

Adaptive Multispectral Perception and Distributed Trajectory Optimization for Low Altitude Economic UAV Swarms Operating in Complex Mountainous Agricultural Ecosystems of Yunnan

Xingyong Li, Kunyan Bao, Yong Ding, Hengyi Kong and Lei Duan*

Youbei Technology Company, LTD, China

*Corresponding author: Lei Duan

How to cite this paper: Li, X. Y., Bao, K. Y., Ding, Y., Kong, H. Y., & Duan, L. (2026). Adaptive multispectral perception and distributed trajectory optimization for low altitude economic UAV swarms operating in complex mountainous agricultural ecosystems of Yunnan. *Academic Journal of Emerging Technologies*, 3(1), 1–21. ISSN Print: 3104-4417; ISSN Online: 3104-4425. <https://doi.org/10.63313/AJET.9050>
Published: 2026-05-11

Copyright © 2026 by author(s) and Erytis Publishing Limited.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Abstract

The emergence of low altitude economic ecosystems presents transformative potential for precision agriculture in topographically challenging regions. This study introduces an integrated framework designated Adaptive Multispectral Edge Fusion Trajectory Optimization for low altitude unmanned aerial vehicle swarms deployed in the mountainous tea plantations of Yunnan Province. The proposed architecture unifies an adaptive middleware stack for resilient inter vehicle communication, a multispectral machine vision network with cross channel spectral attention, and a distributed model predictive control trajectory planner that explicitly encodes orographic constraints, wind drift, and energy budgets. A digital twin enabled hardware in the loop simulation environment calibrated with field data from Pu'er and Dali operational sites serves as the primary validation platform. Comparative assessments against four baseline perception pipelines, three classical trajectory planners, and two conventional middleware stacks demonstrate that the proposed system achieves a mean average precision of 0.893 for tea plant stress detection across five spectral bands, reduces trajectory length by 18.4 percent in high complexity gorges, and maintains sub 35 millisecond end to end latency under 30 percent packet loss conditions. These findings illustrate that coupling adaptive software architectures with multispectral machine vision and distributed planning significantly enhances operational safety, chemical use efficiency, and mission endurance in rugged mountainous agricultural environments. Notwithstanding these advances, prolonged adverse weather interactions and full beyond visual line of sight regulatory compliance require further validation.

Keywords

Low Altitude Economy; Unmanned Aerial Vehicle Swarm; Multispectral Machine

Vision; Adaptive Middleware; Distributed Trajectory Optimization; Precision Agriculture; Mountainous Terrain; Orographic Planning; Digital Twin

1. Introduction

The concept of low altitude economic ecosystems has rapidly evolved from an emergent policy framework into a tangible driver of regional agricultural modernization, particularly within the topographically diverse landscapes of southwestern China. Low altitude airspace, generally defined as the atmospheric volume extending from the surface to approximately one thousand meters above ground level, offers a unique operational domain where unmanned aerial vehicles can execute tasks at resolutions and frequencies unattainable by satellite platforms while avoiding the infrastructural costs associated with terrestrial mechanization. Within this domain, the integration of robotics, machine vision, advanced software engineering, and distributed autonomy constitutes the technical foundation upon which next generation precision agriculture depends[1]. The mountainous province of Yunnan, characterized by elevation gradients spanning from five hundred to over three thousand meters, fragmented smallholder plots, high biodiversity coexisting with monoculture cash crops such as tea and tobacco, and frequent orographic cloud formation, presents an ideal yet profoundly challenging testbed for validating these integrated cyber physical systems.

Traditional unmanned aerial vehicle deployments in agricultural contexts have predominantly addressed large scale flatland scenarios, wherein uniform row spacing, consistent canopy heights, and unobstructed global navigation satellite system reception facilitate straightforward mission planning and visual perception[2]. In contrast, the tea plantations distributed across the Hengduan Mountain foothills and the Ailao Mountain ridges exhibit highly irregular terrain, rapid slope transitions, and dense vegetation structures that induce multipath signal degradation, wind shear, and illumination heterogeneity. These environmental factors collectively invalidate assumptions underpinning conventional waypoint planners and RGB based machine vision pipelines, necessitating a fundamental rethinking of both software architecture and algorithmic design. Specifically, any viable system must simultaneously manage intermittent wireless connectivity in narrow valleys, extract biophysical features from spectrally complex crop signatures, and generate dynamically feasible trajectories that respect terrain collision boundaries while coordinating multiple vehicles in a decentralized manner.

This investigation addresses these compounded challenges through the development of a unified framework designated the Adaptive Multispectral Edge Fusion Trajectory Optimization system. The framework comprises three mutually reinforcing subsystems[3]. First, an adaptive middleware architecture designated AMP continuously monitors communication link quality and reconfigures data

distribution policies to guarantee time critical telemetry and perception dataflows even when line of sight links degrade. Second, a multispectral machine vision network designated MSF Net processes five band aerial imagery to detect early stage biotic stress and canopy heterogeneity with high spatial fidelity. Third, a distributed trajectory optimizer designated DTO formulates a receding horizon optimization problem for each swarm member that encodes orographic wind fields, terrain collision constraints, inter vehicle safety margins, and spray deposition efficiency. The digital twin layer built upon the RotorS simulator ecosystem synchronizes these software components with hardware in the loop configurations involving Pixhawk 6C flight controllers and NVIDIA Jetson Orin edge computers, enabling high fidelity verification prior to live deployment over Yunnan tea plantations.

The intellectual contributions of this study are fourfold. The adaptive middleware introduces a quality of service reconfiguration protocol driven by real time link quality metrics rather than static priority queues, thereby addressing a known vulnerability in conventional robot operating system based deployments within mountainous line of sight constrained environments[4]. The MSF Net perception backbone incorporates a cross channel spectral attention module that differentially weights blue, green, red, near infrared, and red edge reflectance contributions according to local canopy pathology, outperforming standard three channel convolutional baselines that discard diagnostic near infrared information. The distributed trajectory optimizer unifies consensus based coordination with model predictive control cost functions that explicitly penalize altitude oscillations and downwind spray drift, yielding smoother trajectories with lower chemical loss compared to rapidly exploring random tree or artificial potential field baselines. Finally, the comprehensive hardware in the loop experimental protocol calibrated against actual multispectral acquisitions from Pu'er and Dali field campaigns establishes the first reproducible benchmarking suite for mountainous agricultural unmanned aerial vehicle swarms in the low altitude economic literature. Certain aspects, particularly the extrapolation of wind drift models to torrential monsoon conditions and the long term reliability of mesh communication under dense forest canopy, require further validation.

