Ground-Station-Free Multi-UAV Formation with an Airborne Wi-Fi Hub and Distributed Avionics

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

Name omitted for review
Dept name omitted for review
Organization name omitted for review
City, Country omitted for review
Email address omitted for review

***Abstract*—Fieldable multirotor swarms often depend on a ground link or high-bandwidth perception, which limits robustness and range. We report a lightweight system that removes the ground station and still keeps formation tight by pairing an airborne master-slave network with ENU-referenced control and onboard person counting. The master serves as a Wi-Fi access point and timing anchor; followers receive compact formation cues in the local ENU frame and track them. We validate with two followers in AirSim and outdoor flights. On a 5 m circular formation, followers hold sub-meter radial RMSE and show bounded step responses for ±3 m formation perturbations with small overshoot. In-flight networking yields round-trip delay of ≈200 ms; latency-injection trials quantify sensitivity, with orbit RMSE increasing by 0.06 m per additional 100 ms of round-trip time (RTT), explaining part of the sim-to-field gap. The onboard perception stack transmits only integer counts yet attains F1 up to 0.81 across different scenes. Together, these results indicate that a bandwidth-efficient, airborne hub can deliver reliable, low-latency multi-UAV coordination without a ground station. Limitations include scaling beyond two followers, operation in stronger winds and longer missions.**

Keywords—Unmanned aerial vehicles, swarm drones, formation control, localization, airborne network.

I. Introduction

Effective aerial surveillance in remote or rugged border regions demands wide-area coverage and low alert latency. Ground patrols and fixed sensors struggle to maintain continuous monitoring over uneven terrain, whereas unmanned aerial vehicle (UAV) swarms can distribute sensing and extend coverage-provided coordination and communication remain reliable at low latency. Prior work has explored border monitoring with UAVs [1], yet achieving low-latency, long-endurance operation over hilly terrain remains challenging.

Communication architectures for multi-UAV systems are broadly centralized or decentralized. Centralized schemes (ground relay, star/hierarchical star) simplify command and control but introduce a single point of failure and can suffer under weak or congested ground links [2], [3]. Decentralized designs exchange information peer-to-peer or via airborne relays, improving scalability and resilience to link drops; Wi-Fi mesh is often proposed to remove dependence on fixed ground infrastructure [4]. Networking choices interact with routing: traditional protocols degrade under high mobility, motivating geo-aware routing and mesh standards to stabilize latency and throughput in dynamic topologies [1], [4], [5].

Formation control commonly follows leader–follower or virtual-structure paradigms. Leader–follower assigns fixed offsets to followers relative to one or more leaders and is favored for simplicity and predictability; robust variants address leader failure and delay (e.g., vision-based relative pose tolerating in 1 s delay) [6], [7]. Virtual-structure treats the team as a rigid body for shape translation/rotation with flexible reconfiguration, typically at higher communication cost [6]. In all cases, latency and packet loss can degrade stability, motivating delay-aware or robust controllers [6], [7]. Collision avoidance and relative localization are often implemented with lightweight onboard cues rather than bulky range sensors: exchanged state plus radio RSSI can support separation indoors with micro-air-vehicles [8], while vision-based approaches detect neighboring UAVs and motors to estimate relative pose for formation and avoidance [9].

Modern small-form-factor perception increasingly runs onboard. One-stage neural network detectors (YOLO family) offer real-time inference on embedded hardware; lightweight variants tailored to small targets maintain high frame rates with modest accuracy loss [10]. Nonetheless, much of the literature reports simulation or lab-scale demos, with limited outdoor validation that jointly quantifies formation accuracy, collision-avoidance behavior, network latency and onboard detection performance across diverse scenes.

This paper develops and evaluates a multi-copter swarm that integrates formation control, airborne networking, and onboard person counting, with outdoor field validation. Our contributions are:

1) A Master–Slave controller in an ENU frame achieves sub-meter radial error on a 5 m circular orbit using a 36-point track.
2) Formation expansion/contract cues (±3 m) execute with small, repeatable overshoot under a 2 m/s forward-speed limit.
3) A lightweight detector runs on the slaves and transmits only the integer count; results are reported for indoor/daylight with objects, indoor/low-light, and outdoor/in-flight scenes.

