such as swarm coordination and federated learning enable collaborative intelligence and data sharing across multiple platforms [16–18].

At the hardware level, advances in embedded and edge computing platforms, such as Jetson-based systems and low-power AI accelerators, enable real-time onboard inference under strict size, weight, and power constraints [19,20]. These developments allow UAVs to process perception tasks locally, reducing dependence on high-bandwidth communication links and improving operational autonomy.

Energy-aware system design has also become an important research direction. Adaptive power management strategies, intelligent battery monitoring, and task scheduling approaches aim to balance computational load with mission endurance, ensuring reliable operation in long-duration ISR scenarios [21,22].

Despite these significant advances, existing AI-enabled UAV systems are typically developed as collections of loosely coupled subsystems, where perception, navigation, communication, and decision-making are treated independently. While multi-sensor fusion and deep learning techniques have improved detection and navigation performance, several critical limitations remain insufficiently addressed.

First, most approaches rely on probabilistic AI outputs without integrating formally verified safety constraints, which limits their applicability in safety-critical ISR scenarios. Second, existing architectures often assume either fully autonomous operation or operator-driven control, without explicitly addressing intermediate paradigms such as operator-supervised autonomy. Third, although multi-modal sensing is widely adopted, there is a lack of requirement-driven system design frameworks that explicitly link sensor performance with operational ISR objectives.

Moreover, the current literature predominantly focuses on algorithmic performance, such as detection accuracy or SLAM robustness, while less attention is given to system-level integration, including real-time communication between modules, energy-aware computation, and interoperability between perception, control, and safety layers.

To address these limitations, this work proposes a unified architecture for AI-driven UAV-based ISR systems that integrates multi-modal perception, operator-supervised autonomy, and a safety-critical verification layer implemented using Ansys SCADE. Unlike existing approaches, the proposed framework explicitly combines probabilistic AI inference with deterministic safety constraints and aligns hardware–software design with clearly defined operational and sensor performance requirements.

3. System Components

The selection of the hardware and software modules included in the proposed UAV platform is not arbitrary, but is derived from a set of clearly defined UAV Operational Requirements (OR) and Sensor Performance Requirements (SPR) associated with small tactical ISR missions. These requirements establish the functional envelope within which subsystem design choices are evaluated and justified.

From an operational perspective, the platform is specified to support persistent aerial surveillance operations, with a mission endurance of approximately 45–60 min under nominal payload conditions (OR-1). The UAV must maintain autonomous navigation capability under partial GNSS degradation (up to ~50% satellite loss), reflecting the likelihood of jamming or multipath effects in contested environments (OR-2).

Communication performance is constrained by the need for low-latency EO video streaming to support operator supervision and AI-assisted guidance, with an expected

end-to-end latency not exceeding ≈ 250 ms (OR-3), while command-and-control availability must remain above 0.95 during mission execution (OR-4).

Additionally, subsystem selection is shaped by SWaP-C constraints typical of compact tactical UAVs, with strict limitations on payload mass and power consumption to ensure that computational and sensing capabilities do not compromise endurance (OR-5). Finally, the sensing payload must remain operationally meaningful at representative ISR stand-off distances, supporting reliable target detection and recognition within the 500–800 m range envelope (OR-6).

From a sensing perspective, the payload design is further constrained by the classical Johnson criteria for electro-optical target acquisition, which relate target discrimination performance to the number of resolvable details across the target. In surveillance-oriented interpretations, detection is commonly associated with approximately 1–2 pixels across the minimum target dimension, while recognition requires a higher spatial sampling level, typically around 6–8 pixels across the target width. In the context of this work, these thresholds are used as practical design references for evaluating whether the EO/IR payload can support long-range detection and the closer-range recognition of personnel and light vehicles within the representative ISR range of approximately 500–800 m. Accordingly, these values define the minimum sensing-performance envelope used to justify the selection of the EO and thermal modules considered in the proposed architecture.

In the present work, these thresholds are not introduced as experimentally validated performance results of the proposed platform, but as literature-based design references derived from Johnson-type target-acquisition criteria commonly used in electro-optical and infrared surveillance analysis. Within this interpretation, detection is typically associated with approximately 1–2 pixels across the minimum target dimension, while recognition requires a higher spatial sampling level, often expressed in practical surveillance assessments as approximately 6–8 pixels across the target width. The representative ISR range of approximately 500–800 m is likewise used here as an architectural stand-off envelope for compact tactical reconnaissance missions rather than as a fixed operational guarantee. Its practical value depends on the sensor's resolution, optical configuration, target size and aspect, atmospheric conditions, platform altitude, and image-stabilization quality. Accordingly, these values are used in this paper to support requirement definition and sensing-payload selection at the architectural level [23].

Within this requirement-driven design framework, the following subsections present the physical and logical architecture of the UAV platform and explain the role of each subsystem in satisfying the above-defined OR and SPR constraints.

3.1. Component Overview

A high-level logical representation of the proposed UAV system is shown in Figure 1, which summarizes the main onboard and ground-segment components, together with their data, control, telemetry, and power connections.

At the core of the sensing payload lies the electro-optical (EO) camera, implemented through the DJI O4 Air Unit (DJI, Shenzhen, China). This module was selected for its capability to deliver high-resolution 4K video with exceptionally low transmission latency, enabling high-fidelity visual input for AI-driven perception pipelines. Beyond providing FPV for operator situational awareness, the EO stream serves as the primary source for object detection, scene classification, and SLAM-based localization, consistent with recent trends in vision-centric UAV reconnaissance research. This design choice is motivated by the Sensor Performance Requirements, as the 4K EO stream provides sufficient pixel coverage to satisfy Johnson-based detection and recognition thresholds for personnel and vehicles

at 500–800 m, while also meeting the low-latency perception and situational-awareness constraints defined in the Operational Requirements.

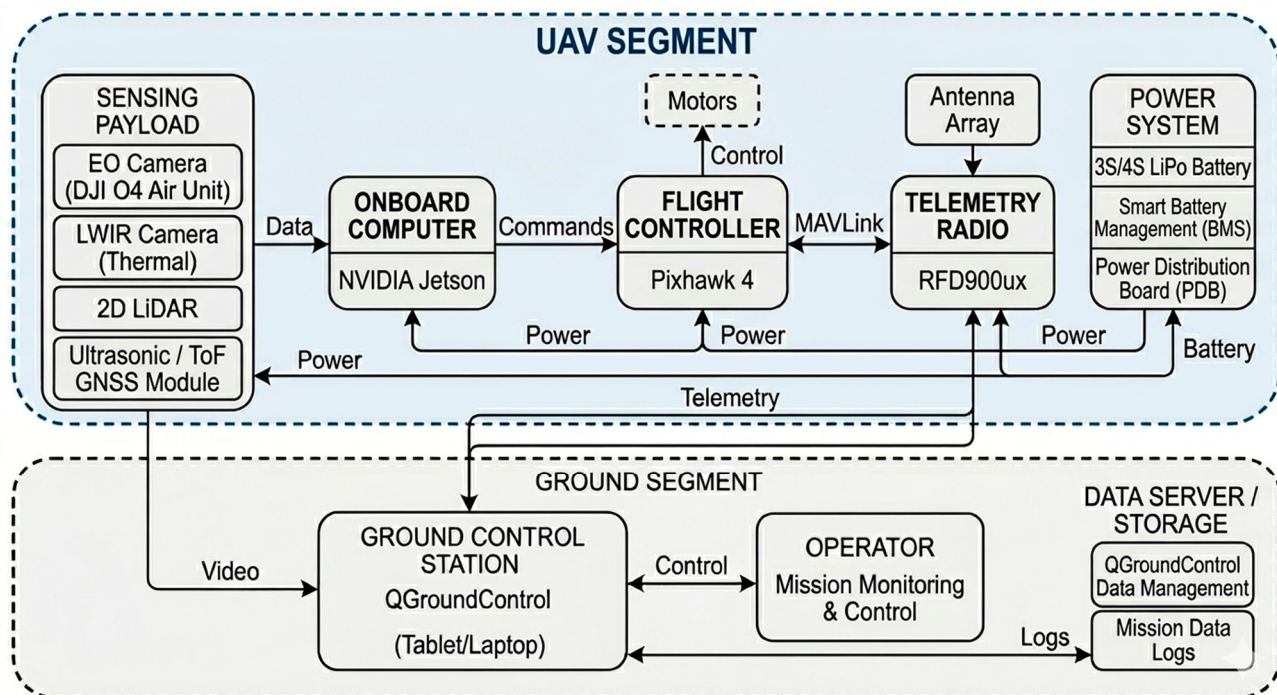


Figure 1. Logical block diagram of UAV system components and connections.