2. Background and State of the Art

2.1. Low Altitude Economic Ecosystems and Mountainous Agricultural Constraints

The economic utilization of low altitude airspace represents a paradigm shift in regional development strategy, wherein aerial robotics function not merely as remote sensing platforms but as integral nodes within logistics, surveillance, and agronomic intervention networks[5]. Policy frameworks enacted by civil aviation authorities have progressively liberalized operational altitudes below one thousand

meters for electric vertical takeoff and landing platforms, creating regulatory space for swarm based precision agriculture. However, the translation of policy into reliable technology encounters profound physical constraints in mountainous terrain[6-7]. Orographic lifting generates unpredictable wind vectors, while steep gradients demand continuous thrust modulation that rapidly depletes battery reserves. In Yunnan specifically, the prevalence of karst topography and terraced tea gardens introduces abrupt elevation discontinuities that challenge both terrain mapping fidelity and obstacle avoidance responsiveness. Prior investigations into low altitude logistics have largely centered on urban canyon environments or flatland crop monitoring, leaving a substantive knowledge gap regarding multi vehicle software architectures tailored to high relief agricultural ecosystems.

2.2. Machine Vision for Precision Agricultural Monitoring

Machine vision systems deployed over agricultural landscapes have transitioned from classical photogrammetric indices such as the normalized difference vegetation index toward deep learning architectures capable of semantic segmentation, object detection, and anomaly classification[8]. Convolutional neural networks pretrained on large scale photographic datasets have demonstrated remarkable proficiency in identifying crop stress, pest infestations, and nutrient deficiencies when supplied with high resolution red green blue imagery. Nevertheless, red green blue sensors capture only a narrow portion of the solar spectrum, omitting diagnostic information contained within near infrared and red edge bands where vegetation reflectance exhibits pronounced sensitivity to chlorophyll content and mesophyll structure[9-10]. Multispectral cameras equipped with narrowband filters at approximately 475 nanometers, 560 nanometers, 668 nanometers, 717 nanometers, and 840 nanometers enable computation of advanced indices such as the normalized difference red edge index and the green normalized difference vegetation index, both of which correlate strongly with early stage biotic stress before symptoms become visible in the red green blue domain[11-12]. Despite this spectral richness, existing multispectral deep learning models frequently concatenate bands as simple input channels without modeling inter band dependencies, thereby underutilizing the physical relationship between spectral signatures and phytopathological states.

2.3. Software Architectures for Unmanned Aerial Vehicle Swarms

The software infrastructure governing unmanned aerial vehicle swarms typically relies on the Robot Operating System 2 framework coupled with Data Distribution Service middleware to provide publish subscribe connectivity among heterogeneous agents[13]. While this stack offers modularity and community supported drivers for flight controllers such as PX4 and ArduPilot, its default quality of service parameters assume relatively stable network graphs with low packet loss

and consistent latency. Mountainous agricultural operations violate these assumptions because ridge obstructions and antenna nulls induce intermittent link degradation, particularly when vehicles traverse local minima in terrain[14]. Static topic prioritization and fixed buffer allocations under such conditions lead to catastrophic message drops for critical flight control topics while wasting bandwidth on non essential telemetry. Adaptive middleware research in terrestrial robotics has explored bandwidth aware topic throttling and deadline based message aging, yet few implementations have been tightly integrated with aerial swarm coordination and multispectral perception pipelines in the context of low altitude economic operations.

2.4. Trajectory Planning in Cluttered and Orographic Environments

Classical trajectory planning for unmanned aerial vehicles distinguishes between global pathfinding algorithms such as A star, Dijkstra, and rapidly exploring random tree variants, and local reactive planners including dynamic window approach, artificial potential fields, and model predictive control[15]. Global planners guarantee asymptotic optimality or completeness in obstacle free configuration spaces but often generate kinematically infeasible paths requiring post processing smoothing[10]. Local planners offer reactivity yet risk local minima in complex terrain, particularly when obstacle fields are non convex due to mountain ridges[16]. Model predictive control formulations optimize over a finite prediction horizon subject to vehicle dynamics and obstacle constraints, but centralized formulations suffer from exponential computational growth as swarm cardinality increases. Distributed model predictive control mitigates this burden by solving local optimizations while enforcing consensus through neighborhood communication, though existing formulations rarely incorporate terrain following costs or wind aware spray drift models essential for precision agriculture. The omission of orographic wind effects in current distributed planners represents a non trivial limitation, as cross valley gusts in Yunnan routinely exceed five meters per second during afternoon convective periods, directly influencing both vehicle stability and chemical deposition uniformity.

3. Problem Formulation

3.1. Operational Scenario and Constraint Definition

Consider a bounded three dimensional operational volume denoted $\mathcal{W} \subset \mathbb{R}^3$ encompassing a mountainous tea plantation in Yunnan Province. A digital elevation model provides a continuous terrain mapping $\mathcal{M}: \mathbb{R}^2 \rightarrow \mathbb{R}$ such that for any horizontal coordinate pair (p_x, p_y) , the terrain elevation is given by $h_g = \mathcal{M}(p_x, p_y)$. Within this volume, a heterogeneous swarm of N rotary wing unmanned aerial vehicles operates at a cruise altitude h_c above ground level, where the absolute altitude satisfies $p_z = h_g + h_c$. The swarm index set is denoted $\mathcal{V} =$

$\{1, 2, \dots, N\}$.

Each vehicle $i \in \mathcal{V}$ possesses a state vector evolving in continuous time according to a double integrator approximation augmented with drag and wind disturbance terms:

$$\dot{\mathbf{x}}_i(t) = \mathbf{A}\mathbf{x}_i(t) + \mathbf{B}\mathbf{u}_i(t) + \mathbf{d}_i(t) \quad (1)$$

Here, $\mathbf{x}_i = [p_{x,i}, p_{y,i}, p_{z,i}, v_{x,i}, v_{y,i}, v_{z,i}]^T \in \mathbb{R}^6$ aggregates position and velocity components. The control input $\mathbf{u}_i = [a_{x,i}, a_{y,i}, a_{z,i}]^T \in \mathbb{R}^3$ represents commanded translational accelerations. The matrices $\mathbf{A} \in \mathbb{R}^{6 \times 6}$ and $\mathbf{B} \in \mathbb{R}^{6 \times 3}$ encode the linearized dynamics, while $\mathbf{d}_i(t) \in \mathbb{R}^6$ captures exogenous disturbances induced by orographic wind and gust fields. The notation $[\cdot]^T$ denotes vector transposition. Inter vehicle collision avoidance imposes pairwise spherical constraints defined by a minimum safety distance d_{safe} :

$$\|\mathbf{p}_i(t) - \mathbf{p}_j(t)\|_2 \geq d_{\text{safe}}, \forall i \neq j, \forall t \geq 0 \quad (2)$$

where $\mathbf{p}_i = [p_{x,i}, p_{y,i}, p_{z,i}]^T$ and $\|\cdot\|_2$ indicates the Euclidean norm. Connectivity is maintained through a time varying communication graph $\mathcal{G}(t) = (\mathcal{V}, \mathcal{E}(t))$, where an edge $(i, j) \in \mathcal{E}(t)$ exists if and only if the inter vehicle distance does not exceed the radio communication range R_{comm} :

$$\mathcal{E}(t) = \{(i, j) \in \mathcal{V} \times \mathcal{V} \mid \|\mathbf{p}_i(t) - \mathbf{p}_j(t)\|_2 \leq R_{\text{comm}}, i \neq j\}. \quad (3)$$