Each entry in Table I indicates whether the prior system was validated in field flight, whether it employed ENU-relative positioning between UAVs, whether quantitative formation accuracy was reported, whether object detection was executed onboard, whether collision avoidance was achieved without external sensors and whether an airborne UAV served as a network hub.

TABLE I. FEATURE COVERAGE IN RELATED WORK

Ref.	AH†	ENU†	FA†	EV†	CA†
Li 2016 [12]	Yes	Yes	No	No	No
Coppola 2018 [8]	No	No	No	No	Yes
Li 2020 [11]	No	Partial	Yes	No	No
Si 2023 [9]	No	No	No	Yes	Yes
Devey 2024 [13]	Partial	Yes	Yes	No	No
This work	Yes	Yes	Yes	Yes	Partial

†AH = airborne network hub; ENU = localization in East-North-Up frame; FA = formation-keeping accuracy reported quantitatively; EV = onboard edge-vision inference; CA = collision avoidance without external ranging/beacon sensors.

II. SYSTEM OVERVIEW

The swarm operates as an airborne star network, with the Master UAV serving as the central hub. The Master UAV hosts a Wi-Fi access point and a lightweight mission server, while the Slave UAVs independently execute their flight paths and perform onboard perception tasks. All vehicles share their positions relative to a common East-North-Up (ENU) coordinate frame anchored at the Master's initial GPS location.

During flight, the Master continuously transmits concise mission updates and periodic heartbeat signals while receiving status reports from each Slave. Slaves, in turn, broadcast their states including position, planar velocity and heading multiple times per second through a single UDP one-to-many message. This approach enables all vehicles to maintain a shared, real-time awareness of the swarm's configuration without additional communication overhead.

Fig. 1 illustrates these links. The Master issues formation updates at a fixed rate and expects brief acknowledgments to confirm delivery. Each Slave broadcasts position, planar velocity and heading several times per second using a single UDP one-to-many message, allowing every vehicle to perceive nearby peers without additional wiring. A simple heartbeat provides liveness: if several beats are missed, the affected vehicle enters hover until the link is restored. When a vehicle detects a problem- low battery, loss of GPS or an onboard fault, it announces the event immediately; the Master relays it to all nodes.

All messages include a timestamp and a sequence number. No ground station is required during flight; the network hub is onboard the Master. Table II lists the message types, and Table III outlines the roles of each platform within the swarm.

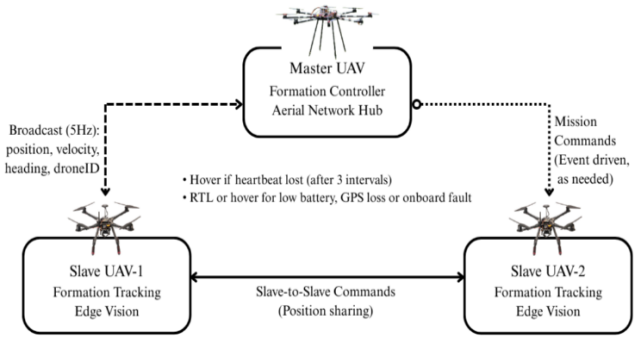

Fig. 1. Airborne master-slave architecture

TABLE II. MESSAGE CATALOG

Type	Direction	Rate (Hz)	Payload (bytes)
Heartbeat	Master < > Slaves	1	8–16
Formation Command	Master > Slaves	10	32–64
Pose	Master < > Slaves	5	48
Alert	Slaves > Master	—	24

TABLE III. MINIMAL PLATFORM SNAPSHOT

UAV	Specs	Companion	Role
Master (octacopter)	Weight: 7 kg Endurance: 24 minutes	Raspberry Pi 5 + Wi-Fi AP + Pixhawk	Mission + Network Hub + Logger
Slave (quadcopter)	Weight: 2 kg Endurance: 13 minutes	Raspberry Pi 5 + Camera + Pixhawk	Formation Tracking + Edge Vision