Complementing the EO system, the platform integrates a long-wave infrared (LWIR) sensor that extends perception to challenging illumination conditions, including nighttime operations, occlusions, and thermal camouflage scenarios. Thermal imagery facilitates human and vehicle detection under concealment and supports multi-modal fusion with EO images, enhancing target discrimination and environmental understanding. LWIR integration is aligned with current ISR methodologies that increasingly rely on EO–IR fusion to ensure operational robustness across diverse combat environments. The inclusion of the LWIR channel is therefore not merely complementary, but requirement-driven: it ensures that target detection performance remains within the SPR envelope under night-time and low-contrast conditions where EO imagery alone would fail to satisfy the Operational Requirements for mission continuity.

For spatial perception and mapping, the UAV employs a suite of range sensors, including a lightweight 2D LiDAR for planar scanning and ultrasonic/ToF sensors for short-range awareness. The 2D LiDAR offers a balance between coverage, weight, power consumption, and computational load, making it suitable for compact tactical UAVs operating in GPS-denied or cluttered settings. High-rate scans contribute to obstacle detection, environment reconstruction, and SLAM, while proximity sensors ensure safe landing, hovering, and avoidance of near-field hazards. The selection of a lightweight 2D LiDAR, complemented by short-range ToF/ultrasonic sensing, is justified by Operational Requirements relating to autonomous navigation in cluttered or partially GPS-denied environments, providing an obstacle-awareness capability without exceeding the SWaP-C constraints of the platform.

Central to the system's autonomy is the onboard AI processing unit, instantiated depending on the mission profile rather than through strict hardware interchangeability. Two processing options are supported: an NVIDIA Jetson module (NVIDIA Corporation, Santa Clara, CA, USA) and the Raspberry Pi 5 (Raspberry Pi Ltd., Cambridge, UK). These platforms address different requirement-driven operating modes derived from the OR and SPR constraints.

At the current stage of development, the Raspberry Pi 5 is considered a practical starting platform for early integration and testing of selected perception functions, including person-detection algorithms and other lightweight AI-assisted routines. Its lower cost, accessibility, and sufficient capability for initial prototyping make it suitable for validating data flow, communication, and basic onboard processing within the proposed architecture. However, for more advanced ISR configurations involving dense EO/IR perception, multi-modal fusion, continuous target tracking, and higher-throughput onboard inference, the NVIDIA Jetson platform is considered the preferred solution. This is due to its stronger parallel-computing capability, integrated GPU acceleration, and better suitability for sustained real-time execution of computationally demanding AI pipelines. Accordingly, the architecture distinguishes between an initial prototyping configuration based on Raspberry Pi 5 and an advanced deployment configuration based on Jetson for higher-performance onboard autonomy.

The perception subsystem is governed by a sensor fusion and processing stack, which integrates inertial data, visual inputs, and LiDAR measurements via an Extended Kalman Filter (EKF). This approach yields a unified spatial representation that supports stable SLAM, visual odometry, and real-time obstacle mapping, as demonstrated in the multi-sensor fusion literature. EKF-based fusion ensures temporal alignment and robust estimation even under partial sensor degradation, which is critical in contested or GPS-denied tactical missions. The adoption of an EKF-based fusion framework is thus aligned with the Operational Requirement for resilient state estimation under partial sensor degradation, ensuring navigation robustness in contested or GPS-degraded environments.

Low-level stabilization and actuation are managed by a Pixhawk 4 flight controller (Holybro, Hong Kong, China) running PX4 firmware (PX4 Autopilot open-source project), a widely adopted open-source autopilot architecture recognized for reliability, configurability, and compatibility with ROS 2 (Open Robotics, Mountain View, CA, USA). Through MAVROS (ROS package for MAVLink communication), high-level trajectory commands, mission updates, and perception-driven adjustments are communicated seamlessly to the flight controller. PX4 executes PID-based control loops and integrates inertial and visual measurements to maintain precise flight dynamics. Communication between the autonomy stack and the flight controller is performed via MAVROS over the MAVLink protocol, the de facto standard in both industrial and research UAV ecosystems, ensuring interoperability and satisfying the Operational Requirement for deterministic and certifiable flight-control behavior.

Communication is supported by a dual-link system combining the DJI O4 digital video link with the RFD900ux long-range telemetry radio (RFDesign Pty Ltd., Brisbane, Australia). This architecture separates high-bandwidth video streaming from control-critical telemetry, providing redundancy and resilience against interference. ROS 2's DDS middleware orchestrates inter-node communication, ensuring deterministic data distribution and supporting potential mesh networking for multi-UAV collaboration. This dual-link design is selected to satisfy Operational Requirements on link robustness and low-latency command-and-control, by decoupling control-critical telemetry from high-bandwidth EO video traffic and increasing resilience under RF interference or contested conditions.

A Ground Control Station (GCS)—typically a ruggedized tablet or laptop running QGroundControl (Dronecode Foundation open-source ground control software)—facilitates mission supervision, waypoint management, and visualization of telemetry, maps, detections, and system diagnostics. The GCS acts as the operational bridge between autonomous onboard processing and human decision-makers. Although the platform does not claim full STANAG 4586 compliance, the communication structure is intentionally aligned with

its interoperability principles, separating command-and-control messaging from payload and telemetry channels and supporting integration with standardized UAV-GCS interfaces commonly used in defense-oriented environments.

Power is supplied by 3S/4S LiPo batteries, managed through a smart Battery Management System (BMS) and a redundant Power Distribution Board (PDB). The choice of LiPo batteries is motivated by their favorable compromise between low-mass, high-discharge capability, and practical integration on compact UAV platforms operating under strict SWaP-C constraints. In small and medium UAV applications, lithium-based batteries are widely used because they provide high energy at a relatively low weight, while remaining suitable for propulsion peaks and onboard payload demands. This makes LiPo-based solutions particularly appropriate for ISR configurations that require short-duration high-power bursts during take-off, maneuvering, onboard computing, and payload operation [24].

At the same time, the use of LiPo batteries reflects a design trade-off. Although lithium battery systems are well suited to compact UAVs, their limited energy density compared with hybrid or fuel cell-supported architectures constrains endurance and payload growth. For this reason, the proposed platform adopts LiPo batteries not as a universal optimum, but as a practical choice for a compact tactical ISR UAV where responsiveness, modularity, and weight efficiency are prioritized over very-long-endurance operation [25].

Within this architecture, the PDB allocates power dynamically, prioritizing critical components such as flight control and AI processing, while the BMS provides cell-level monitoring to prevent over-discharge, imbalance, or thermal risk, thereby supporting operational reliability during ISR missions [25].

The airframe combines carbon-fiber structural elements with lightweight, 3D-printed components for sensor mounting and payload modularity. Vibration isolation mechanisms protect IMUs and cameras from high-frequency disturbances, improving SLAM performance and object detection accuracy. The modularity of the frame allows rapid mission-specific reconfiguration, sensor replacement, or incremental upgrades [9].

3.2. System Integration and Architecture

The system architecture follows a layered, modular design that tightly integrates sensing, perception, control, and communication subsystems. High-speed data links and deterministic scheduling ensure that sensor measurements are processed in real time and consistently delivered to AI inference pipelines and control loops. This architecture reflects emerging standards in autonomous ISR UAV design, emphasizing robustness, redundancy, and scalability.

Operationally, the system processes data in a staged sequence: sensor acquisition (EO, LWIR, LiDAR, ultrasonic/ToF, GNSS/IMU), pre-processing and synchronization, AI-based perception and target detection, multi-sensor fusion and state estimation, decision-support generation, and finally command transmission through MAVROS to the PX4 flight controller. This staged flow clarifies how raw sensor data are progressively transformed into navigation-relevant and mission-relevant outputs.

The physical placement of the main onboard components is illustrated in Figure 2, which shows the distribution of sensing, computing, communication, and power modules on the UAV platform. As shown in Figure 2, the sensing payload is concentrated in the forward section of the platform, while the computing, flight-control, communication, and power-management modules are distributed around the central mass to preserve balance and modularity.

