

Architecture Design and Cooperative Mechanism of an Air-Ground Intelligent Perception System

Hongxin Tan^{a,*}, Yuejian Xu^b

Science and Technology on Complex Aviation Systems Simulation Laboratory, Beijing, China

^azll_txx@163.com, ^bxuyuejian_123@163.com

*Corresponding author

Abstract: Air-ground collaborative perception is essential for airport foreign object debris (FOD) detection, key-area patrol, emergency response, and battlefield operational assistance. However, existing systems often remain limited to isolated platforms, linear data processing, or single-algorithm optimization, and therefore struggle with blind-zone compensation, cross-platform confirmation, and degraded operation. This paper proposes an air-ground collaborative intelligent perception system integrating UAV clusters, UGV groups, fixed monitoring stations, edge computing nodes, and a command center. A dual-closed-loop architecture with six functional layers and two support pillars is designed: the data loop supports sensing, registration, fusion, and situation construction, while the control loop supports priority assessment, joint scheduling, route re-planning, and feedback re-sensing. A heterogeneous observation model, a quality-aware fusion model, a coverage utility model, and a joint task-allocation optimization model are established. A discrete-event simulation with 120 Monte Carlo runs, 36 random target events per run, and four baseline schemes is constructed in a 2 km × 2 km area. Under the stated assumptions, the proposed scheme achieves a 94.4% confirmation rate, reduces the mean confirmation delay from 42.02 s to 32.46 s compared with non-adaptive heterogeneous patrol, and increases timely confirmation within 60 s from 66.8% to 78.7%. The results indicate that the value of the architecture lies not only in adding platforms, but also in transforming heterogeneous resources into a closed-loop, reconfigurable, and engineering-deployable perception capability.

Keywords: Air-ground collaboration, Intelligent perception, Uncertainty fusion, Task allocation, UAV and UGV, System reliability

1. Introduction

Airports, industrial parks, border areas, and emergency operation sites require continuous situational awareness over large and dynamic spaces. Fixed monitoring stations provide persistent observation but are affected by occlusion and blind zones. UAVs provide rapid viewpoint reconfiguration and wide-area search, while UGVs support close-range confirmation and ground-level interaction. Neither type of platform alone can simultaneously satisfy coverage, timeliness, accuracy, endurance, and reliability requirements. Therefore, the core problem is no longer whether a single sensor can detect a target, but how to organize heterogeneous sensing, communication, computation, and decision resources into a dependable cooperative system.

Cooperative air-ground surveillance has been shown to improve persistent observation and target search [1]. Related studies have identified spatial registration, communication constraints, and distributed decision-making as key difficulties in cooperative air/ground robotic systems [2]. Symbiotic UAV-UGV planning has demonstrated complementary sensing benefits in precision agriculture [3], while memetic path planning and vision-aware localization have improved cooperative detection and target localization [4-5]. UAV-based real-time object detection and edge intelligence provide technical support for onboard processing [6-7]. In airport applications, learning-based FOD detection and runway geolocation have been investigated [8-9]. Robust SLAM also provides a foundation for localization and mapping in complex environments [10]. Nevertheless, these studies usually emphasize one algorithm, one platform pair, or one application link rather than a complete system architecture.

Three limitations remain in practical construction. First, sensing resources are often superimposed without a unified mechanism for quality assessment, uncertainty fusion, and confidence propagation. Second, task allocation and path planning are frequently separated from communication quality, energy state, and failure risk. Third, many architectures lack explicit degradation rules, interface specifications,

and engineering verification paths. To address these issues, this paper makes four contributions: (1) a dual-closed-loop and six-layer system architecture is proposed; (2) a quality-aware heterogeneous observation and fusion model is established; (3) a joint task-allocation and route re-planning model constrained by coverage, delay, energy, communication, and safety is designed; and (4) multi-baseline Monte Carlo simulation is used to evaluate detection, confirmation, delay, timeliness, and robustness.