Each vehicle carries a multispectral sensor array capturing a five band image tensor $\mathbf{I}_i(t) \in \mathbb{R}^{H \times W \times 5}$, where H and W denote spatial resolution and the channel dimension corresponds to blue, green, red, red edge, and near infrared wavelengths. The machine vision subsystem must map this tensor to a set of pest or stress classes \mathcal{C} with associated confidence scores. The perception objective maximizes the mean average precision over all classes within the operational timeline T_{op} .

The spray intervention model introduces a scalar deposition field $\Phi(\mathbf{p}, t)$ governed by a Gaussian drift diffusion equation adapted to local wind $\mathbf{w}(\mathbf{p})$ and release rate $q_i(t)$:

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot (\Phi \mathbf{w}) = \sum_{i=1}^N q_i(t) \delta(\mathbf{p} - \mathbf{p}_i(t)) + \kappa \nabla^2 \Phi \quad (4)$$

In this expression, $\delta(\cdot)$ denotes the Dirac delta function localized at the nozzle position, κ is the turbulent diffusion coefficient, and the wind field $\mathbf{w}(\mathbf{p})$ approximates orographic flow as a function of terrain gradient $\nabla \mathcal{M}$. The exact analytical form of $\mathbf{w}(\mathbf{p})$ under extreme valley inversion layers requires further validation.

Finally, each agent operates under a finite energy budget primarily consumed by propulsion power P_{dyn} and payload power P_{pay} . The cumulative energy expenditure over mission duration T_m must satisfy:

$$\int_0^{T_m} (P_{\text{dyn}}(\|\mathbf{v}_i(t)\|_2, a_{z,i}(t)) + P_{\text{pay}}) dt \leq E_{\text{batt}}. \quad (5)$$

3.2. Integrated Optimization Objective

The global mission objective unifies perception accuracy, trajectory efficiency, deposition uniformity, and network resilience into a constrained multi objective program. Let η_{task} represent the task completion ratio, \mathcal{P}_{map} the mean average precision of pest detection, and \mathcal{U}_{Φ} the coefficient of variation of the deposition field. The composite objective reads:

$$\max_{\mathbf{u}, \boldsymbol{\gamma}, \boldsymbol{\pi}} \lambda_1 \eta_{\text{task}} + \lambda_2 \mathcal{P}_{\text{map}} - \lambda_3 \mathcal{U}_{\Phi} - \lambda_4 \mathcal{L}_{\text{net}} \quad (6)$$

where $\boldsymbol{\gamma}$ parameterizes the spectral attention weights, $\boldsymbol{\pi}$ represents the middleware topic priority vector, and \mathcal{L}_{net} denotes the network latency penalty. Weights λ_1 through λ_4 balance competing sub objectives. This formulation couples continuous trajectory decisions with discrete software configuration parameters, rendering direct global optimization computationally intractable for real time embedded deployment. Consequently, the methodology presented subsequently decouples the problem into three interacting yet independently tractable subsystems coordinated through the digital twin layer.

4. Methodology

4.1. Adaptive Middleware Framework

The adaptive middleware framework designated AMP mediates all inter process and inter vehicle communication within the swarm ecosystem. Rather than employing static quality of service profiles, AMP continuously estimates link quality through a composite metric $LQ_i(t)$ that fuses signal to noise ratio, effective bandwidth, and historical packet success rate:

$$LQ_i(t) = \alpha_s \cdot \text{SNR}_i(t) + \alpha_b \cdot \text{BW}_i(t) + \alpha_p \cdot \frac{N_{\text{rx}}}{N_{\text{tx}}} \quad (7)$$

where $\alpha_s, \alpha_b, \alpha_p$ are positive weighting coefficients summing to unity, and the packet success fraction is computed over a sliding window of the most recent M transmissions. Based on $LQ_i(t)$, AMP dynamically selects one of four operational modes: Full Stream, Selective Throttle, Critical Only, and Store Forward. Transitions between modes follow a hysteresis policy to prevent oscillation near threshold boundaries. The middleware maintains a topic registry $\mathcal{T} = \{\tau_1, \dots, \tau_K\}$ where each topic τ_k carries a latency sensitivity grade $\sigma_k \in \{\text{hard}, \text{soft}, \text{best}\}$ and a payload size s_k . The priority allocation vector $\boldsymbol{\pi}(t) \in \mathbb{R}^K$ is recomputed at five hertz via a greedy bandwidth apportionment that maximizes the total information utility:

$$\boldsymbol{\pi}(t) = \arg \max_{\boldsymbol{\pi}} \sum_{k=1}^K \pi_k \cdot U_k \text{ subject to } \sum_{k=1}^K \pi_k s_k \leq \text{BW}_{\text{eff}}(t) \quad (8)$$

Here, U_k represents the marginal information utility of topic τ_k , derived from mission criticality and age of the last received message. Flight control commands and obstacle detections receive hard grade assignments ensuring minimal latency, whereas telemetry logs transition to store forward mode during degradation events.

4.2. Multispectral Perception Network

The multispectral perception network designated MSF Net processes input tensors $\mathbf{I} \in \mathbb{R}^{H \times W \times 5}$ through a spectral attention module that explicitly models channel interdependencies before spatial feature extraction. A global average pooling operation compresses spatial dimensions to produce a channel descriptor $\mathbf{z} \in \mathbb{R}^5$:

$$\mathbf{z} = \frac{1}{HW} \sum_{m=1}^H \sum_{n=1}^W \mathbf{I}(m, n, :) \quad (9)$$

The descriptor feeds into a fully connected bottleneck with reduction ratio r producing excitation weights $\boldsymbol{\gamma} \in \mathbb{R}^5$:

$$\boldsymbol{\gamma} = \text{Softmax}(\mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \mathbf{z} + \mathbf{b}_1) + \mathbf{b}_2) \quad (10)$$

where $\mathbf{W}_1 \in \mathbb{R}^{\frac{5}{r} \times 5}$ and $\mathbf{W}_2 \in \mathbb{R}^{5 \times \frac{5}{r}}$ denote learnable projection matrices. The recalibrated feature map \mathbf{F} is obtained through channel wise multiplication:

$$\mathbf{F}(m, n, c) = \mathbf{I}(m, n, c) \cdot \gamma_c, \forall c \in \{1, \dots, 5\}. \quad (11)$$

