arXiv:2509.14210v3 [cs.RO] 5 Mar 2026

GLIDE: A Coordinated Aerial-Ground Framework for Search and Rescue in Unknown Environments

Seth Farrell[1], Chenghao Li[1], Hesam Mojtahedi[1] and Henrik I. Christensen[1]

***Abstract*— We present a cooperative aerial-ground search-and-rescue (SAR) framework that pairs two unmanned aerial vehicles (UAVs) with an unmanned ground vehicle (UGV) to achieve rapid victim localization and obstacle-aware navigation in unknown environments. We dub this framework Guided Long-horizon Integrated Drone Escort (GLIDE), highlighting the UGV's reliance on UAV guidance for long-horizon planning. In our framework, a goal-searching UAV executes real-time onboard victim detection and georeferencing to nominate goals for the ground platform, while a terrain-scouting UAV flies ahead of the UGV's planned route to provide mid-level traversability updates. The UGV fuses aerial cues with local sensing to perform time-efficient A* planning and continuous replanning as information arrives. Additionally, we present a hardware demonstration (using a GEM e6 golf cart as the UGV and two X500 UAVs) to evaluate end-to-end SAR mission performance and include simulation ablations to assess the planning stack in isolation from detection. Empirical results demonstrate that explicit role separation across UAVs, coupled with terrain scouting and guided planning, improves reach time and navigation safety in time-critical SAR missions.**

I. Introduction

Search and rescue (SAR) operations stand to benefit from recent advances in autonomous aerial and ground robotics. Unmanned Aerial Vehicles (UAVs) enable rapid, large-area coverage due to their agility and mobility. The adoption of drones across civilian and military applications has highlighted advantages in speed and perspective. In SAR, aerial platforms facilitate rapid victim localization and provide global views of disaster zones. They enable first responders to identify both victims and safe access routes, thereby reducing response time. Unmanned Ground Vehicles (UGVs) offer persistence, greater payload capacity, and the ability to physically reach and assist victims. However, realizing these complementary strengths in a single operational system remains challenging due to limited UAV endurance, local myopic planning on UGVs, and robust online coordination in unstructured, time-critical environments [1], [2].

Prior SAR deployments have heavily relied on teleoperation [3], wherein human operators must continuously monitor video feeds and sensor data, make decisions, and directly control the robots in real time. In disaster environments, this quickly becomes overwhelming, particularly when multiple robots must be managed simultaneously. Moreover, teleoperation requires a reliable, high-bandwidth communication link between robots and the operator control center. In SAR scenarios (e.g., collapsed buildings, tunnels, wildfire zones), wireless signals are often obstructed or delayed. Latency in control commands or video feedback can render precise maneuvering extremely difficult or even infeasible. Finally, teleoperation underutilizes recent advances in learning-based perception, onboard inference, and edge autonomy, limiting both scalability and effectiveness.

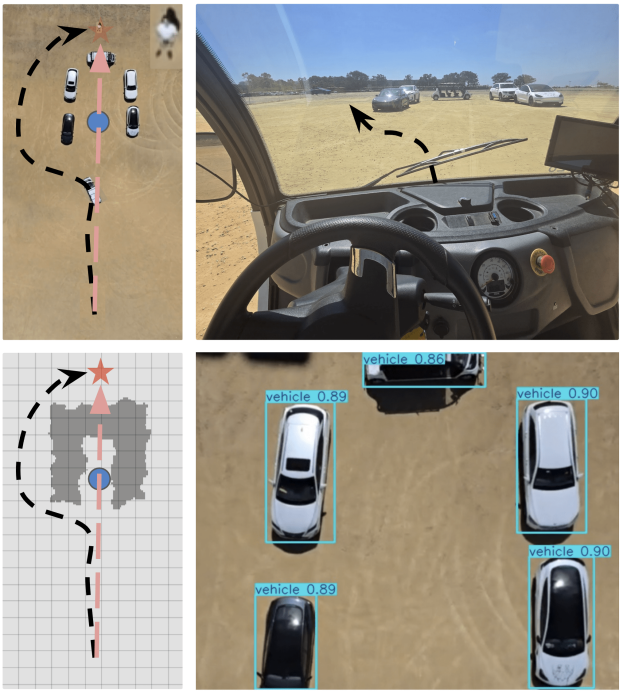

Fig. 1: (Top Left) Real-world experimental visualization of the proposed end-to-end aerial-ground SAR framework. The test environment includes a U-shaped corridor at the center, with the victim located at the top. The dashed pink line indicates the UGV's planned trajectory prior to receiving terrain information from the terrain-scouting UAV, while the blue dot marks the footprint of the terrain-scouting UAV. The black dashed line indicates the ego-vehicle's planned trajectory after re-planning occurs. (Top Right) Ego-vehicle perspective as it approaches the U-shaped obstacles. (Bottom Left) Occupancy grid generated by our GLIDE method, smoothed over time, where dark gray regions denote detected non-traversable areas. The ego vehicle uses this map to plan avoidance maneuvers. (Bottom Right) Object detection results from aerial imagery captured by the terrain-scouting UAV. The detection module outputs bounding boxes for identified objects along with their associated confidence scores.