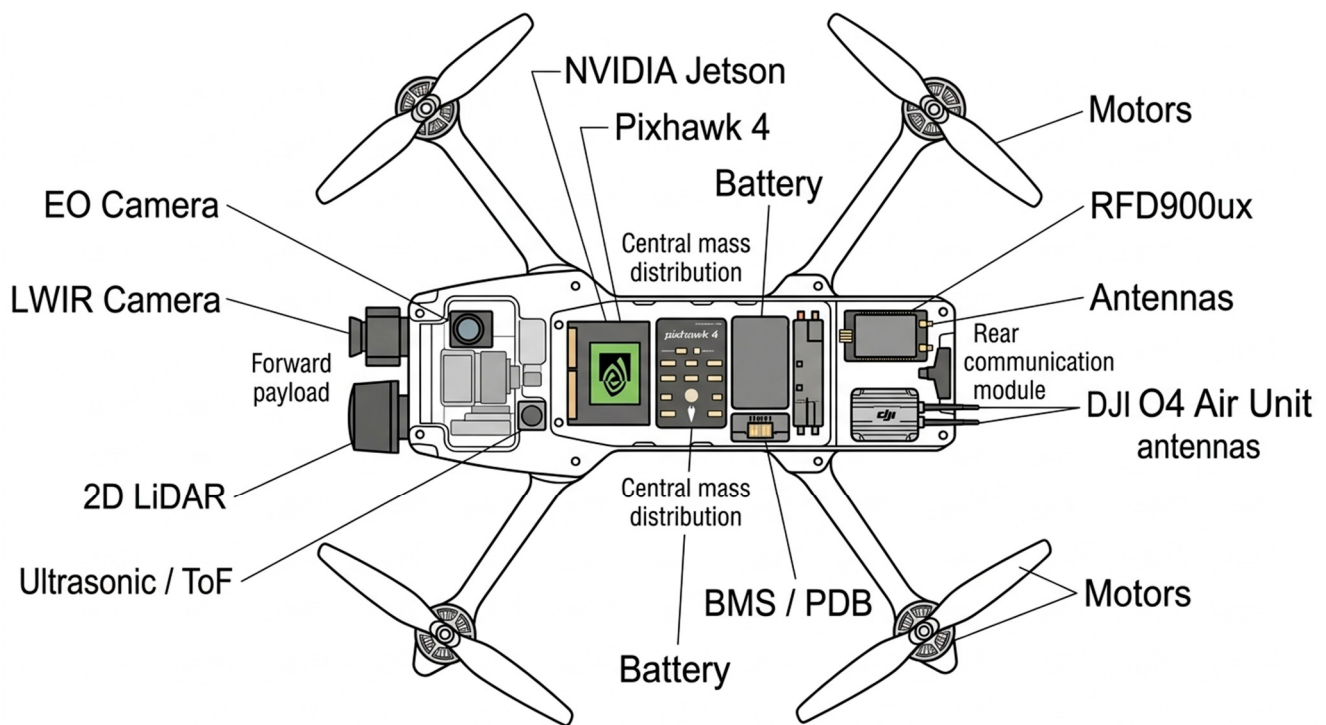


Figure 2. Physical layout of the UAV platform and component placement.

EO data acquired through the DJI O4 module flows directly into the onboard AI processor, where real-time inference is executed. The pipeline includes frame pre-processing, neural inference, and post-processing for bounding-box refinement and classification confidence estimation. This enables the identification of tactical targets such as personnel, vehicles, defensive structures, and unmanned platforms. In parallel, LWIR imagery provides complementary thermal cues, allowing for segmentation and detection even under low-visibility or deception conditions. Fusing EO and IR modalities significantly enhances robustness, as documented in recent EO–IR fusion studies.

Spatial perception is driven by the 2D LiDAR, which continuously generates ranging data used to infer environmental structure. These measurements are fused with visual-inertial odometry in an EKF-based framework, producing a temporally consistent estimate of UAV pose and environment geometry. This fused output supports local mapping, reactive obstacle avoidance, and trajectory planning. Such multi-sensor fusion approaches are well established in modern UAV navigation literature and contribute critically to operations in degraded or GPS-denied environments.

The onboard AI unit orchestrates the perception–action cycle. All processing operations—deep inference, SLAM, sensor fusion, and mission logic—are encapsulated within ROS 2 nodes. The use of containerized nodes ensures fault isolation, modularity, and scalability, enabling rapid deployment of updated models or new perception pipelines. ROS 2’s DDS backbone provides deterministic communication between nodes, and its use of namespaces supports future extensions to multi-UAV collaborative missions.

While the current system emphasizes data-driven autonomy, compatibility with Ansys SCADE (Ansys Inc., Canonsburg, PA, USA) enables the integration of formally verified control logic. Safety-critical behaviors—such as emergency shutdown, fail-safe transitions, or actuator diagnostics—can be modeled and validated through SCADE, enhancing certifiability and trustworthiness, particularly in defense applications where system integrity is paramount.

At the control layer, Pixhawk 4 receives trajectory updates and transmits telemetry through MAVROS. Visual-inertial odometry enhances navigation precision, while PX4’s

flight management ensures stable execution of autonomous commands. The system supports seamless switching between manual, stabilized, and autonomous modes, enabling human override when required by mission constraints.

Communication between the UAV and GCS is maintained through redundant channels. High-bandwidth EO video is streamed via the DJI O4 link, while telemetry, mission data, and AI-generated alerts are transmitted using encrypted RFD900ux communication. This dual-link architecture mitigates the risk of signal degradation or jamming and aligns with modern ISR communication strategies that emphasize link redundancy and security.

Energy management plays a pivotal role in sustaining real-time AI workloads during flight. The BMS monitors cell status and communicates energy metrics to the AI unit, allowing adaptive throttling of non-critical computations when necessary. This ensures that essential perception and control tasks retain priority even under high-load or low-power conditions.

To support the development of stabilization algorithms, the system incorporates a dedicated experimental rig. Built using an Arduino Mega 2560 (Arduino S.r.l., Monza, Italy), an MPU-9050 IMU (TDK InvenSense, San Jose, CA, USA), and three servomotors aligned to simulate roll, pitch, and yaw disturbances, this platform enables controlled evaluation of inertial compensation strategies. Threshold filtering and adaptive correction weighting were tested to mitigate drift and vibration, demonstrating the feasibility of lightweight, low-cost stabilization methods. Although not yet integrated into the main UAV, this experimentation framework offers valuable insights into gimbal-less stabilization approaches suitable for compact or disposable tactical drones.

In summary, the proposed system architecture integrates state-of-the-art perception, control, and communication technologies into a unified platform tailored for autonomous ISR missions. By coupling multi-modal sensing with onboard AI, resilient communications, and robust navigation, the system is capable of generating actionable intelligence in dynamic, constrained, and contested environments. Its modular design facilitates future upgrades, multi-UAV cooperation, and the integration of additional sensing modalities, positioning the platform as a foundation for next-generation tactical reconnaissance systems.

4. System Diagram and Communication Flow

The architecture of the proposed UAV reconnaissance system is organized into a layered, modular structure that reflects contemporary design practices for autonomous ISR platforms. Each subsystem—ranging from sensing and perception to flight control, communication, and collaborative networking—plays a dedicated role within the broader data flow that enables real-time situational awareness and tactical decision-making. The overall design prioritizes low-latency processing, robustness under contested conditions, and the ability to scale toward multi-UAV operational modes. A high-level representation of the system architecture and inter-module communication is illustrated in Figure 3, which highlights the interaction between onboard components, the Ground Control Station (GCS), and potential swarm networks:

The architecture diagram adopts a layered representation in which each horizontal band corresponds to a functional domain of the UAV system (sensing, perception, flight control, communication, energy management, and ground-segment interaction). Rectangular boxes denote functional subsystems, while grouping within the same band indicates that the modules operate at the same architectural level and exchange information within that domain.