III. METHODS OF CONTROL AND PERCEPTION

A. Reference Frames and ENU Transformation

All formation and avoidance logic operates in a local East-North-Up (ENU) frame anchored at the Master UAV's initial GPS fix ($\varphi_0, \lambda_0, h_0$). Each slave converts its geodetic position (φ, λ, h) to Earth-Centered Earth-Fixed (ECEF): a = 6378137 m & $e^2 = 6.69437999014 \times 10^{-3}$.

$$N_\varphi = \frac{a}{\sqrt{1-e^2\sin^2\varphi}} \tag{1}$$

$$\begin{aligned} x &= (N_\varphi + h)\cos\varphi\cos\lambda \\ y &= (N_\varphi + h)\cos\varphi\sin\lambda \\ z &= \left((1-e^2)N_\varphi + h\right)\sin\varphi \end{aligned} \tag{2}$$

The displacement vector $\mathbf{\Delta r} = [x - x_0, y - y_0, z - z_0]^T$ is rotated into ENU using:

$$\mathbf{R}_{ecef\to enu} = \begin{bmatrix} -\sin\lambda_0 & \cos\lambda_0 & 0 \\ -\sin\varphi_0\cos\lambda_0 & -\sin\varphi_0\sin\lambda_0 & \cos\varphi_0 \\ \cos\varphi_0\cos\lambda_0 & \cos\varphi_0\sin\lambda_0 & \sin\varphi_0 \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} E \\ N \\ U \end{bmatrix} = \mathbf{R}\,\Delta\mathbf{r} \tag{4}$$

For missions confined within ≈1km, we also validated a small-area approximation:

$$N = (M_0 + h_0)(\varphi - \varphi_0),$$
$$E = (N_0 + h_0)\cos\varphi_0(\lambda - \lambda_0),\ U = h - h_0 \tag{5}$$

where M and N are radii of curvature at latitude φ.

B. Formation Tracking

The Master broadcasts formation cues at 10 Hz. Each Slave computes its target waypoint by rotating its assigned body-frame offset ($\Delta x_i, \Delta y_i, \Delta z_i$) into the master's ENU frame:

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + R(\psi_0)\begin{bmatrix} \Delta x_i \\ \Delta y_i \end{bmatrix} \tag{6}$$

With $R(\psi_0)$ the 2D rotation by master yaw ψ. Three primitives are supported:

1) Circular Orbit: Slaves evenly spaced on a circle of radius r around the Master.
2) Echelon: Fixed lateral offsets aligned with the Master's heading.
3) Expansion-Contraction: Master broadcasts a Δr; slaves update $r' = r \pm \Delta r$.

Each slave tracks its ENU setpoint using a cascaded PID:

- Outer Loop (10Hz): position error → desired velocity.
- Inner Loop (50Hz, ArduPilot): velocity → attitude commands.

C. Distributed Collision Avoidance

Each Slave maintains peer states from POSE broadcasts (5–10 Hz). At 20 Hz, a constant-velocity predictor checks pairwise separations 1 s ahead. A neighbor j is considered risky if $d_{ij(t+1)} < 3$m. The avoidance policy is:

1) Conflict Resolution: The lower-priority agent adjusts its trajectory.
2) Yaw Deflection: Heading increment $\Delta\psi_{cmd} = \text{clip}(\psi^* - \psi, -2.5°, +2.5°)$.
3) Speed Scaling: When saturated, forward speed is reduced to 50% for 0.5 s.
4) Recovery: Once separation exceeds 3.2 m, the slave realigns with formation.

This runs in 0.5 ms per cycle on Raspberry Pi 5, ensuring feasibility for larger swarms. The full decision flow is illustrated in Fig. 2.