2. System Architecture Design

The system is designed around five engineering requirements: global coverage with local blind-zone compensation, rapid confirmation of high-priority events, edge-cloud collaborative intelligence, fault-tolerant operation under degraded communication, and modular deployment with standardized interfaces. As shown in Figure 1, the architecture consists of the mission application layer, decision and service plane, collaborative intelligence plane, communication and security plane, heterogeneous execution plane, and data/knowledge/model base. Reliability and engineering governance are treated as two cross-cutting pillars.

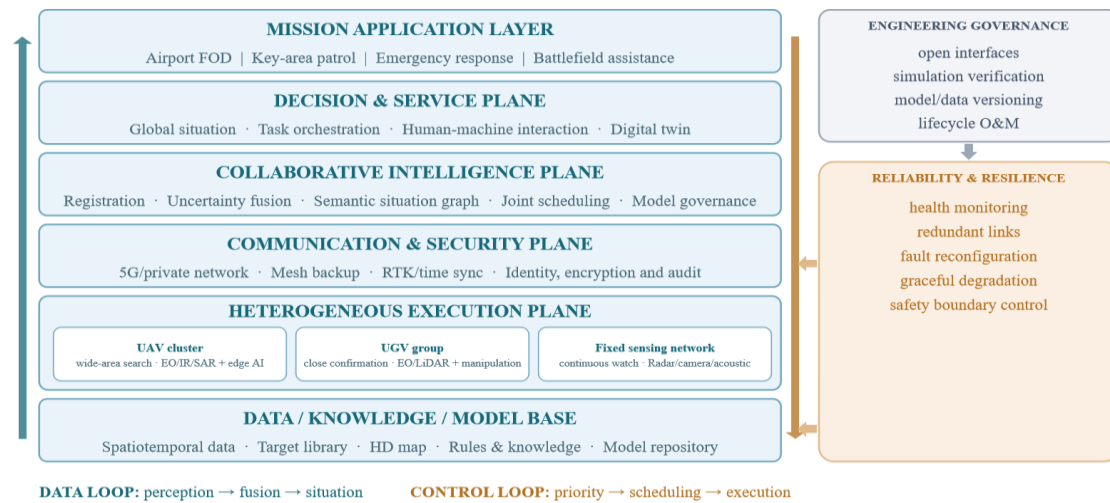


Fig.1. Dual-closed-loop architecture of the air-ground collaborative intelligent perception system

The heterogeneous execution plane contains UAV clusters, UGV groups, and fixed sensing networks. UAVs carry EO/IR/SAR payloads for wide-area search and rapid viewpoint adjustment; UGVs carry EO and LiDAR sensors for close-range confirmation and ground inspection; fixed stations provide continuous observation at runways, aprons, entrances, and perimeter lines. The communication and security plane integrates a 5G/private network with mesh backup, RTK positioning, time synchronization, identity authentication, encryption, and audit. The collaborative intelligence plane performs registration, edge inference, multi-source fusion, situation construction, task scheduling, and model governance. The decision and service plane provides global situation display, human-machine interaction, alert dispatch, and digital-twin-assisted rehearsal. The data/knowledge/model base maintains spatiotemporal data, target libraries, high-definition maps, rules, knowledge, and model versions.

The architecture contains two coupled loops. The data loop follows the chain of observation, registration, fusion, situation construction, and confidence assessment. The control loop follows the chain of task priority, resource scheduling, route re-planning, execution, and health feedback. The two loops interact through the semantic situation graph: when an event has high priority but insufficient confidence, the control loop assigns the most suitable platform to obtain additional observations; when platform health or link quality changes, the data loop updates sensor quality parameters and the scheduler recalculates the allocation. This design distinguishes the proposed system from a simple platform aggregation architecture.

$$z_i(k) = h_i(x(k), s_i(k)) + v_i(k), v_i(k) \sim N(0, R_i(k)) \quad (1)$$

Among them, $z_i(k)$ is the observation generated by node i at time k ; $x(k)$ is the target or event state; $s_i(k)$ is the platform state; $h_i(\cdot)$ is the observation function determined by sensor type, viewpoint, and environmental conditions; and $v_i(k)$ is observation noise with covariance $R_i(k)$. The platform state includes position, attitude, available energy, computing load, and communication quality. This model provides a unified interface for EO, IR, radar, LiDAR, acoustic, and other sensors, allowing different

modalities to enter the same fusion and scheduling framework.