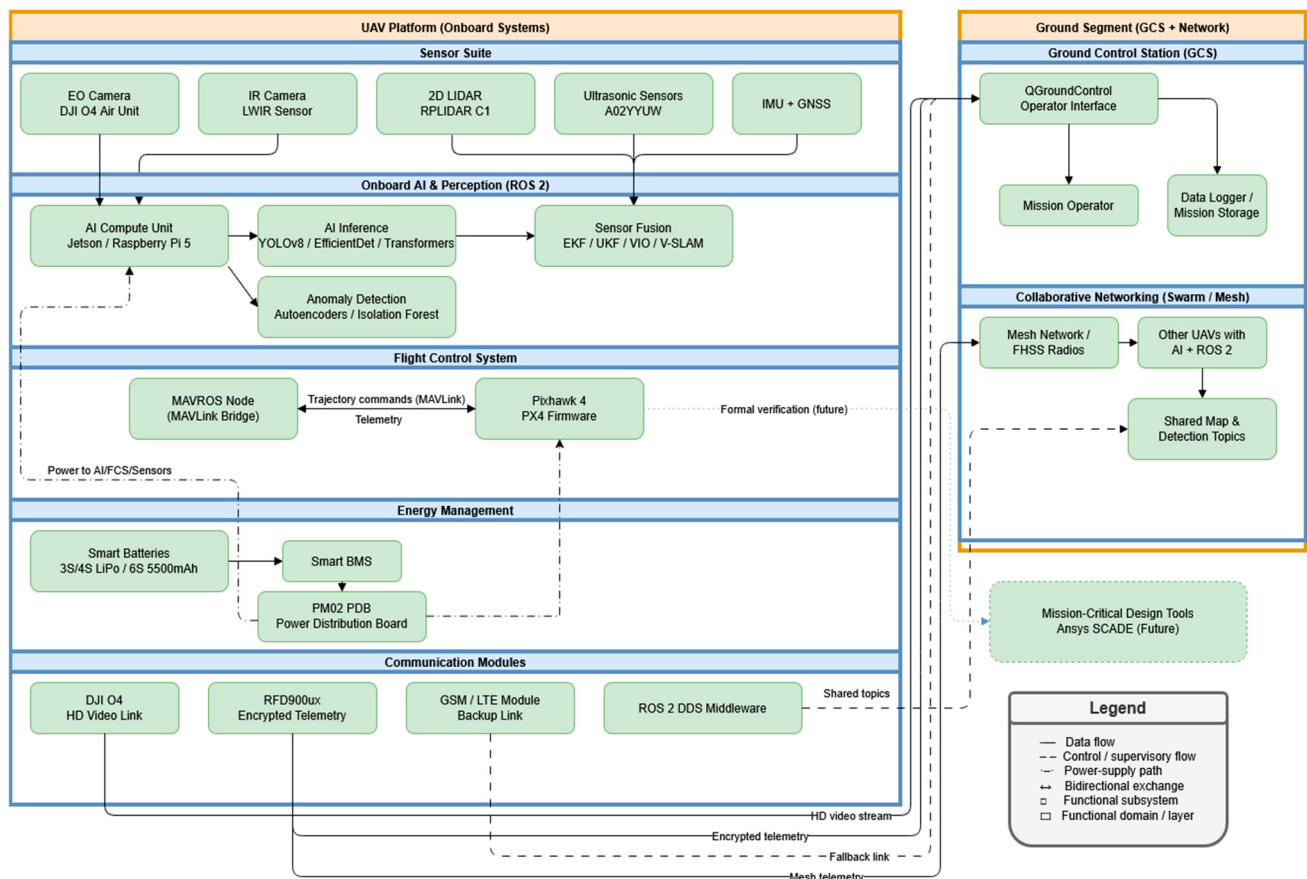


Figure 3. System architecture and communication flow.

As shown in Figure 3, the architecture differentiates between three categories of links, each with a distinct visual convention. Solid arrows represent data-flow connections (for example sensor streams, ROS 2 topics, fused-state outputs, or AI inference results). Dashed arrows represent control-flow or supervisory interactions (such as trajectory commands, MAVLink telemetry, or operator-level mission updates). Dot-dash arrows indicate power-supply paths between the Energy Management subsystem and autonomy-critical components; these links correspond to electrical power rails and do not represent data or control interfaces. Bidirectional arrows denote continuous exchange, while unidirectional arrows indicate producer–consumer flow.

In Figure 3, the left side represents the onboard UAV subsystems, whereas the right-hand side corresponds to the Ground Segment and collaborative networking elements. Vertical alignment across layers reflects the end-to-end propagation of information and resources—from sensing and perception, through flight-control execution, to communication, operator interaction, and potential multi-UAV cooperation—following the logical progression of decision-making within the system.

4.1. Functional Domains of the System Architecture

The end-to-end data flow of the UAV is structured across seven functional domains that collectively ensure mission autonomy and operational resilience. These include (1) sensor data acquisition, (2) multi-sensor fusion and AI-based perception, (3) flight control and trajectory adaptation, (4) secure communication pipelines, (5) operator interaction through the GCS, (6) energy monitoring and power allocation, and (7) collaborative networking for swarm capabilities. Although conceptually distinct, these domains operate concurrently,

with continuous bidirectional information exchange mediated through ROS 2 middleware and PX4 autopilot interfaces.

Table 1 summarizes these seven functional domains in structured form, together with their representative inputs, main processing functions, and mission-level outputs, and relates them to the architecture and data-flow stages presented earlier in Figure 3 and in the associated system-flow description. More specifically, the sensor data acquisition domain corresponds to the sensing layer of the architecture, the multi-sensor fusion and AI-based perception domain reflects the onboard processing chain, the flight-control and trajectory-adaptation domain corresponds to the MAVROS–PX4 control path, and the remaining domains synthesize the communication, operator interaction, energy management, and collaborative networking branches shown in the overall system diagram.

Table 1. Functional domains, representative inputs, processing functions, and mission-level outputs of the proposed UAV ISR system.

Functional Domain	Representative Inputs	Main Processing/Method	Mission-Level Outputs
Sensor data acquisition	EO video, LWIR frames, LiDAR scans, ultrasonic/ToF measurements, GNSS/IMU data	Sensor interfacing, acquisition, timestamping, synchronization	Time-aligned raw sensory data for perception and navigation
Multi-sensor fusion and AI-based perception	EO/LWIR imagery, LiDAR data, inertial measurements	Pre-processing, EO–IR fusion, EKF/UKF-based state estimation, VIO/V-SLAM, object detection, tracking, anomaly detection	Target detections, fused state estimate, local environmental representation, anomaly cues
Flight control and trajectory adaptation	Fused navigation state, AI-generated guidance cues, operator commands	MAVROS communication, MAVLink exchange, PX4 control loops, waypoint and trajectory correction	Stabilization, obstacle-aware trajectory adaptation, mission execution
Secure communication pipelines	Video stream, telemetry packets, detections, mission updates	Link separation, encryption, redundancy handling, fallback communication	Robust operator connectivity, secure telemetry, resilient C2 exchange
Operator interaction through GCS	Annotated video, telemetry, SLAM maps, AI alerts, mission status	Visualization, supervision, waypoint editing, geofence update, manual override	Mission supervision, operator decision support, command refinement
Energy monitoring and power allocation	Battery voltage, current, temperature, subsystem power demand	BMS monitoring, PDB distribution, load prioritization, adaptive throttling of non-critical tasks	Protected critical subsystems, endurance preservation, safe energy usage
Collaborative networking for swarm capabilities	Detection metadata, shared map fragments, mission status, task updates	DDS-based data exchange, mesh coordination, distributed situational sharing	Multi-UAV awareness, task sharing, scalable surveillance coverage

4.2. System Architecture and Data Flow

The system begins with data acquisition, performed by a heterogeneous sensor suite composed of an EO camera (DJI O4 Air Unit), an LWIR module, a 2D LiDAR (RPLIDAR C1 <Slamtec, Shanghai, China>), ultrasonic rangefinders (A02YYUW), and inertial/GNSS units. These sensors were selected to cover complementary perceptual modalities: visual imagery for object detection and scene understanding, thermal signatures for low-visibility target detection, planar ranging for spatial mapping, and inertial cues for attitude estimation. All raw data streams are timestamped and synchronized using ROS 2 time and onboard GNSS, ensuring temporal consistency across modalities.

Upon acquisition, sensory inputs enter the AI and perception stack, executed on a Jetson or Raspberry Pi 5 compute unit. EO and IR imagery undergo hardware-accelerated pre-processing operations including frame stabilization, deblurring, and adaptive contrast

normalization. LiDAR scans and ultrasonic measurements are fused with visual-inertial odometry (VIO) using Extended and Unscented Kalman Filters (EKF, UKF), producing a unified spatial representation consistent with recent approaches to multi-sensor UAV state estimation. The perception module further integrates V-SLAM pipelines to maintain real-time mapping and localization performance even in GPS-denied environments.

As illustrated in Figure 3, the proposed architecture follows a staged processing chain: sensor data acquisition, timestamping and synchronization, pre-processing of EO/IR and ranging data, multi-sensor fusion and state estimation, AI-based detection and tracking, mission-level decision support, and finally command transmission through MAVROS to the PX4 flight controller. This sequence clarifies how raw sensor data are progressively transformed into navigation-relevant and mission-relevant outputs within the proposed ISR framework.

Following fusion, the system activates multiple AI inference pipelines. Deep neural detectors such as YOLOv8, EfficientDet, and transformer-based architectures perform detection and tracking of tactical objects including vehicles, personnel, weapon systems, and defensive structures [10,26].

Supplementary classification networks refine object categories, while anomaly detection modules—based on autoencoders and Isolation Forests—identify deviations from normal patterns, enabling early indicators of hostile activity or unexpected environmental features. This layered perception strategy enhances robustness against noise, occlusion, and adversarial conditions.