D. Detection Pipeline

Each Slave UAV is equipped with a Raspberry Pi 5 (8 GB) and a Pi Camera V3 NoIR. The camera is front-mounted with a 70° horizontal field of view. The Pi executes all perception routines locally and transmits compact detection metadata to the Master via the airborne Wi-Fi AP. This design avoids raw video streaming, conserving bandwidth. The perception stack runs at 8 frames per second (fps). We utilize the pre-trained YOLOv8n object detector, selected for its lightweight architecture and favorable balance between inference speed and detection accuracy, as demonstrated on public benchmark datasets. The processing pipeline is:

- Image acquisition (8 fps): Frames are timestamped with the same monotonic clock as telemetry.
- Pre-processing: Resized to 640 x 640 px; normalized to [0,1].
- YOLOv8n inference: Obtain detections (b_k, c_k, γ_k).
- Filtering: Keep only the person class; apply NMS. (Confidence is used internally for NMS but is not transmitted).
- Counting and packaging
  $n_{person} = \#\{k \mid c_k = \text{person (after NMS)}\}$
  Broadcast a single UDP packet per frame:
  $\{ \text{id}, t_{ms}, E, N, U, n_{person} \}$.

  This replaces any per-detection payloads or confidence values.

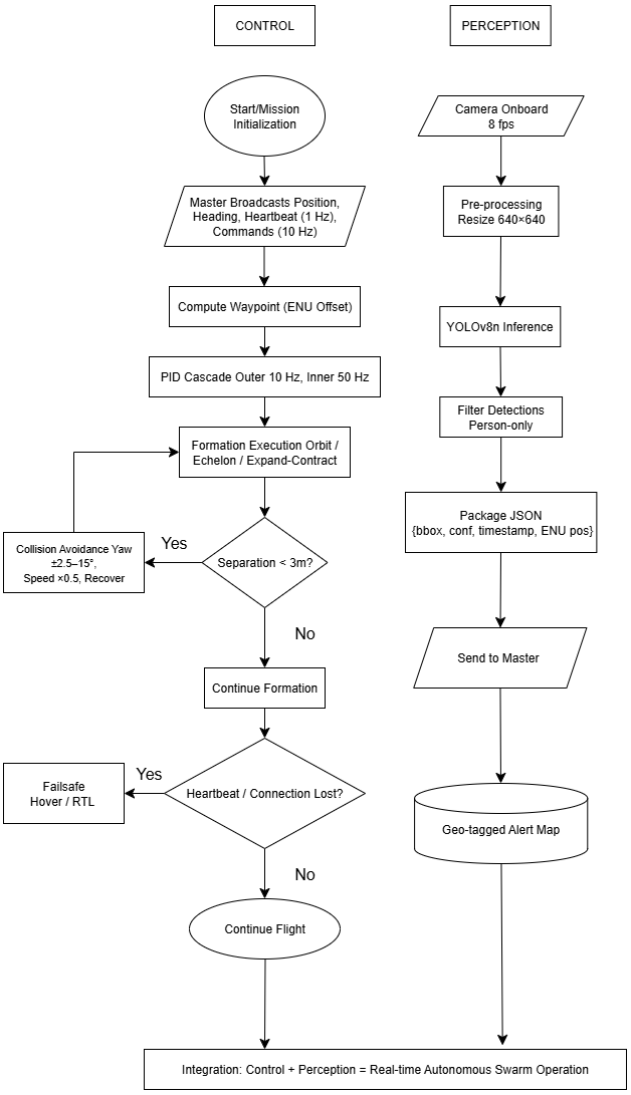

Fig. 2. Control-perception flowchart.

E. Failsafe Policies

Safety logic runs concurrently with control:

- Heartbeat Loss: Three missed master beats → Hover.
- Low battery: Broadcast ALERT; vehicle performs RTL, peers re-form.
- GPS loss: Climb +5m and hover until recovery.

Failsafe behaviors are implemented using ArduPilot safety modes triggered by the companion computer.

IV. Experimental Setup

Flights were performed on an open soccer field under sunny condition with wind speeds of 3–4 m/s. GNSS quality supported ArduPilot Loiter throughout. All coordination used an airborne Wi-Fi AP mounted on the Master; no ground station was used during flight.

The fleet comprised one Master octocopter and two Slave quadcopters. The Master hosted the AP and ran the mission script and logger. Slaves executed formation tracking, local safety and onboard person detection; only a per-frame person count with timestamp and ENU pose was transmitted to the Master. Manual RC override was available at all times.