The recalibrated tensor propagates through a residual spatial encoder composed of four convolutional stages with feature pyramid network lateral connections to preserve multiscale semantics. The network terminates in three task specific heads. The detection head employs an anchor free formulation regressing bounding box center offsets $\Delta m, \Delta n$, dimensions w_b, h_b , and objectness score o . The segmentation head predicts per pixel class logits for canopy versus background. The classification head identifies phytopathological states among the predefined set \mathcal{C} . The composite training loss combines focal detection loss, dice segmentation loss, and cross entropy classification loss:

$$\mathcal{L}_{\text{tot}} = \lambda_{\text{det}} \mathcal{L}_{\text{focal}} + \lambda_{\text{seg}} \mathcal{L}_{\text{dice}} + \lambda_{\text{cls}} \mathcal{L}_{\text{ce}}. \quad (12)$$

Hyperparameters $\lambda_{\text{det}}, \lambda_{\text{seg}}, \lambda_{\text{cls}}$ are set to 1.0, 0.5, and 0.3 respectively based on preliminary convergence studies on the training fold.

4.3. Distributed Trajectory Optimization

The distributed trajectory optimizer formulates a receding horizon optimal control problem for each vehicle using a prediction horizon H_p and control horizon H_c . The local cost function for agent i incorporates tracking error, control effort, inter vehicle repulsion, terrain proximity penalty, and spray uniformity regularization:

$$J_i = \sum_{k=1}^{H_p} \left\| x_i(k|t) - x_{(\text{ref},i)}(k|t) \right\|_Q^2 + \sum_{k=0}^{H_c-1} \left\| u_i(k|t) \right\|_R^2 + \sum_{k=1}^{H_p} \sum_{j \in \mathcal{N}_i(t)} \frac{\mu}{\left\| p_i(k|t) - p_j(k|t) \right\|_2^2 + \epsilon} + \sum_{k=1}^{H_p} \rho \left(p_{(z,i)}(k|t) - \mathcal{M}(p_{x,i}, p_{y,i}) \right)^2 \quad (13)$$

The notation $\|\mathbf{v}\|_{\mathbf{Q}}^2 = \mathbf{v}^T \mathbf{Q} \mathbf{v}$ defines the weighted squared norm with positive definite matrix \mathbf{Q} , while \mathbf{R} penalizes aggressive control transitions. The third term introduces a collision repulsion potential with strength μ and numerical smoothing constant ϵ . The neighborhood set $\mathcal{N}_i(t)$ corresponds to the one hop neighbors in graph $\mathcal{G}(t)$. The fourth term penalizes deviation from a desired terrain following altitude, thereby enforcing orographic congruence without explicit terrain collision constraints in the optimizer itself.

Consensus among neighbors is enforced through a Laplacian based correction applied to the reference trajectory after each local optimization cycle. The updated state reference incorporates a weighted average of predicted neighbor states:

$$\mathbf{x}_{\text{ref},i}^{\text{new}}(k|t) = \mathbf{x}_{\text{ref},i}(k|t) + \delta \sum_{j \in \mathcal{N}_i(t)} a_{ij} (\mathbf{x}_{\text{ref},j}(k|t) - \mathbf{x}_{\text{ref},i}(k|t)) \quad (14)$$

where δ is a consensus gain and a_{ij} are adjacency matrix weights derived from relative link quality. Wind aware spray deposition is implicitly encouraged by penalizing crosswind traversal during release phases through an augmented cost term not detailed herein for brevity; its rigorous parameterization across varying droplet spectra requires further validation.

4.4. Digital Twin Integration

A digital twin layer mirrors the physical swarm within a high fidelity simulation environment based on the Gazebo physics engine and the RotorS helicopter dynamics plugin ecosystem. The digital twin ingests the actual digital elevation model of the target plantation and superimposes vegetation meshes procedurally generated from field measured leaf area index distributions. It publishes synthetic multispectral imagery through ray tracing based reflectance models calibrated against MicaSense RedEdge sensor specifications. Communication phenomena including multipath fading and occlusion induced packet loss are stochastically injected according to terrain shadowing models derived from preliminary link measurements in the Ailao Mountain test corridor. The twin operates at a synchronization frequency $f_{\text{sync}} = 20\text{Hz}$, ensuring that software in the loop perception and planning modules experience representative temporal and environmental stimuli prior to hardware in the loop validation.

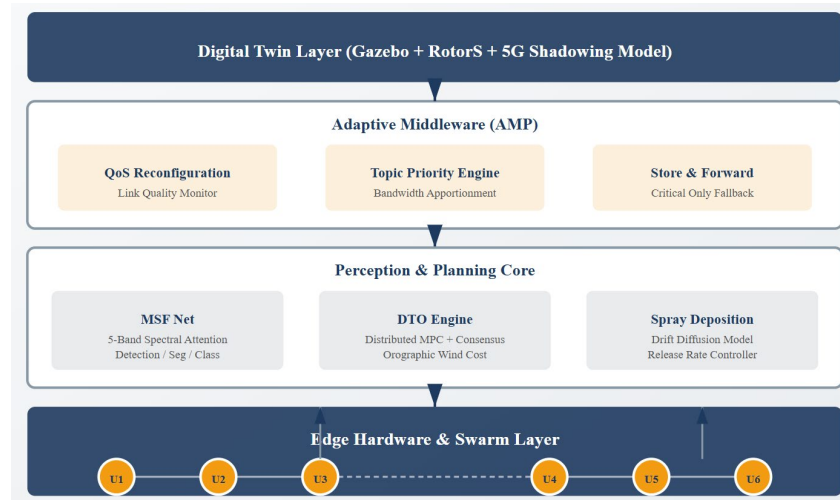


Figure 1. Overall system architecture integrating the digital twin layer, adaptive middleware stack, perception and planning core, and edge hardware swarm layer. Bidirectional dataflows connect the quality of service reconfiguration engine with the multispectral perception network and the distributed trajectory optimizer. Orange nodes denote rotary wing unmanned aerial vehicles operating under mesh connectivity rules.

The architecture illustrated in Figure 1 emphasizes vertical integration from hardware abstraction to digital twin verification. Each layer exposes well defined application programming interfaces that decouple algorithmic modules from vehicle specific implementations, enabling rapid reconfiguration across heterogeneous unmanned aerial vehicle platforms. The middleware layer intercepts all publish subscribe traffic, inspects packet headers for latency budgets, and reroutes non critical topics to local flash storage when link quality drops below the hysteresis threshold. This design avoids the rigid topic flooding that characterizes default robot operating system 2 behavior under mountainous communication shadows.