Outputs from perception directly inform the flight control system, composed of the MAVROS bridge and the Pixhawk 4 autopilot running PX4 firmware. MAVROS publishes trajectory corrections and receives telemetry updates encapsulated in MAVLink messages. PX4's internal control loops, operating at up to 400 Hz, integrate sensor fusion outputs and AI-driven guidance signals to achieve precise stabilization, obstacle avoidance, and mission-level navigation. This tight coupling between perception and control is essential for ISR operations requiring dynamic repositioning, target tracking, or rapid threat avoidance [27].

4.3. Communication Infrastructure and Ground Control Integration

As illustrated in Figure 3, communication is distributed across redundant channels to increase mission robustness. The DJI O4 digital link handles high-definition video transmission, ensuring minimal delay for operator monitoring. Concurrently, the RFD900ux provides encrypted long-range telemetry, which includes system status, MAVLink messages, and AI-generated detections. A GSM/LTE backup link can be activated when line-of-sight communication is degraded, enabling mission continuity under adverse radio conditions. All telemetry is secured through AES-256 encryption layers.

At the operational interface, the Ground Control Station (GCS)—running QGroundControl—displays annotated video feeds, telemetry, real-time SLAM maps, AI detections, and status indicators. Operators can modify mission parameters, upload waypoints, adjust geofences, or enable automated behaviors. Additionally, the GCS integrates with data logging systems that store flight footage, detection results, and mission metadata for intelligence analysis, retraining of AI models, and post-mission evaluation.

4.4. Energy Management and Power Allocation

As shown in the energy management branch of Figure 3, sustaining continuous AI inference during flight imposes significant energy demands. To address this, the UAV integrates smart 3S/4S LiPo batteries, a dedicated Smart BMS, and a PM02 Power Distribution Board. The BMS monitors voltage, current, and thermal conditions, transmitting power metrics to the AI compute unit. These metrics feed an energy prediction model that

dynamically adjusts computation loads—such as temporarily reducing inference frequency or disabling non-critical sensors—to extend mission endurance.

The PDB provides redundancy, enabling seamless transitions between power rails and ensuring constant supply to autonomy-critical modules including the flight controller and perception pipelines.

At the current stage of the work, the energy management layer is presented at the architectural level, with emphasis on battery monitoring, regulated power distribution, and adaptive reduction in non-critical processing loads. A detailed quantitative evaluation of subsystem-level power consumption and endurance impact will be addressed in future implementation and validation stages.

4.5. Collaborative Networking and Swarm-Capable Extensions

As indicated in Figure 3, beyond single-UAV operations, the architecture supports collaborative networking, a capability increasingly emphasized in modern ISR doctrine. ROS 2 DDS middleware can be extended into a mesh topology that enables multiple UAVs to exchange fused maps, detection results, and mission updates. In such configurations, UAVs can divide surveillance areas, share real-time target information, and maintain operational coherence even if one platform becomes disabled. Other UAVs equipped with equivalent AI and ROS 2 stacks can join or leave the mesh dynamically, enabling adaptive and resilient multi-agent behaviors.

4.6. Integration with Mission-Critical Design Tools

As reflected in the safety-oriented architectural path of the proposed system, while the current implementation focuses on operational deployment, future integration with mission-critical design tools such as Ansys SCADE is planned. SCADE's formal verification frameworks allow rigorous validation of safety-critical components, including recognition gating, emergency landing logic, power-failure responses, and fault diagnostics. Such verification pathways enhance certifiability, security, and reliability—key requirements for ISR systems operating in adversarial environments.

4.7. Summary of Architectural Coherence

In summary, the system architecture constitutes a comprehensive framework that tightly integrates sensing, fusion, AI perception, flight control, communication, energy management, and collaborative networking. The layered design—illustrated in Figure 3—enables real-time intelligence extraction and autonomous navigation in complex, contested environments.

Its modular structure ensures extensibility toward future enhancements, including swarm-based surveillance, additional sensing modalities, or formally verified mission-critical behaviors. This positions the UAV platform as a versatile and resilient ISR asset capable of supporting a diverse range of defense and security operations.

5. AI Architecture and Cognitive Functions in the ISR System

Artificial Intelligence constitutes the cognitive core of the UAV reconnaissance platform, transforming raw sensory input into actionable intelligence and mission-relevant insights. Unlike traditional ISR architectures in which perception and decision-making remain largely operator-driven, the proposed system employs AI to enhance mission efficiency, reduce operator cognitive load, and ensure consistent performance under variable or degraded environmental conditions. Within this framework, AI does not operate as a fully autonomous agent; instead, it assumes the role of an “assisted autonomy layer,” supporting an operator-supervised workflow in which human judgment remains central to mission execution.

Figure 4 provides a detailed functional view of the internal artificial intelligence processing chain within the proposed UAV architecture. Unlike Figure 3, which presents the overall system architecture and communication flow between subsystems, Figure 4 focuses specifically on the perception, tracking, prioritization, and operator-supervised decision-support stages implemented within the onboard AI layer. This structure therefore reflects a functional decomposition of the cognitive pipeline rather than a repetition of the global system architecture.

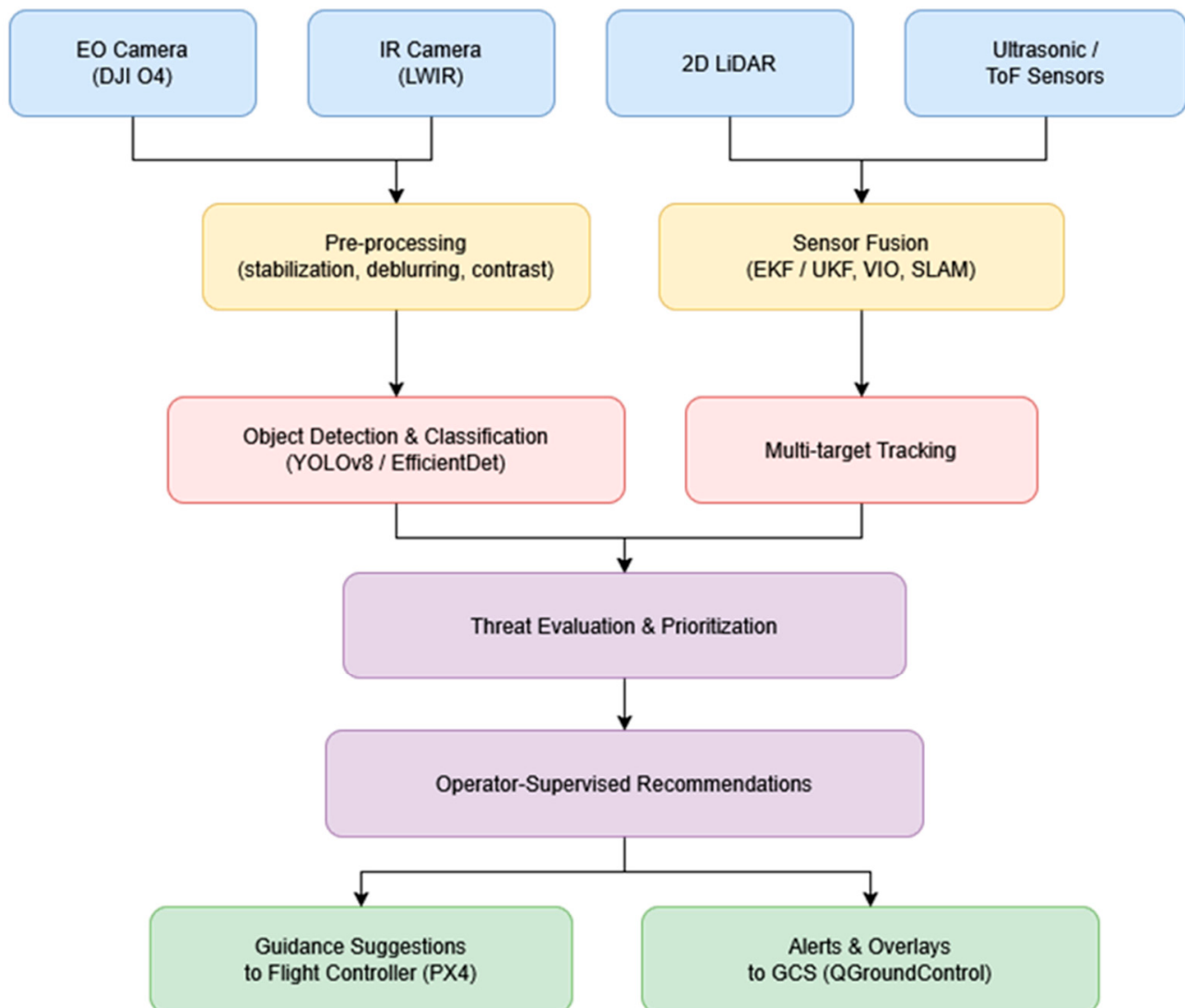


Figure 4. AI-Based Perception and Decision-Making Pipeline for the UAV ISR Platform.