Each run followed the same script: the Master took off to 8–10 m AGL and hovered; on cue the Slaves armed and took off, executed the assigned formation, then landed; the Master landed last after a mission-complete handshake. Two scenarios were evaluated:

1) Circular Orbit: Slaves orbited the hovering Master at r = 5 m, maintaining tangent heading.
2) Expansion-Contraction: Vehicles began as a right triangle on the ground projection with the Master at the 90° node; after take-off to a formation hover, the Master issued +3 m expand and -3 m contract commands, and Slaves stepped their ENU radial offsets accordingly.

Pixhawk logs provided GNSS, ENU pose, heading, and timestamps; the Master logged command sequence/ACK.

To isolate networking effects, we repeated the mission in Microsoft AirSim using the identical mission script and controller. AirSim's Multirotor API supplied ground-truth poses at 100 Hz. Communication delay was emulated on both paths (Master→Slave commands and Slave→All state broadcasts) with fixed one-way delay $\delta \in \{0, 25, 50, 100, 150\}$ ms and 10% jitter (RTT = 2δ); packet loss was held at 4% to isolate delay. For each δ, we flew a 2-min orbit at r = 5 m and executed two ±3 m steps. Error metrics were computed per run; results are reported as the median of three runs. Sensitivity is summarized by the slope of RMSE radial error versus RTT (m per 100 ms).

From this experimental setup, we computed:

a) Circular orbit (agent i, samples k=1..N, position $p_i(t_k)$, orbit center p_0):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_k \left(\|p_i(t_k)-p_0\|-r\right)^2} \quad (7)$$

$$\text{Max radial error} = \max_k \left|\|p_i(t_k)-p_0\|-r\right| \quad (8)$$

$$\delta_L = \frac{L_{flown}-2\pi r}{2\pi r}\times 100\% \quad (9)$$

$$\overline{|e_\psi|} = \frac{1}{N}\sum_k \left|\text{wrap}\left(\psi_i(t_k)-\psi_{tan}(t_k)\right)\right| \text{ (deg)} \quad (10)$$

Where ψ_{tan} is the ideal tangent heading and wrap(.)∈(-π,π].

b) Expansion-contraction (radial step Δr = ±3m):

$$t_{rise} = \min\{t \geq t_{cmd} : |r(t)-r(t_{cmd})| \geq 0.95|\Delta r|\} \quad (11)$$

$$\%OS = 100\cdot \frac{\max_t |\Delta r_{ach}(t)|-|\Delta r|}{|\Delta r|} \quad (12)$$

Command→response latency is

$$\Delta t_{cmd} = t_{first\text{-}motion} - t_{cmd} \quad (13)$$

With $t_{first\text{-}motion}$ the first sustained increase in speed satisfying |v|>0.2 m/s for 0.2s.

c) Detection pipeline (per frame):

$$\text{Precision} = \frac{TP}{TP+FP},\ \text{Recall} = \frac{TP}{TP+FN},\ F1 = \frac{2(\text{Precision})(\text{Recall})}{\text{Precision}+\text{Recall}} \quad (14)$$

V. Result and Discussion

With the Master fixed at the origin and a target radius of 5 m, both Slaves executed a 36-point CCW orbit (10° spacing). The realized tracks in Fig. 3 closely follow the ideal circle. Table IV shows sub-meter radial accuracy (RMSE < 0.7 m) and tightly matched performance between the two vehicles. Path-length inflation is 10–11%, which is expected from chord-to-arc discretization plus small tangential jitter; it does not indicate a persistent radial bias. The mean absolute heading error is on the order of ten degrees, consistent with piecewise-linear motion around a curved path, and has negligible impact on inter-vehicle spacing because radial control dominates separation.

Fig. 4 plots the median RMSE versus latency. Error increases monotonically from 0.35 m at 75 ms to 0.48 m at 300 ms, i.e., a slope of ≈0.06 m per 100 ms RTT. These values lie below the field RMSE in Table IV (0.631–0.638 m) because the simulator provides ideal pose sensing and no wind; however, at the ≈200 ms delays measured in flight using MAVLink ping message, the curve predicts 0.36–0.40 m due to networking alone, attributing the remaining gap to real-world disturbances, estimator noise and waypoint discretization.