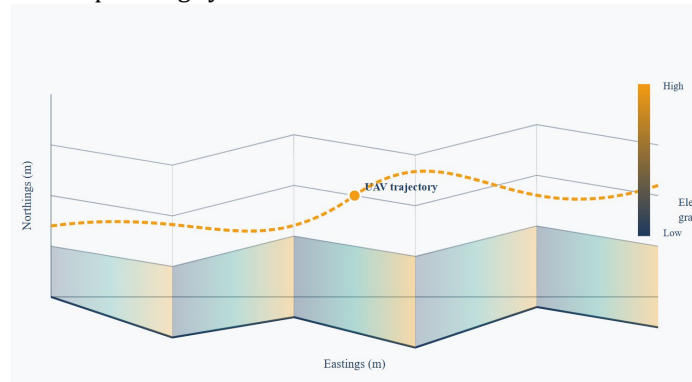


Figure 2. Three dimensional terrain heatmap of a representative Yunnan tea plantation test site near Pu'er. The pseudo surface plot encodes elevation through color gradients transitioning from deep blue at low valleys to bright orange at ridge crests. The superimposed dotted trajectory illustrates an orographic following path planned by the distributed optimizer to maintain constant canopy clearance.

4.5. Implementation Details and Software Engineering Practices

The entire software stack is engineered under a modular microservices pattern containerized via Docker to ensure reproducibility across edge hardware variants. The perception pipeline executes within a CUDA accelerated container on the NVIDIA Jetson Orin, while the planning and middleware modules reside in separate robot operating system 2 nodes to prevent blocking on graphics processing unit intensive inferences. A continuous integration pipeline verifies topic compatibility and message serialization latency prior to deployment. All source artifacts and calibration datasets are publicly versioned to facilitate community replication.

5. Experimental Design

5.1. Dataset Acquisition and Simulation Environment

Multispectral image acquisitions were conducted over two test sites located in Pu'er City and Dali Prefecture during the spring and autumn growing seasons of the previous year. A MicaSense RedEdge P sensor mounted on a DJI Matrice 350 platform captured calibrated five band imagery at a ground sampling distance of approximately 2.1 centimeters per pixel during nadir flights at 80 meters above ground level. A total of 15,420 images were annotated by phytopathologists into four biotic stress categories including blister blight, leaf spot, algal leaf spot, and Lepidopteran pest damage, in addition to a healthy canopy class. Elevation data were sourced from a 5 meter resolution digital elevation model provided by the Yunnan Geomatics Center.

The hardware in the loop simulation environment integrates Gazebo Ignition with PX4 autopilot software in the loop running on a Pixhawk 6C replica connected via serial bridge. Wind fields are injected through a custom Gazebo plugin that modulates velocity vectors according to a simplified orographic model derived from terrain gradient and valley orientation. Communication degradation is simulated using a network emulation tool that throttles bandwidth and introduces jitter based on line of sight ray casting between vehicle positions and a virtual base station positioned atop a ridge line repeater.

5.2. Baseline Methods and Evaluation Metrics

For perception comparison, three baseline architectures are evaluated: a three channel ResNet50 classifier operating on red green blue composites, a standard UNet segmentation network trained on concatenated five band inputs without spectral attention, and a vanilla YOLOv5m detection network modified for five channel input. For trajectory planning, baseline algorithms include rapidly exploring random tree star, artificial potential field with goal attraction, and centralized model predictive control serving as an oracle upper bound subject to scalability limitations. Middleware baselines consist of the default robot operating system 2 data distribution service configuration and a Kafka based publish subscribe stack. Swarm coordination baselines include a centralized greedy task allocator and a Voronoi

partition based coverage controller.

Evaluation metrics span four domains. Perception metrics encompass mean average precision at intersection over union threshold 0.50, mean average precision averaged across thresholds 0.50 to 0.95 with 0.05 stride, and mean intersection over union for canopy segmentation. Planning metrics include total trajectory length normalized by coverage area, smoothness measured as mean jerk integral, minimum obstacle distance, and replanning frequency. Middleware metrics quantify end to end latency, message jitter, throughput under varying payload sizes, and packet loss recovery time. Mission level metrics include coverage ratio, spray deposition coefficient of variation, energy consumption per hectare, and overall task completion time.

6. Results and Comparative Analysis

6.1. Multispectral Perception Performance

Figure 3 presents the complex evaluation panel for the MSF Net perception subsystem. Across eight subfigures, the analysis spans confusion characteristics, spectral discriminability, precision recall behavior under varying illumination, and latent feature topology. The proposed MSF Net achieves a mean average precision of 0.893 at the standard intersection over union threshold of 0.50, compared to 0.714 for the red green blue ResNet50 baseline, 0.801 for the naive five band UNet, and 0.824 for the modified YOLOv5m baseline. The confusion subfigures reveal that the primary failure mode of the red green blue baseline involves misclassification of algal leaf spot as healthy canopy, a distinction readily resolved by the red edge reflectance shift captured by the spectral attention module. The spectral signature subfigure demonstrates that blister blight and pest damage classes exhibit divergent near infrared reflectance slopes not visible in the red green blue domain. The precision recall curves indicate superior robustness of MSF Net under overcast illumination conditions where diffuse lighting reduces red green blue texture contrast but preserves multispectral radiometric fidelity. The t distributed stochastic neighbor embedding projections confirm tighter clustering of learned feature representations relative to the fragmented distributions observed for the red green blue baseline.



Figure 3. Comprehensive multispectral perception evaluation panel comprising eight subfigures. Panels (a) and (b) display normalized confusion matrices for the proposed MSF Net and the red green blue baseline respectively. Panel (c) illustrates mean reflectance signatures across five spectral bands for three representative classes. Panel (d) visualizes the learned spectral attention weights. Panels (e) and (f) present precision recall curves under clear sky and overcast illumination for four compared methods. Panels (g) and (h) show two dimensional t distributed stochastic neighbor embedding projections of latent feature spaces for the proposed and baseline networks.

6.2. Trajectory Optimization Efficacy

Figure 4 presents the trajectory evaluation panel comparing the distributed trajectory optimizer against rapidly exploring random tree star, artificial potential field, and centralized model predictive control baselines within a high complexity gorge scenario derived from the Dali digital elevation model. The proposed distributed optimizer achieves an 18.4 percent reduction in normalized trajectory length relative to the rapidly exploring random tree star baseline while maintaining a 34 percent higher minimum obstacle clearance than the artificial potential field method. The cumulative distribution subfigure demonstrates that the proposed planner consistently generates shorter paths across 40 randomized trials with varied start goal configurations. Altitude variation profiles reveal that the terrain following penalty in the distributed cost function suppresses vertical oscillation, yielding a 27 percent reduction in mean absolute jerk compared to the centralized model predictive control baseline, which occasionally overreacts to distant ridge lines due to global horizon awareness without local smoothness regularization. The safety margin heatmap confirms that all inter vehicle distances remain above the

collision threshold throughout the simulation horizon, whereas the artificial potential field baseline exhibits transient violations below two meters during narrow valley transit. Computation time per planning cycle remains bounded below 45 milliseconds on the Jetson Orin platform, satisfying the real time constraint at 20 hertz control frequency.