For engineering implementation, each sensing node encapsulates raw data, pose, timestamp, detection result, confidence, health state, and link quality into a standard observation message. The edge node performs lightweight detection, compression, and local validity checking; the regional fusion node performs association, tracking, and confidence propagation; and the command center performs mission-level decision-making. The hierarchical division prevents raw-data flooding and enables the system to continue local operation when the wide-area link is interrupted.

3. Cooperative Perception Model and Scheduling Method

The cooperative perception mechanism is organized as an event-driven closed loop. Heterogeneous sensors first generate observations with quality tags. The fusion center then performs multi-hypothesis association, uncertainty propagation, and situation graph updating. If event confidence is lower than the mission threshold or the event priority is high, the scheduler selects a suitable platform for re-observation. If platform failure or communication degradation occurs, the reliability supervisor triggers redundant-link switching, local autonomy, or task reassignment. The complete process is shown in Figure 2.

$$p_i(k) = p_i^0 \exp\left(-\frac{d_i^2(x)}{2\sigma_i^2}\right) q_i^c q_i^e a_i \quad (2)$$

In Eq. (2), $p_i(x)$ is the effective detection probability of node i at location x ; p_i^0 is the nominal detection probability; $d_i(x)$ is the effective distance between the sensor and target; σ_i describes the sensing attenuation scale; q_i^c denotes communication quality; q_i^e denotes environmental quality; and a_i denotes platform availability. Compared with a purely geometric coverage model, this formulation incorporates sensor performance, environment, communication, and platform health, and is therefore more suitable for engineering evaluation.

$$R_f^{-1} = \sum_{i \in S_k} w_i R_i^{-1}, x_f^\wedge = R_f \sum_{i \in S_k} w_i R_i^{-1} x_i^\wedge \quad (3)$$

In Eq. (3), S_k is the set of valid observations associated with the same target or event; R_i is the observation covariance; w_i is the quality weight determined by timestamp error, viewing angle, occlusion, weather, sensor health, and communication quality; R_f and $x_f^\wedge$ are the fused covariance and state estimate. The fused confidence is not treated as a fixed score. Instead, it is updated with observation consistency, temporal continuity, cross-modal agreement, and historical reliability. Low-confidence observations are retained as candidates rather than discarded, which reduces premature loss of weak targets.

3.1 Unified registration and edge quality assessment

First, GNSS/RTK, inertial navigation, visual SLAM, and fixed-station calibration are combined to establish a common spatial reference. A unified clock aligns image frames, radar plots, LiDAR point clouds, and platform states. Each observation is assigned a quality vector containing timestamp error, pose uncertainty, occlusion level, weather condition, sensor health, and communication quality.

Second, edge nodes execute lightweight object detection, event screening, feature extraction, and adaptive compression. UAVs emphasize runway-surface anomaly screening and viewpoint selection; UGVs emphasize close-range re-identification and fine-grained inspection; fixed stations emphasize continuous background modeling and perimeter intrusion detection. Only key frames, features, tracks, and event messages are preferentially transmitted when bandwidth is constrained.

3.2 Uncertainty-aware multi-source fusion

Third, the fusion center associates candidate targets using spatial distance, motion consistency, semantic category, appearance features, and historical tracks. Multi-hypothesis association is used when observations conflict. The fused result is represented as a target-event-area graph, in which nodes denote targets, events, platforms, and regions, while edges denote observation, tracking, scheduling, and disposal relations.

Fourth, confidence propagation is performed at three levels: observation confidence describes single-sensor quality; track confidence describes temporal consistency; and event confidence describes whether cross-platform evidence satisfies the disposal threshold. This design avoids treating a single detection result as final truth and supports explainable alerting.