Starting from a right-triangle formation with 5 m legs, a radial step of +3 m followed by -3 m was commanded under a 2 m/s forward-speed limit. Plan-view traces in Fig. 5 exhibit small, repeatable overshoot with clean settling. It also indicates larger overshoot on expansion than contraction,

TABLE IV. Circular Orbit Error Metrics

Metric	Slave 1	Slave 2
RMSE Radial Error	0.638 m	0.631 m
Max Radial Error	1.342 m	1.405 m
Path Length Error	11.30%	10.52%
Mean Absolute Heading Error	13.18°	11.54°

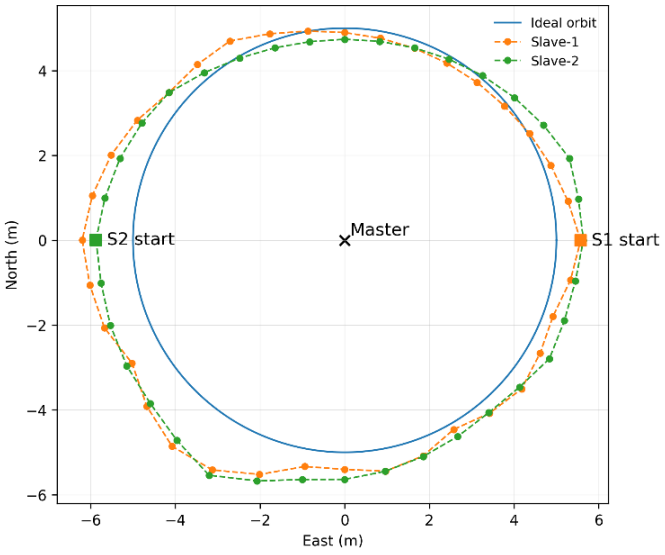

Fig. 3. Circular-orbit tracking at r = 5 m: ideal circle and realized paths for Slave-1 and Slave-2 (Master at origin; CCW slaves' start points marked).

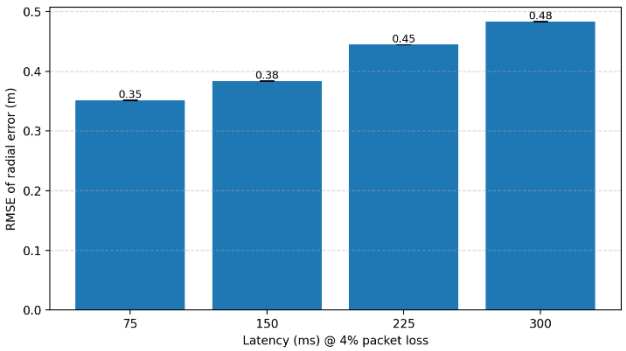

Fig. 4. Formation radius RMSE vs latency under 4% packet loss

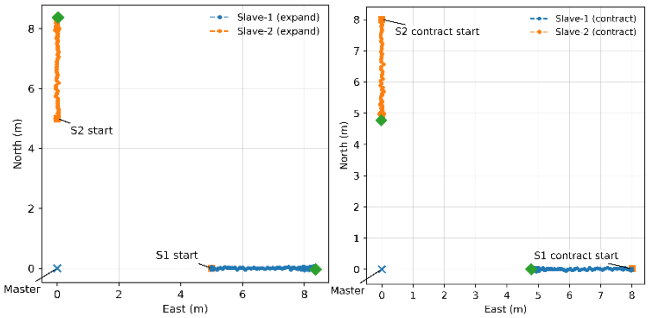

Fig. 5. (a) Expansion step from 5 m to 8 m. (b) Contraction step from 8 m to 5 m. Overshoot peaks annotated.

consistent with outward acceleration and derivative damping under velocity saturation. The responses are monotone with no oscillation, indicating adequate phase margin for the chosen gains.