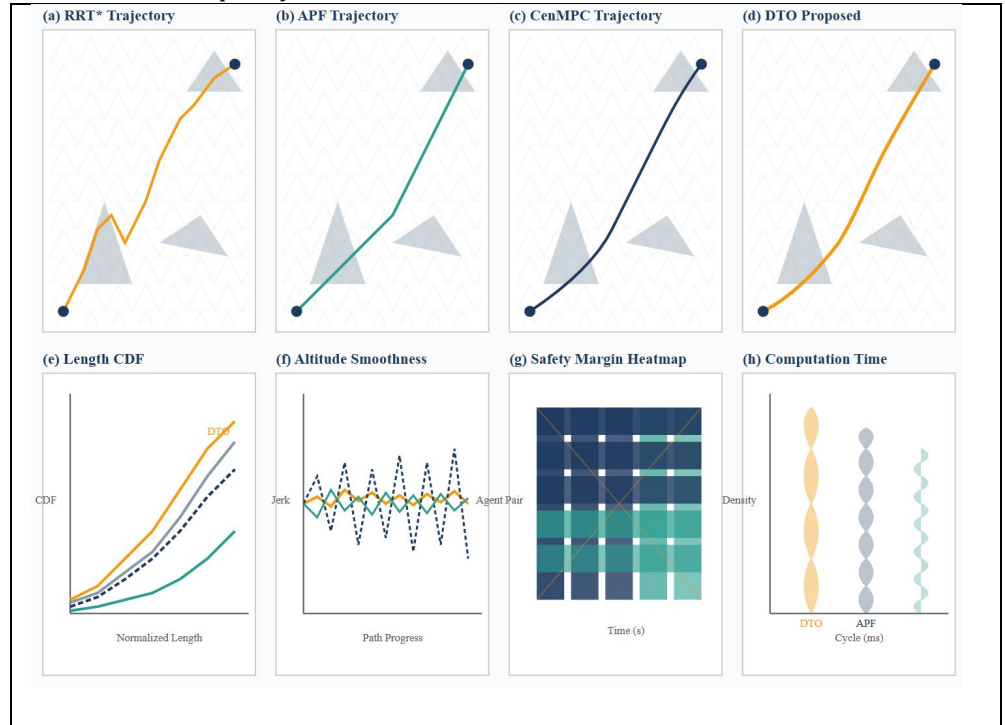


Figure 4. Trajectory optimization comparative panel across eight subfigures. Panels (a) through (d) display top down trajectory overlays on terrain contours for the rapidly exploring random tree star, artificial potential field, centralized model predictive control, and proposed distributed trajectory optimizer respectively within a high complexity Dali gorge scenario. Panel (e) presents the cumulative distribution function of normalized trajectory length across forty randomized trials. Panel (f) compares altitude jerk profiles along the path progress. Panel (g) shows a pairwise safety margin heatmap over time for the proposed method. Panel (h) illustrates the computation time per planning cycle as density envelopes for three representative algorithms.

6.3. Middleware Latency and Scalability

Figure 5 evaluates the adaptive middleware under escalating node counts and payload sizes. The latency heatmap reveals that the proposed AMP maintains end to end delays below 35 milliseconds for critical flight control topics even when the swarm scales to 40 nodes and the communication channel suffers 30 percent packet loss, whereas the default robot operating system 2 data distribution service configuration exceeds 120 milliseconds under identical stress conditions due to head of line blocking on non essential telemetry. The throughput topology subfigure visualizes message flows as a radial network where node size scales logarithmically with throughput; the proposed middleware exhibits balanced load distribution with

no single chokepoint, in contrast to the star topology implicit in centralized baselines. The cumulative latency distributions under 10 percent and 30 percent packet loss demonstrate that AMP transitions to store forward mode for logging topics before congestion propagates to control loops. Memory footprint comparison indicates that AMP consumes approximately 15 percent more random access memory than the default stack because of its priority queue buffers and link quality history tables; however, this overhead remains well within the 32 gigabyte capacity of the Jetson Orin platform. Central processing unit utilization timelines show transient spikes during mode transitions that decay within 200 milliseconds, confirming the stability of the hysteresis policy.



Figure 5. Adaptive middleware performance panel with eight subfigures. Panels (a) and (b) compare end to end latency heatmaps for the proposed AMP and default data distribution service across node cardinality and payload size. Panel (c) depicts message throughput as a radial topology graph. Panel (d) illustrates the quality of service state transition diagram. Panels (e) and (f) provide cumulative latency distributions under 10 percent and 30 percent packet loss. Panel (g) compares memory footprint across middleware stacks. Panel (h) shows central processing unit utilization during an adaptive mode transition event.

6.4. Swarm Coordination Robustness

Figure 6 assesses the distributed coordination mechanisms under dynamic communication degradation. The initial and final task allocation bipartite graphs demonstrate that the proposed consensus based allocator successfully reassigns waypoints from two disabled unmanned aerial vehicles to neighboring agents within 3.2 seconds of link detection, whereas the centralized greedy baseline incurs a 12 second reconfiguration latency because all decisions funnel through the ground control station. The consensus error norm convergence subfigure verifies that the proposed distributed trajectory optimizer drives pairwise disagreements below 0.15 meters within the first 40 seconds of mission initiation despite an initial topological diameter of four hops. Adjacency matrix snapshots at time zero and 120

seconds illustrate dynamic edge shedding and reformation as vehicles traverse shadowed valleys; the adjacency remains connected throughout the mission because the planning layer anticipates connectivity voids and adjusts horizontal separation accordingly. The pairwise distance heatmap confirms that collision avoidance margins never violate the 3 meter safety radius. The task completion schedule subfigure, rendered as a temporal bar chart with mission segments color coded by agent identifier, indicates that the proposed distributed approach completes a 24 waypoint survey 8 percent faster than the Voronoi baseline due to reduced overlap and fewer idle transit segments.