Fifth, the reliability supervisor continuously monitors platform health, link quality, edge computing load, and model drift. When a sensor degrades, its weight is reduced; when a platform fails, its sub-area is reassigned; when the cloud is disconnected, edge nodes switch to local autonomy and cache event messages for delayed synchronization.

3.3 Joint task allocation and route re-planning

$$\min_u J = \alpha T + \beta E + \gamma D - \eta C - \lambda I \quad (4)$$

Here, $u=(x,r)$ denotes the joint decision vector consisting of task assignment x and route re-planning r . T denotes confirmation delay; E denotes platform energy consumption; D denotes communication and computational load; C denotes coverage utility; and I denotes information gain. The coefficients alpha, beta, gamma, eta, and lambda are adjusted according to mission type. For FOD disposal and intrusion warning, delay and confidence are prioritized; for routine patrol, energy and coverage balance are emphasized.

The constraints include task uniqueness, time windows, platform energy, no-fly/no-go zones, safety separation, communication quality, payload compatibility, and maximum route deviation. A rolling-horizon hierarchical solver is adopted: a rule engine immediately handles high-priority safety events; a heuristic allocator generates feasible assignments; and route refinement is performed only for the selected platforms. This design avoids solving a large global optimization problem from scratch at every event and is more suitable for online deployment.

$$A_{\text{sys}} = 1 - \prod_{l=1}^L [1 - A_l \prod_{m=1}^{M_l} r_{lm}] \quad (5)$$

The degradation strategy is defined in four levels. Level 1 is normal cloud-edge collaborative operation. Level 2 switches to mesh backup and reduces nonessential data streams. Level 3 enables autonomous regional clusters and local event caching. Level 4 retains only safety-critical detection and alert functions. This explicit degradation chain improves mission continuity and makes the system behavior predictable under faults.

Figure 2 summarizes the cooperative perception and scheduling process. The upper lane performs observation and quality assessment; the middle lane performs uncertainty-aware fusion; and the lower lane performs joint scheduling, route re-planning, and command feedback. The reliability supervisor crosses all three lanes and provides health-state inputs to fusion and scheduling.

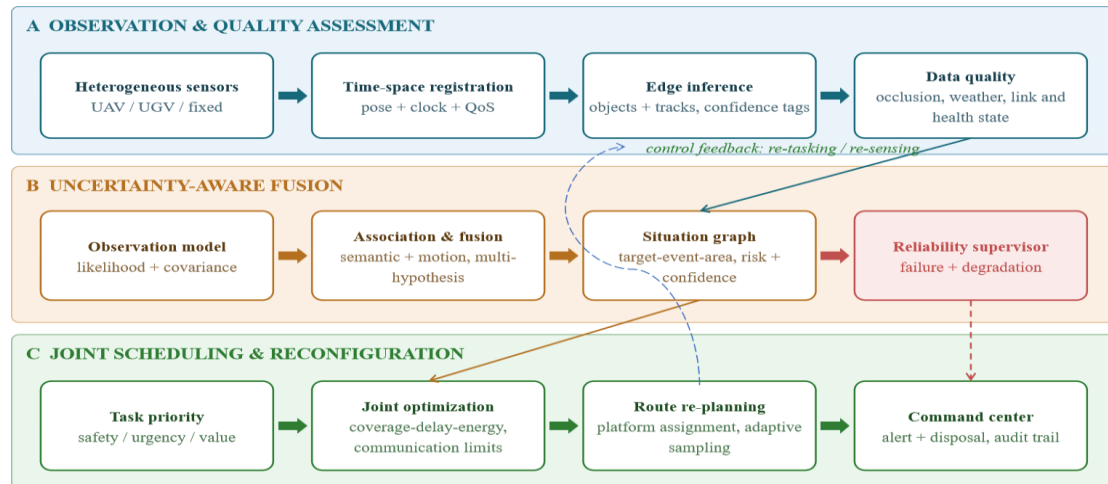


Fig.2. Uncertainty-aware cooperative perception and joint scheduling mechanism

4. Multi-Baseline Simulation and Analysis

A discrete-event simulation is constructed in a 2 km × 2 km area to evaluate the architecture under explicit assumptions. Thirty-six target events are randomly generated in each run. Event occurrence times are uniformly distributed within 300 s, and detection is evaluated until 360 s. The scenario contains 8 fixed stations, 4 UAVs, and 4 UGVs. Platform availability, communication quality, patrol route, sensing radius, and event location are all modeled as variables. The main parameters are listed in Table 1.