Fig. 6 shows the per-frame output used to form the transmitted Person count. Across the three evaluation conditions in Table V, indoor/daylight, indoor/low-light and outdoor/in-flight performance follows the expected pattern. In indoor/daylight with objects, illumination is stable and targets are near scale, so recall is highest and precision is limited mainly by object look-alikes and shadows. In indoor/low-light, reduced SNR and motion-induced blur lower both precision and recall, producing the weakest F1; locking exposure and biasing the detector toward higher ISO or longer integration would help. In outdoor/in-flight, platform vibration, perspective changes and small apparent size increase false negatives relative to indoor daylight, but temporal continuity moderates false positives. Since the system transmits only the count, we mitigate jitter by applying a short temporal median on the last 5–7 frames; small-scale recall can be improved further with multi-scale inference or a lower score threshold combined with this temporal filter, without increasing network load. Overall, the swarm holds sub-meter radial accuracy on a 5 m orbit while executing discrete formation cues with bounded overshoot. The symmetry between the two Slaves suggests the control behavior is robust to minor platform differences. The path-

TABLE V. PERSON-DETECTION PERFORMANCE METRICS

Scenario	Precision (%)	Recall (%)	F1-Score (%)
Indoor-daylight	77	85	81
Indoor-low-light	64	70	67
In Flight (Sunny day)	72	78	75

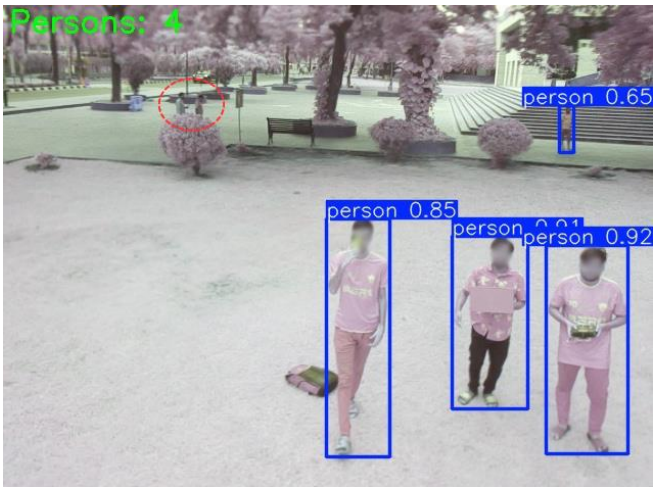

Fig. 6. Onboard person detection from a slave during an outdoor trial (faces de-identified for publication, consent was taken from the individuals). Blue boxes are detections; the overlay is the transmitted person count; two upper-left individuals (dashed red) are false negatives.

length inflation and heading error, are consistent with waypoint discretization and can be reduced by denser angular sampling or a tangent-aligned spline tracker. For expansion/contraction, a small integral term with anti-windup or a scheduled derivative gain would further reduce overshoot at the 2 m/s limit. For perception, the dominant failure mode is small/occluded targets; multi-scale inference or a short temporal vote on the count would be the most impactful next steps.

VI. CONCLUSION

This paper presented and field-validated a multi-copter swarm that couples an airborne master-slave network with ENU-referenced formation control and onboard person counting. Outdoor flights achieved sub-meter formation accuracy on a 5 m orbit and bounded step responses for ± 3 m formation cues with small overshoot. The perception stack runs on the slaves and transmits only the integer count, achieving F1 up to 81% across indoor-daylight, indoor-low light and in-flight scenes. With an airborne Wi-Fi hub and sequence-numbered messaging, round-trip delay remained ≈200 ms in-flight; AirSim sensitivity indicates orbit RMSE increases by about 0.06 m per 100 ms RTT, explaining part of the gap between field and simulation.

These results indicate that a lightweight airborne hub with message-efficient perception can deliver reliable, low-latency coordination without a ground station. Remaining limitations include scaling beyond two followers, operation in stronger winds and longer missions, and small-target recall. Future work will explore (i) anti-windup and scheduled gains to further reduce overshoot, (ii) multi-hop/mesh networking for range and resilience, and (iii) multi-scale and temporal counting to raise F1 while keeping bandwidth constant.