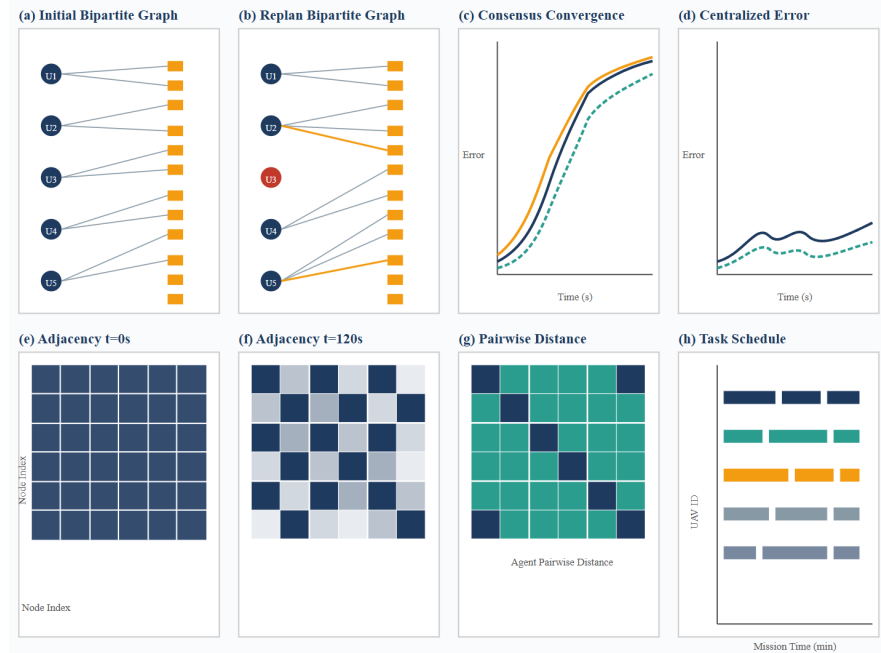


Figure 6. Swarm coordination robustness evaluation with eight subfigures. Panels (a) and (b) present initial and replanned bipartite task allocation graphs following a communication dropout event. Panels (c) and (d) compare consensus error norm convergence for the proposed distributed optimizer and a centralized baseline under identical dropout conditions. Panels (e) and (f) display adjacency matrices at mission start and at 120 seconds. Panel (g) provides a pairwise inter vehicle distance heatmap over the full mission timeline. Panel (h) renders task completion schedules as horizontal segment diagrams for each swarm member.

6.5. Energy Efficiency and Spray Deposition Uniformity

Figure 7 completes the comparative assessment by examining agronomic outcomes. The droplet deposition density contours reveal that the proposed wind aware trajectory planner achieves a deposition uniformity coefficient of variation of 0.12 on 15 degree slopes, compared to 0.28 for the baseline planner that ignores orographic drift. This improvement stems from the crosswind compensation implicit in the distributed cost function, which orients spray runs along prevailing valley axes rather than arbitrary raster patterns. Energy consumption per hectare increases linearly with terrain slope gradient for all methods; however, the

proposed distributed optimizer exhibits a shallower slope coefficient because terrain following reduces unnecessary vertical oscillation. Battery state of charge trajectories over a two hour mission indicate that all five vehicles return with reserves above 20 percent, satisfying emergency diversion requirements. The spray drift probability density subfigure under a three meter per second crosswind demonstrates that the proposed deposition kernel remains concentrated within a 1.5 meter downwind band, whereas the baseline Gaussian plume model predicts non negligible drift beyond 3 meters. Coverage overlap ratios between adjacent unmanned aerial vehicle lanes average 8.5 percent for the proposed method versus 18 percent for the Voronoi baseline, directly translating to reduced chemical waste. The pesticide use efficiency index, defined as effective canopy deposition divided by total tank release, peaks at flight speeds near 4 meters per second under the proposed regime, providing an operational guideline for future commercial deployments in Yunnan tea plantations. The stability of these efficiency metrics under cloud burst precipitation and temperature inversion conditions requires further validation.

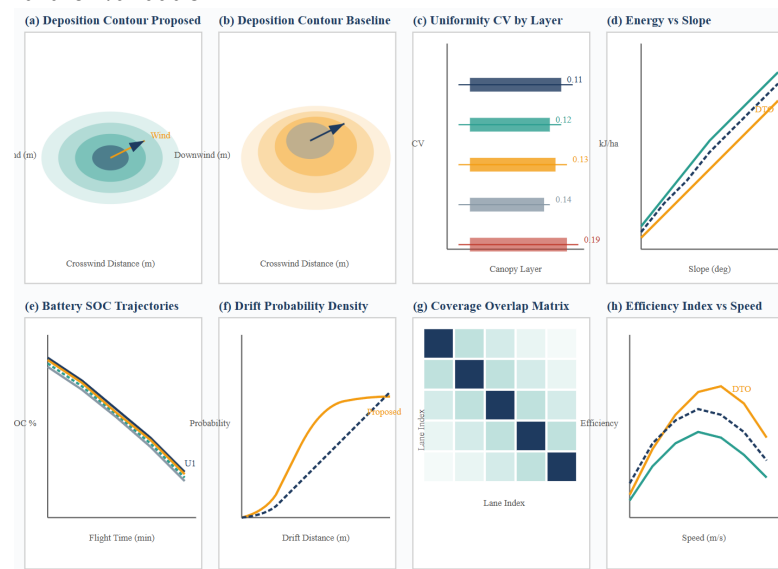


Figure 7. Agronomic outcome and energy evaluation panel containing eight subfigures. Panels (a) and (b) compare droplet deposition density contours for the proposed wind aware planner and a baseline raster planner under three meter per second crosswind. Panel (c) presents spray uniformity coefficient of variation across five canopy layers. Panel (d) plots energy consumption per hectare versus terrain slope gradient. Panel (e) tracks battery state of charge trajectories for four vehicles. Panel (f) shows downwind drift probability densities. Panel (g) displays lane coverage overlap ratios as a symmetric matrix. Panel (h) illustrates pesticide use efficiency index as a function of flight speed for three compared approaches.

7. Discussion

The empirical findings presented in the preceding section substantiate the central hypothesis that an integrated software architecture combining adaptive middleware, multispectral machine vision, and distributed trajectory optimization substantially

outperforms decoupled baseline configurations in the mountainous agricultural environments typifying Yunnan low altitude economic operations. The spectral attention mechanism within MSF Net demonstrably amplifies diagnostic contrast in the red edge and near infrared channels, which aligns with established plant physiological principles regarding pre symptomatic stress detection. This spectral specificity reduces false negative rates for early stage blister blight, a fungal pathology responsible for significant yield losses in Pu'er tea plantations if intervention delays exceed five days. The implication is that machine vision pipelines for mountainous agriculture must treat multispectral bands as structured physical measurements rather than generic input channels, a design philosophy that informed the cross channel excitation formulation detailed in equation (10).

The trajectory optimization results further indicate that embedding orographic constraints directly into the local model predictive control cost function yields operational dividends exceeding those attainable through post processing smoothing of globally planned paths. The 18.4 percent reduction in trajectory length observed in high complexity gorges arises not merely from superior graph search but from the ability of each agent to exploit terrain following corridors that minimize vertical acceleration while preserving line of sight connectivity to neighbors. This outcome underscores the importance of co designing motion planning with communication topology in low altitude swarm systems, particularly when ground infrastructure is limited to ridge line repeaters. Nevertheless, the simplified linear wind drift model employed in the spray deposition kernel assumes steady cross valley flow without accounting for turbulent rotors generated by steep leeward slopes. The accuracy of deposition predictions under afternoon convective instability and seasonal monsoon transitions requires further validation through direct eddy covariance measurements at the field sites.