In this sense, Figure 4 should be interpreted as a detailed expansion of the onboard AI and perception block introduced earlier in Figure 3.

The perception layer of the AI system is responsible for extracting meaningful information from the electro-optical and infrared sensing modalities. Real-time object detection models identify and classify personnel and vehicles—two of the most tactically relevant categories in reconnaissance missions—while feature extractors and spatiotemporal trackers maintain persistent awareness of moving targets across variable viewpoints. Thermal cues from the LWIR sensor contribute to human detection under camouflage, partial concealment, or nighttime conditions, complementing EO-based classifiers and reducing ambiguity in uncertain environments. This multi-modal perceptual strategy ensures resilient detection

capabilities even when visual conditions degrade, a challenge commonly encountered in real-world ISR scenarios.

Beyond static detection, the AI system incorporates a lightweight threat evaluation framework. Rather than issuing autonomous engagement or navigation directives, this mechanism estimates the relevance, persistence, and potential trajectory of detected entities, providing operators with prioritized situational cues. For instance, recurrent detections of a vehicle within a restricted perimeter or prolonged human movement near a strategic asset are scored and elevated to the operator interface. This prioritization mechanism supports rapid decision-making and reduces the risk of information overload—one of the core limitations of modern ISR workflows characterized by dense sensor data streams.

As illustrated by the sensing, fusion, and onboard processing stages shown earlier in Figures 3 and 4, the notion of high-volume sensory input in the proposed architecture refers to the continuous inflow of EO frames, LWIR imagery, LiDAR scans, proximity measurements, navigation states, and time-persistent detections generated during flight. To reduce operator overload, the AI layer is intended to filter this information using criteria such as detection confidence, temporal persistence across consecutive frames, target class relevance, proximity to areas of interest or restricted zones, and anomaly related behavioral cues. Mission-relevant events are then prioritized according to their estimated tactical significance, recurrence over time, spatial context, and possible threat implications. Compared with conventional operator-driven ISR workflows, in which most interpretation and prioritization tasks remain manual, the proposed framework is intended to forward a reduced and ranked subset of detections and guidance cues for operator review and supervisory decision-making.

To further clarify the internal logic of the lightweight threat assessment framework, Figure 5 illustrates its schematic workflow. As shown, the framework receives detection- and context-related inputs from the perception layer, extracts prioritization cues such as detection confidence, temporal persistence, target category, spatial relevance, and anomaly related indicators, and then performs rule-guided assessment before assigning a relative priority level to the detected event. The resulting ranked alerts and prioritized situational cues are subsequently forwarded to the operator interface for supervisory decision-making.

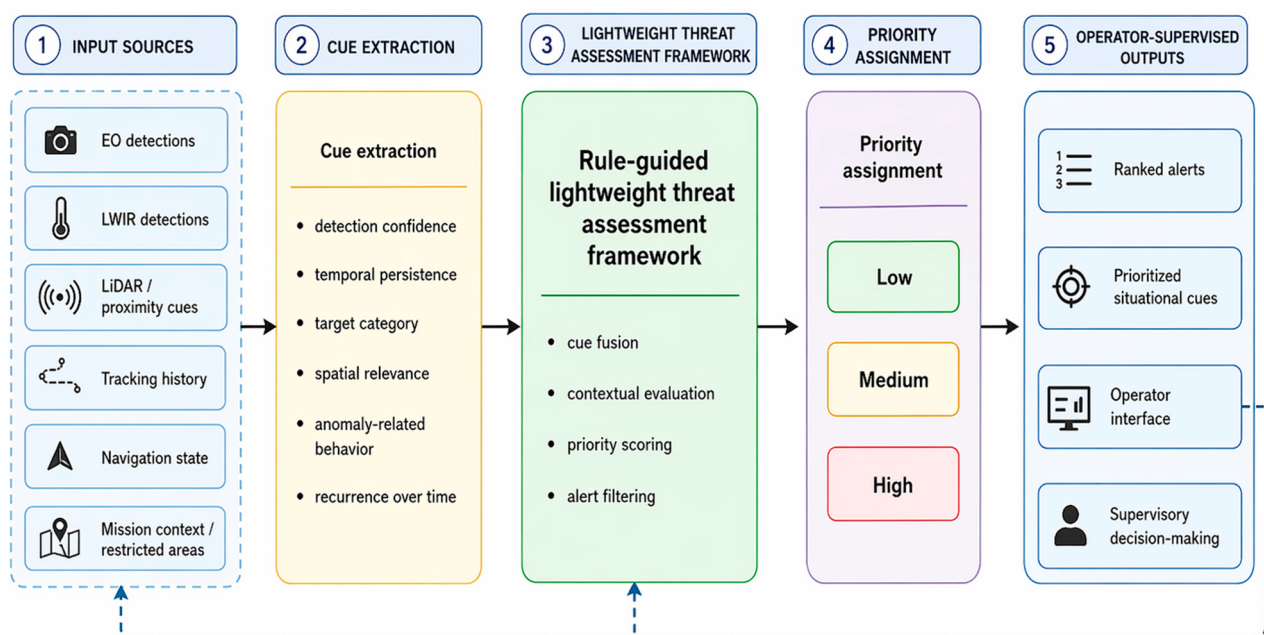


Figure 5. Schematic workflow of the lightweight threat assessment framework.

In the proposed ISR architecture, the lightweight threat assessment framework is conceived as a rule-guided prioritization layer rather than a fully autonomous threat-classification engine. Its role is to convert multiple low-level perception outputs into operator-relevant alerts by combining several cues, including detection confidence, temporal persistence across consecutive frames, target category, proximity to protected or restricted areas, anomaly related behavioral indicators, and recurrence over time. Based on these inputs, the framework is intended to assign a relative priority level to detected events and to forward only the most relevant cues to the operator interface. In this way, isolated or low-confidence detections can be deprioritized, while repeated or context-sensitive detections can be highlighted for supervisory decision-making.

The AI module also contributes to navigation by providing contextual awareness to the flight controller. While the UAV retains operator supervision and preserves manual override as the highest authority, AI-assisted navigation continuously evaluates environmental structure, obstacle distribution, and potential occlusion patterns to recommend viewpoint repositioning or surveillance routes. These recommendations are non-binding but significantly aid in maintaining optimal geometry for detection tasks, especially when tracking moving objects or when operating in cluttered or unpredictable environments. By coupling perception-derived insights with route optimization heuristics, the system enhances mission efficiency without compromising operator authority.

Energy-aware computation is another function embedded within the AI architecture. The system monitors real-time consumption metrics provided by the BMS and dynamically refines non-critical workloads such as inference frequency, tracking update rates, or the activation of secondary perception modules. This adaptive management framework extends operational endurance while ensuring that essential ISR capabilities remain fully functional. It also supports future mission configurations where limited-energy platforms or long-duration loitering missions require intelligent load balancing.

A defining component of this research lies in exploring how safety-critical elements of the AI-control interface can be formalized using tools such as Ansys SCADE. Unlike perception models, which rely on probabilistic inference and machine learning, mission-critical behaviors—such as transitions to emergency modes, constraints on UAV positioning near obstacles, or operational limits when communication links degrade—must adhere to deterministic safety rules. SCADE allows these rules to be specified, verified, and integrated in a manner compatible with certification paths common in aerospace systems. Within the operator-supervised autonomy paradigm, this separation ensures that while AI may analyze, recommend, and infer, the final gating of system-critical actions passes through a formally validated decision layer.

In the proposed architecture, the Ansys SCADE-oriented layer is intended to supervise deterministic, mission-critical control logic rather than probabilistic AI inference itself. More specifically, this layer is envisaged to validate or reject AI-generated guidance commands against explicit safety constraints such as geofence violations, minimum safe altitude conditions, obstacle-proximity limits, communication-loss fallback logic, and emergency return or landing triggers. In this sense, the safety layer acts as a rule-based supervisory barrier between the AI decision module and the PX4 autopilot. At the current stage of the work, this functionality is defined at the architectural level; detailed implementation, verification workflow, and formal safety assurance allocation remain part of future system development and validation.