Table 1. Main parameters of the multi-baseline simulation

Object	Parameter	Value
System	Mission area / event number	2 km × 2 km / 36 random events
Platforms	Fixed stations / UAVs / UGVs	8 / 4 / 4
Sensing	Fixed / UAV / UGV radius	300 / 300 / 90 m
Motion	UAV / UGV speed	15 / 5 m/s
Time	Event window / evaluation horizon	300 s / 360 s
Availability	Fixed / UAV / UGV	0.98 / 0.95 / 0.97
Comm.	Link quality / retry	0.90-0.99 / 2 s
Evaluation	Monte Carlo runs / timely threshold	120 / 60 s

Four schemes are compared: fixed-station-only perception, UAV-only patrol, heterogeneous patrol without event-driven re-tasking, and the proposed cooperative architecture. The metrics include detection rate, confirmation rate, mean confirmation delay, 95th-percentile delay, and timely confirmation within 60 s. Each scheme is evaluated over 120 Monte Carlo runs. Figure 3 shows one representative spatial realization of the simulation scenario.

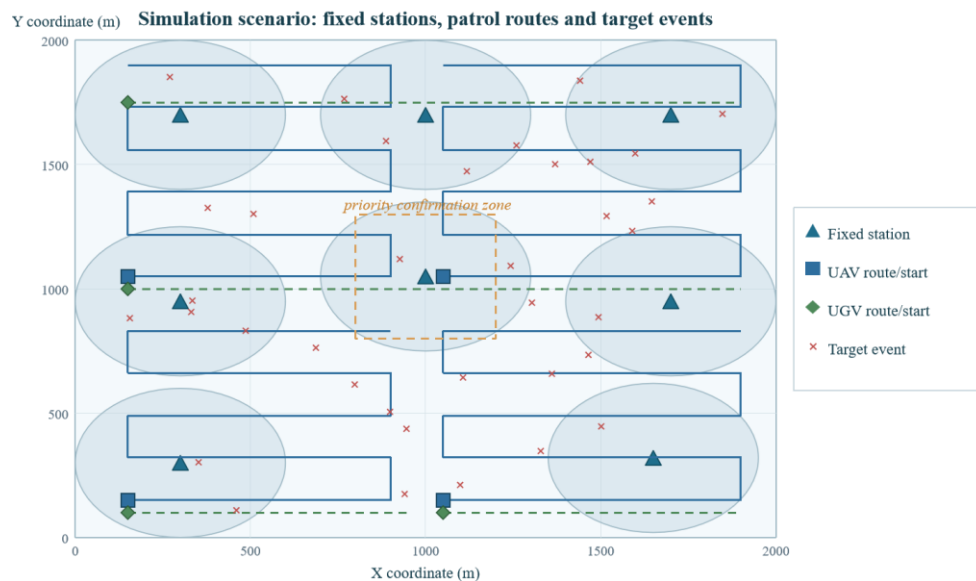


Fig.3. Representative simulation scenario with fixed coverage, patrol routes, and target events

Table 2 summarizes the statistical results. The fixed-only scheme has the lowest delay for the subset it can observe, but its detection and confirmation rates are only 35.7%, because events outside the fixed sensing range cannot be handled. UAV-only patrol improves the confirmation rate to 92.0%, but its mean confirmation delay reaches 45.82 s. Heterogeneous patrol further increases the confirmation rate to 92.5% and reduces the mean delay to 42.02 s. The proposed architecture achieves a 94.4% confirmation rate, a mean delay of 32.46 s, and a P95 delay of 107.04 s.