REFERENCES

[1] X. Chen, J. Tang, and S. Lao, "Review of unmanned aerial vehicle swarm communication architectures and routing protocols," *Applied Sciences*, vol. 10, no. 10, May 2020, doi: 10.3390/app10103661.

[2] Manju, Prabha, Farithkhan, Sivagurunathan and M. Mahesh, "Decentralized control design for UAV swarms communication,"

Discover Applied Sciences, vol. 7, no. 2, Feb. 2025, doi: 10.1007/s42452-024-06408-w.

[3] M.L. Cummings, “Operator Interaction with Centralized Versus Decentralized UAV Architectures,” in *Handbook of Unmanned Aerial Vehicles*, K. P. Valavanis and G. J. Vachtsevanos, Eds. Dordrecht, Netherlands: Springer, 2015, doi: 10.1007/978-90-481-9707-1_117.

[4] M. A. Lopez, M. Baddeley, W. T. Lunardi, A. Pandey and J. -P. Giacalone, "Towards Secure Wireless Mesh Networks for UAV Swarm Connectivity: Current Threats, Research, and Opportunities," in *Proc. 17th Int. Conf. Distributed Computing in Sensor Systems (DCOSS)*, Pafos, Cyprus, 2021, pp. 319–326, doi: 10.1109/DCOSS52077.2021.00059.

[5] G. Asaamoning and P. Mendes, “A Position-based Hybrid Routing Protocol for Clustered Flying Ad Hoc Networks,” *Drones and Autonomous Vehicles*, vol. 1, no. 2, 2024, doi: 10.35534/dav.2024.10001.

[6] Y. Bu, Y. Yan, and Y. Yang, “Advancement Challenges in UAV Swarm Formation Control: A Comprehensive Review,” *Drones*, vol. 8, no. 7, Jul. 2024, art. 320, doi: 10.3390/drones8070320.

[7] N. Evangeliou, D. Chaikalis, A. Tsoukalas, and A. Tzes, “Visual Collaboration Leader-Follower UAV-Formation for Indoor Exploration,” *Frontiers in Robotics and AI*, vol. 8, Jan. 2022, doi: 10.3389/frobt.2021.777535.

[8] M. Coppola, K. N. McGuire, K. Y. W. Scheper, et al., “On-board communication-based relative localization for collision avoidance in Micro Air Vehicle teams,” *Autonomous Robots*, vol. 42, pp. 1787–1805, 2018, doi: 10.1007/s10514-018-9760-3.

[9] Si, Xiaokun & Xu, Guozhen & Ke, Mingxing & Zhang, Haiyan & Tong, Kaixiang & Qi, Feng, “Relative Localization within a Quadcopter Unmanned Aerial Vehicle Swarm Based on Airborne Monocular Vision” *Drones*, vol. 7, no. 10, 2023, art. 612, doi: 10.3390/drones7100612.

[10] Y. Feng, T. Wang, Q. Jiang, C. Zhang, S. Sun, and W. Qian, “A Efficient and Accurate UAV Detection Method Based on YOLOv5s,” *Applied Sciences (Switzerland)*, vol. 14, no. 15, Aug. 2024, doi: 10.3390/app14156398.

[11] S. Li, F. Shan, J. Liu, M. Coppola, C. de Wagter and G. C. H. E. de Croon, "Onboard Ranging-Based Relative Localization and Stability for Lightweight Aerial Swarms," *IEEE Robotics and Automation Letters*, vol. 10, no. 10, pp. 10066-10073, early access, Oct. 2025, doi: 10.1109/LRA.2025.3597039

[12] B. Li, Y. Jiang, J. Sun, L. Cai, and C. Y. Wen, “Development and testing of a two-UAV communication relay system,” *Sensors (Switzerland)*, vol. 16, no. 10, Oct. 2016, doi: 10.3390/s16101696.

[13] J. Devey, P. S. Gill, G. Allen, E. Shahra, and M. Idrissi, “Network-Centric Formation Control and Ad Hoc Communication with Localisation Analysis in Multi-UAV Systems,” *Machines*, vol. 12, no. 8, Aug. 2024, doi: 10.3390/machines12080550.