To visualize how the deterministic safety layer interacts with AI-generated commands, Figure 6 presents the proposed SCADE-based safety gate. This component sits between the AI decision module and the PX4 autopilot, enforcing mission rules and emergency procedures before any maneuver is executed.

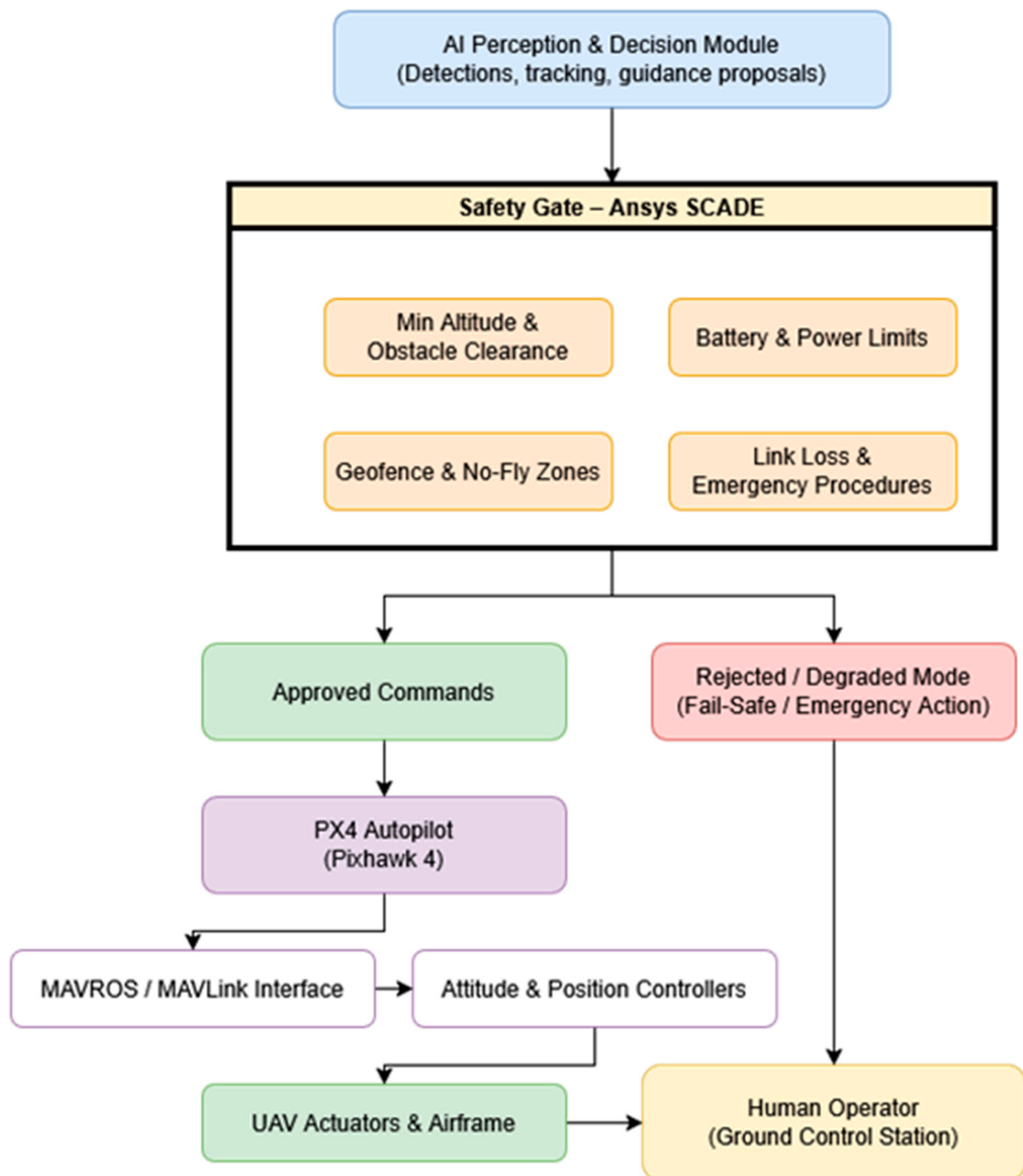


Figure 6. Safety Gate Architecture Using Ansys SCADE for Mission-Critical Validation.

As shown in Figure 6, the safety gate acts as a safety-verification layer that separates probabilistic AI behavior from mission-critical actuation. By validating or rejecting AI proposals, the SCADE module ensures that the UAV remains within certified operational limits, preventing unsafe actions even in environments with degraded sensor data, adversarial interference, or AI misclassification. This hybrid architecture—AI assistance combined with deterministic safety gating—forms a central contribution of the proposed research.

In this architecture, AI is positioned not as a replacement for human decision-making but as a force multiplier that enhances ISR effectiveness. It filters high-volume sensory input, extracts and ranks mission-relevant events, guides the UAV toward advantageous observation positions, and optimizes energy consumption, all while maintaining a transparent

and predictable operational envelope. The integration of formal verification tools further strengthens system reliability, ensuring that AI-driven behaviors remain bounded, interpretable, and compatible with safety constraints expected in defense-oriented applications.

As shown in Table 2, the literature in related UAV detection, tracking, and energy-aware navigation tasks reports measurable gains in detection performance, real-time operation, and energy efficiency when AI-based perception and adaptive planning are introduced. In the present manuscript, these results are not claimed as direct performance outcomes of the proposed system, but are used to position the expected advantages of the proposed architecture with respect to established results in the literature.

Table 2. Representative literature-based performance indicators relevant to the proposed AI-assisted ISR UAV architecture.

Reference	Function/Task	Reported Metric(s)	Relevance to the Proposed Architecture
Tan et al., 2020 (EfficientDet) [26]	Efficient object detection	EfficientDet-D7 achieved 52.2 AP on COCO test-dev; EfficientDet-D6 reported 190 ms GPU latency	Supports the choice of efficient detectors for onboard perception under resource constraints
Xu et al., 2024 [15]	UAV intrusion detection and tracking	Embedded deployment on Jetson TX2/Xavier reported a speed improvement of more than 10× after optimization, with negligible loss of accuracy	Supports the feasibility of real-time UAV-based AI detection and tracking in field conditions
Yilmaz, 2024 [15]	YOLOv8-based drone detection	F1 = 0.588; recall = 0.97; precision = 0.832	Illustrates representative detector behavior for drone-target detection with strong recall
Ahmed and Sheltami, 2024 [11]	Energy-aware UAV path planning	Average normalized energy evaluated over 30 runs; smoothed paths showed lower normalized energy than non-smoothed paths; dynamic environments consumed more than static ones	Supports the inclusion of energy-aware workload and path-planning logic in the proposed ISR architecture
Present work	Requirement-driven AI-assisted ISR UAV architecture	Architecture-level contribution; no experimental metrics at current stage	Integrates multi-modal sensing, operator-supervised autonomy, resilient communication, and safety-oriented validation within a unified UAV ISR framework

Through this combination of perception, contextual analysis, operator-assist functionality, and safety-aware integration, the AI module serves as a foundational component that elevates the UAV from a sensor platform to an intelligent reconnaissance system capable of supporting complex missions in contested and dynamic environments.

6. Conclusions

This paper presented an AI-driven unmanned aerial vehicle architecture intended for tactical intelligence, surveillance, and reconnaissance missions. The study focused on the system-level integration of multi-modal sensing, onboard AI-assisted perception, resilient communication, and safety-oriented supervisory logic within a unified ISR framework. The proposed architecture was defined from operational and sensor-performance requirements and was structured to support target detection, tracking, environment awareness, and mission-level decision support under degraded visibility, contested communication conditions, and constrained onboard energy budgets. At its current stage, the work remains

architectural and conceptual, but it provides a coherent basis for future implementation, validation, and progressive extension toward collaborative multi-UAV operation.

The main features of the proposed system can be summarized as follows:

- The integration of electro-optical, long-wave infrared, LiDAR, inertial, communication, and flight-control subsystems within a unified UAV ISR architecture;
- Operator-supervised autonomy supported by AI-based perception, prioritization, and decision-support functions;
- A safety-oriented validation path intended for future integration with Ansys SCADE in order to supervise mission-critical behaviors;
- A modular and scalable ROS 2–PX4-based framework suitable for future implementation, validation, and multi-UAV extension.

7. Future Research Directions and Main Contributions

The main contributions of the proposed work can be summarized as follows:

- A requirement-driven UAV ISR architecture integrating electro-optical, long-wave infrared, and LiDAR sensing for multi-modal perception in complex operational environments;
- An operator-supervised autonomy concept that combines AI-assisted perception with a safety-oriented validation layer intended for future integration with Ansys SCADE;
- A scalable ROS 2–PX4-based system architecture designed to support real-time sensing, perception, communication, and future multi-UAV extension.