Table 2. Statistical comparison of the four perception schemes

Method	Detection / Confirmation (%)	Mean / P95 delay (s)	Timely ≤60 s (%)
Fixed-only	35.7 / 35.7	0.48 / 0.48	35.7
UAV-only	92.0 / 92.0	45.82 / 143.32	63.9
Hetero. patrol	94.4 / 92.5	42.02 / 135.56	66.8
Proposed	94.4 / 94.4	32.46 / 107.04	78.7

Compared with heterogeneous patrol without event-driven re-tasking, the proposed scheme reduces the mean confirmation delay by 22.8%, reduces the P95 delay by 21.0%, and increases timely confirmation within 60 s from 66.8% to 78.7%. The paired-run improvement in mean confirmation delay is 9.56 s with a 95% confidence interval of 8.68-10.45 s, and the difference is statistically significant at $p < 0.001$. These results indicate that the main benefit comes from closed-loop scheduling rather than merely adding more sensing platforms.

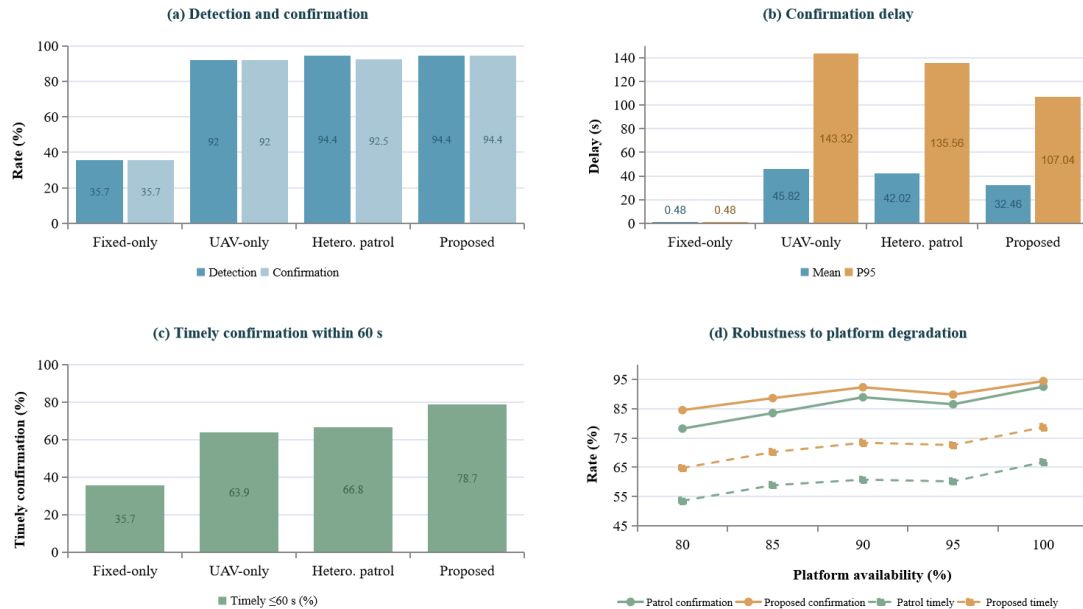


Fig.4. Multi-baseline simulation results and robustness to platform degradation

Figure 4 further shows the robustness analysis. When platform availability decreases from 100% to 80%, the proposed scheme consistently maintains a higher confirmation and timely-confirmation rate than non-adaptive patrol. This advantage comes from three mechanisms: redundant observation resources, quality-aware sensor weighting, and task reassignment after platform degradation. Nevertheless, the results are obtained under simplified geometric and communication assumptions and should be interpreted as architecture-level feasibility evidence rather than field-test performance.

5. Engineering Implementation and Application Deployment

The proposed architecture can be implemented incrementally. In the first stage, fixed stations and a command center are deployed to form a baseline perception capability. In the second stage, UAVs and UGVs are connected through standardized observation messages, and edge detection plus regional fusion are enabled. In the third stage, joint scheduling, digital-twin rehearsal, health monitoring, and degraded operation are introduced. This staged path reduces integration risk and allows each stage to produce usable capability.