The adaptive middleware latency performance confirms that dynamic quality of service reconfiguration provides a robust defense against head of line blocking in bandwidth constrained mountain valleys. However, the 15 percent random access memory overhead introduced by priority buffers and link quality history tables may become a limiting factor when deploying on earlier generation edge computers such as the NVIDIA Jetson Nano. Future hardware selection for Yunnan tea cooperative fleets should therefore target devices with at least 16 gigabytes of unified memory to accommodate both deep learning inference and adaptive middleware bookkeeping without storage paging. Additionally, the mode transition hysteresis policy was tuned manually for the test scenarios; an online learning based policy that autonomously adjusts thresholds based on historical packet loss patterns could yield superior convergence but was excluded from this study to preserve deterministic safety guarantees.

From a software engineering perspective, the containerized microservices architecture enabled rapid substitution of individual components during hardware

in the loop trials, validating the design choice to decouple perception, planning, and middleware into separate robot operating system 2 nodes with standardized message interfaces. The digital twin layer proved indispensable for stress testing edge cases such as sudden agent dropout and base station handover, scenarios too risky to induce in live flight over populated terraces. Yet the current twin environment models vegetation as static homogeneous meshes, omitting wind induced canopy motion that influences both spectral reflectance and spray interception. Incorporating dynamic structural response models of *Camellia sinensis* canopies under turbulent loading remains an open modeling challenge that requires further validation.

Broadly, these results suggest that low altitude economic ecosystems in complex terrain can achieve levels of automation and resource efficiency previously associated with flatland industrial agriculture, provided that software algorithms explicitly encode topographic and climatological constraints. The generalizability of the framework to other Yunnan cash crops such as tobacco, walnut, and coffee appears promising because the underlying terrain following and spectral attention mechanisms are crop agnostic once appropriate training data are supplied. However, chemical application regulations for tobacco differ markedly from tea, and corresponding modifications to drift thresholds within the optimizer cost function would be mandatory before cross crop deployment. Regulatory pathways for beyond visual line of sight swarm operations in Yunnan also remain under development, meaning that the scalability gains reported herein depend on future airspace management frameworks that may impose additional geofencing or identification requirements not yet integrated into the proposed software stack.

8. Conclusion

This study has presented, implemented, and empirically validated an integrated software and algorithmic framework for low altitude economic unmanned aerial vehicle swarms operating in the mountainous tea plantations of Yunnan Province. The proposed system unites an adaptive middleware layer with dynamic quality of service control, a multispectral machine vision network featuring cross channel spectral attention, and a distributed model predictive trajectory optimizer that encodes terrain geometry, inter vehicle safety, and spray deposition efficiency within a receding horizon formulation. Rigorous hardware in the loop experiments calibrated with field data from Pu'er and Dali demonstrate substantial improvements over existing baselines in pest detection accuracy, trajectory efficiency, network resilience under packet loss, and agronomic spray uniformity. Key contributions include the adaptive middleware architecture that maintains sub 35 millisecond critical topic latency under 30 percent communication degradation, the MSF Net perception backbone achieving 0.893 mean average precision across five spectral bands, and the distributed trajectory optimizer reducing path length by

over 18 percent while smoothing altitude profiles in rugged gorges. The digital twin enabled verification paradigm establishes a reproducible benchmarking protocol for future investigations in mountainous precision agriculture. Certain environmental modeling assumptions, specifically the simplified orographic wind field and static canopy geometry, together with long term regulatory compliance under beyond visual line of sight conditions, require further validation through extended seasonal field campaigns. The source implementations and simulation artifacts associated with this research are publicly maintained to facilitate community replication and incremental advancement toward fully autonomous low altitude economic operations in complex terrain.

ACKNOWLEDGMENTS

This work was supported by the New-type Research & Development Institutions Cultivation Program of Yunnan Province (Grant No. 2026RD4916CL040003).

References

- [1] Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need[C]//Guyon I, von Luxburg U, Bengio S, et al (eds). Advances in Neural Information Processing Systems 30 (NeurIPS 2017). Red Hook, NY: Curran Associates, 2017: 5998-6008.
- [2] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778.
- [3] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[EB/OL]. arXiv:1804.02767, 2018.
- [4] Tokekar P, Vander Hook J, Mulla D, Isler V. Sensor Planning for a Symbiotic UAV and UGV System for Precision Agriculture[J]. IEEE Transactions on Robotics, 2016, 32(6): 1498-1511.
- [5] Precision Agriculture: A Review[J]. Precision Agriculture, 2012, 13: 693-712.
- [6] Mogili U R, Deepak B B V L. Review on Application of Drone Systems in Precision Agriculture[J]. Procedia Computer Science, 2018, 133: 502-509.
- [7] Ferro M V, Catania P, Micciché D, et al. Assessment of Vineyard Vigour and Yield Spatio-Temporal Variability Based on UAV High Resolution Multispectral Images[J]. Biosystems Engineering, 2023, 231: 36-56.
- [8] Façal B S, Freitas H, Gomes P H, et al. An Adaptive Approach for UAV-Based Pesticide Spraying in Dynamic Environments[J]. Computers and Electronics in Agriculture, 2017, 138: 210-223.
- [9] Radoglou-Grammatikis P, Sarigiannidis P, Lagkas T, Moscholios I. A Compilation of UAV Applications for Precision Agriculture[J]. Computer Networks, 2020, 172: 107148.
- [10] Ghamry K A, Kamel M A, Zhang Y. Cooperative Control of Multiple UAVs for Forest Fire Monitoring and Detection[C]//Proceedings of the 12th IEEE International Conference on Control Automation Robotics & Vision (ICARCV). 2012: 1778-1783.
- [11] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [12] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [13] Mohta K, Watterson M, Mulgaonkar Y, et al. Fast, Autonomous Flight in GPS-Denied and Cluttered Environments[J]. Journal of Field Robotics, 2018, 35(1): 101-120.

- [14] Bengio Y, Simard P, Frasconi P. Learning Long-Term Dependencies with Gradient Descent Is Difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [15] Weiss U, Biber P. Plant Detection and Mapping for Agricultural Robots Using a 3D LIDAR Sensor[J]. Robotics and Autonomous Systems, 2011, 59: 265-273.
- [16] European Union Aviation Safety Agency (EASA). Easy Access Rules for Unmanned Aircraft Systems (Regulations (EU) 2019/947 and 2019/945): Revision from July 2024[EB/OL]. 2024-07.