Future work will focus on the implementation and validation of the proposed architecture in several complementary directions. A first priority concerns the development of an AI-assisted perception layer capable of detecting, classifying, and tracking tactically relevant targets, especially personnel and vehicles, under varying illumination, visibility, and environmental conditions. Particular emphasis will be placed on EO–IR fusion, low-visibility robustness, and resource-constrained onboard inference.

A second research direction concerns the refinement of the operator-supervised autonomy layer. This includes the development of mechanisms for target prioritization, contextual threat evaluation, and viewpoint optimization, with the objective of improving decision support while preserving human authority over mission-critical actions. In this sense, the architecture is intended to support human–machine teaming rather than unrestricted autonomous behavior.

A third direction involves the safety-critical dimension of the system. Future implementation will examine the use of Ansys SCADE for the modeling and verification of deterministic supervisory functions, including fail-safe transitions, emergency procedures, communication-loss responses, and operational constraints related to obstacle proximity or restricted zones. This step is intended to strengthen the reliability and certifiability of the proposed AI-assisted autonomy framework.

Finally, the proposed architecture will be evaluated progressively through simulation, hardware-in-the-loop testing, and future field experimentation using a physical UAV prototype. Additional work will also investigate the extension of the framework toward collaborative multi-UAV operation through shared detections, exchanged map fragments, and mission-level coordination over ROS 2 DDS-based communication.

Author Contributions: Conceptualization, A.-D.A. and A.N.P.; methodology, A.-D.A. and J.G.; software, A.-D.A.; validation, A.-D.A., A.N.P. and J.G.; formal analysis, A.-D.A.; investigation, A.-D.A.; resources, A.-D.A.; data curation, A.N.P.; writing—original draft preparation, A.-D.A.; writing—review and editing, A.N.P. and J.G.; visualization, J.G.; supervision, A.N.P.; project adminis-

tration, A.N.P.; funding acquisition, A.N.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: Author Jair Gonzalez was employed by the Ansys Inc. (France). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Ku, J.; Harakeh, A.; Waslander, S.L. In defense of classical image processing: Fast depth completion on the CPU. In Proceedings of the 15th Conference on Computer and Robot Vision (CRV), Toronto, ON, Canada, 8–10 May 2018; pp. 16–22. [\[CrossRef\]](#)
2. Mustafa, A.; Khan, S.H.; Hayat, M.; Shen, J.; Shao, L. Image super-resolution as a defense against adversarial attacks. *IEEE Trans. Image Process.* **2019**, *29*, 1711–1724. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Kumar, C.A.; Giridhar, O. UAV detection multi-sensor data fusion. *J. Res. Sci. Eng.* **2024**, *6*, 6–12. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Grieco, G.; Amendola, D.; Anderson, D. *Counter-Drone Systems and Data Fusion*; Final report; Publications Office of the European Union: Luxembourg, 2024. [\[CrossRef\]](#)
5. Liu, S.; Lin, Z.; Huang, W.; Yan, B. Technical development and future prospects of cooperative terminal guidance based on knowledge graph analysis: A review. *J. Zhejiang Univ.-Sci. A* **2025**, *26*, 605–634. [\[CrossRef\]](#)
6. Alagurajan, V.; Rajagopal, S.; Kulkarni, V. Performance metrics of unmanned aerial vehicles from the perspective of ISR applications. *Trans. Indian Natl. Acad. Eng.* **2025**, *10*, 243–255. [\[CrossRef\]](#)
7. Arsene, C. Fusion of real-time thermal image and depth laser readings for remote thermal sensing via UAVs. *arXiv* **2020**, arXiv:2006.01286.
8. Vandersteegen, M.; Vanbeeck, K.; Goedemé, T. Super accurate low-latency object detection on a surveillance UAV. *arXiv* **2019**, arXiv:1904.02024.
9. Katkuri, A.V.R.; Madan, H.; Khatri, N.; Abdul-Qawy, A.; Patnaik, K. Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review. *Array* **2024**, *23*, 100361. [\[CrossRef\]](#)
10. Iqra; Giri, K.J. SO-YOLOv8: A novel deep learning-based approach for small object detection with YOLO beyond COCO. *Expert Syst. Appl.* **2025**, *280*, 127447. [\[CrossRef\]](#)
11. Ahmed, G.; Sheltami, T. Novel energy-aware 3D UAV path planning and collision avoidance using receding horizon and optimization-based control. *Drones* **2024**, *8*, 682. [\[CrossRef\]](#)
12. Motayyeb, S.; Samadzadegan, F.; Dadrass Javan, F.; Hosseinpour, H. Fusion of UAV-based infrared and visible images for thermal leakage map generation of building facades. *Heliyon* **2023**, *9*, e14551. [\[CrossRef\]](#)
13. Tang, P.; Li, J.; Sun, H. A review of electric UAV visual detection and navigation technologies for emergency rescue missions. *Sustainability* **2024**, *16*, 2105. [\[CrossRef\]](#)
14. Pal, O.K.; Shovon, M.S.H.; Mridha, M.F.; Shin, J. In-depth review of AI-enabled unmanned aerial vehicles: Trends, vision, and challenges. *Discov. Artif. Intell.* **2024**, *4*, 97. [\[CrossRef\]](#)
15. Xu, Y.; Liu, Y.; Li, H.; Wang, L.; Ai, J. A deep learning approach of intrusion detection and tracking with UAV-based 360° camera and 3-axis gimbal. *Drones* **2024**, *8*, 68. [\[CrossRef\]](#)
16. Jia, Y.; Ji, Y. Research progress and application of UAV multi-sensor data fusion technology. *MATEC Web Conf.* **2025**, *410*, 04009. [\[CrossRef\]](#)
17. Irfan, M.; Dalai, S.; Vishwakarma, K.; Trslie, P.; Riordan, J.; Dooly, G. Multi-sensor fusion for efficient and robust UAV state estimation. In Proceedings of the 2024 12th International Conference on Control, Mechatronics and Automation (ICCMA), London, UK, 11–13 November 2024.
18. Pereira, A.; Warwick, S.; Moutinho, A.; Suleman, A. Infrared and visible camera integration for detection and tracking of small UAVs: Systematic evaluation. *Drones* **2024**, *8*, 650. [\[CrossRef\]](#)
19. NVIDIA Corporation. *Jetson Orin Whitepaper*; NVIDIA: Santa Clara, CA, USA, 2024. Available online: <https://developer.nvidia.com/embedded/jetson-orin> (accessed on 12 October 2025).
20. Zhang, S.; Gao, H.; Rao, Q. Defense against adversarial attacks by reconstructing images. *IEEE Trans. Image Process.* **2021**, *30*, 6117–6129. [\[CrossRef\]](#)
21. Samaras, S.; Diamantidou, E.; Ataloglou, D.; Sakellariou, N.; Vafeiadis, A.; Magoulaniotis, V.; Lalas, A.; Dimou, A.; Zarpalas, D.; Votis, K.; et al. Deep learning on multi sensor data for counter UAV applications—A systematic review. *Sensors* **2019**, *19*, 4837. [\[CrossRef\]](#)

22. Zolfaghari, B.; Abbasmollaei, M.; Hajizadeh, F.; Yanai, N.; Bibak, K. Secure UAV (drone) and the great promise of AI. *ACM Comput. Surv.* **2024**, *56*, 1–37. [[CrossRef](#)]
23. Johnson, J. Analysis of image forming systems. In *Selected Papers on Infrared Design. Part I and II*; Johnson, R.B., Wolfe, W.L., Eds.; Society of Photo-Optical Instrumentation Engineers (SPIE): Bellingham, WA, USA, 1985; Volume 513, pp. 761–781.
24. Mohsan, S.A.H.; Othman, N.Q.H.; Khan, M.A.; Amjad, H.; Żywiłłek, J. A comprehensive review of micro UAV charging techniques. *Micromachines* **2022**, *13*, 977. [[CrossRef](#)]
25. Xiao, C.; Wang, B.; Zhao, D.; Wang, C. Comprehensive investigation on lithium batteries for electric and hybrid-electric unmanned aerial vehicle applications. *Therm. Sci. Eng. Prog.* **2023**, *38*, 101677. [[CrossRef](#)]
26. Tan, M.; Pang, R.; Le, Q.V. EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 10781–10790.
27. Virginio, R.; Adiprawita, W. Utilization of MAVROS for electro optical seeker subsystem in the design of ROS based architecture for autonomous guided platform. *AIP Conf. Proc.* **2021**, *2366*, 060002. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.