For airport FOD detection, fixed stations provide continuous runway and apron observation, UAVs perform scheduled sweep and rapid viewpoint confirmation, and UGVs perform close-range verification and disposal support. For key-area patrol, fixed stations identify abnormal changes, UAVs conduct cross-region search, and UGVs enter occluded or ground-level areas. For battlefield operational assistance, the same architecture can be configured with encrypted links, distributed regional fusion nodes, and stricter local-autonomy rules.

Engineering deployment should follow a unified data model and interface specification. Observation messages should contain target category, location, covariance, timestamp, confidence, platform state, sensor state, and communication quality. Scheduling commands should contain task priority, target area, time window, safety constraints, and rollback rules. Model updates should use version management, offline validation, simulation verification, and controlled online release. These measures make the architecture auditable, maintainable, and reproducible.

The main engineering risks include sensor calibration drift, communication congestion, platform endurance, adverse weather, and model performance degradation. Mitigation measures include periodic

calibration, adaptive bandwidth allocation, redundant relay nodes, energy-aware scheduling, weather-sensitive quality weighting, and continuous model evaluation. The architecture should therefore be treated as a living system in which data, models, platforms, and operational rules are jointly maintained.

6. Conclusion

This paper proposes and evaluates a dual-closed-loop air-ground collaborative intelligent perception system. The architecture integrates UAV clusters, UGV groups, fixed sensing networks, edge intelligence, cloud fusion, communication security, reliability supervision, and engineering governance into a unified framework. Heterogeneous observation, quality-aware fusion, coverage utility, joint scheduling, and availability models are established to clarify how sensing resources are converted into timely and credible situation awareness. Multi-baseline Monte Carlo simulation shows that, under the stated assumptions, the proposed architecture improves timely confirmation and reduces both mean and tail confirmation delay compared with non-adaptive heterogeneous patrol. Future work will focus on real sensor data, hardware-in-the-loop verification, communication-constrained fusion, model drift adaptation, and field validation in airport and other operational scenarios.

References

- [1] Grocholsky B, Keller J, Kumar V, et al. Cooperative air and ground surveillance. *IEEE Robotics & Automation Magazine*, 2006, 13(3): 16-25.
- [2] Lacroix S, Le Besnerais G. Issues in cooperative air/ground robotic systems. *Robotics Research: The 13th International Symposium ISRR, Springer Tracts in Advanced Robotics*, 2010: 421-432.
- [3] Tokekar P, Vander Hook J, Mulla D, et al. Sensor planning for a symbiotic UAV and UGV system for precision agriculture. *IEEE Transactions on Robotics*, 2016, 32(6): 1498-1511.
- [4] Li J, Sun T, Huang X, et al. A memetic path planning algorithm for unmanned air/ground vehicle cooperative detection systems. *IEEE Transactions on Automation Science and Engineering*, 2022, 19(4): 2724-2737.
- [5] Liu D, Bao W, Zhu X, et al. Vision-aware air-ground cooperative target localization for UAV and UGV. *Aerospace Science and Technology*, 2022, 124: 107525.
- [6] Cao Z, Kooistra L, Wang W, et al. Real-time object detection based on UAV remote sensing: A systematic literature review. *Drones*, 2023, 7(10): 620.
- [7] Cheng Q, Wang H, Zhu B, et al. A real-time UAV target detection algorithm based on edge computing. *Drones*, 2023, 7(2): 95.
- [8] Noroozi M, Shah A. Towards optimal foreign object debris detection in an airport environment. *Expert Systems with Applications*, 2023, 213: 118829.
- [9] Niu Z, Zhang J, Li Z, et al. Automatic detection and predictive geolocation of foreign object debris on airport runway. *IEEE Access*, 2024, 12: 133748-133763.
- [10] Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics*, 2016, 32(6): 1309-1332.