Search-and-Rescue (SAR) operations are driven by two urgent imperatives: the need to rapidly localize victims and the necessity of delivering immediate support in unstructured, partially known environments such as uneven terrain, forests, or flooded streets. However, achieving these objectives is highly challenging due to the inherent trade-offs of different robotic platforms. UAVs, despite their speed

[1] Affiliation: 1. UCSD Email: swfarrel@ucsd.edu

and wide-area coverage, are constrained by limited flight time and payload capacity. UGVs, on the other hand, offer endurance and robustness on the ground but lack the rapid mobility and global situational awareness of aerial systems. The cooperation between UAVs and UGVs has emerged as a central research direction in robotics, driven by their complementary strengths [4] [5] [6]. This recognition has inspired numerous research efforts and competitions designed to evaluate and extend the capabilities of UAV-UGV teams in realistic settings.

We address these challenges with a cooperative multi-agent framework that integrates two aerial roles with a ground ego-vehicle for end-to-end autonomous SAR. A *goal-searching* UAV conducts fast site surveys, detects potential victims onboard, and transmits compact georeferenced messages to the ground platform. A *terrain-scouting* UAV scouts ahead of the UGV's current plan, providing mid-level situational awareness by flagging non-traversable or hazardous regions. The UGV fuses these aerial cues to perform long-horizon planning and online re-planning toward locations of confirmed victims. This division of responsibilities accelerates victim discovery while enabling safe and time-efficient routing. The key contributions of this paper are summarized as follows:

- A cooperative aerial-ground SAR framework that explicitly separates goal-searching and terrain-scoutting aerial roles to complement the ground platform.
- A lightweight perception stack for real-time detection on resource-constrained UAVs.
- An aerial-guided planning stack that uses situational maps (from the terrain-scouting UAV) to generate time-efficient, collision-free UGV trajectories.
- A hardware demonstration using a GEM e6 UGV and two X500 UAVs that provides a proof of concept for end-to-end SAR automation.
- Simulation ablations that robustly evaluate the proposed planning stack in isolation from perception.

II. Related Works

A landmark initiative in advancing search-and-rescue (SAR) robotics was the *DARPA Subterranean (SubT) Challenge*, in which participants developed robotic platforms capable of navigating underground environments and detecting designated objects. Similar to how the DARPA Urban Grand Challenge accelerated self-driving research, the SubT Challenge spurred significant progress in SAR robotics. Key outcomes included advances in UAV control within confined spaces, edge-deployable detection models, and both homogeneous and heterogeneous multi-agent planning approaches [7], [8].

Following the SubT Challenge, research on combined UGV–UAV systems has expanded considerably. Survey works [9], [10], [11] highlight the complementary strengths of UAVs and UGVs, reviewing progress in communication, perception, and coordination. These studies also identify persistent challenges such as autonomy in unstructured environments, interoperability across platforms, and adaptation to dynamic operational conditions.

Building on these insights, several frameworks have been proposed to operationalize UAV–UGV cooperation. Wang et al. [12] introduced an aerial–ground exploration strategy where UAVs scout unexplored regions to accelerate UGV mapping, significantly reducing mission time. Gawel et al. [13] demonstrated that UAVs providing third-person perspectives can improve UGV teleoperation, underscoring the benefits of complementary sensing viewpoints. Oh et al. [14] examined SAR scenarios under degraded communication and GPS-denied conditions. They proposed a genetic search algorithm to enable coordination when no control-station link is available, though their approach requires survivors to carry transmission devices.

More recent works extend SAR autonomy beyond UAV-focused solutions. Zafar et al. [15] employed a quadruped robot equipped with a fine-tuned YOLOv8 detector, coupled with gesture-based control, to improve victim detection and interaction in disaster settings. Other efforts [16] highlight the limitations of assuming static environments, proposing UAV–UGV cooperative path planning that integrates aerial image segmentation to identify safe paths in dynamic, evolving settings.

Collectively, these studies emphasize the importance of autonomy and cooperation in SAR, particularly for rapid victim localization and safe navigation. The integration of UAV–UGV frameworks stands out as a promising direction, leveraging the speed, coverage, and complementary capabilities of aerial and ground systems [9].

III. Guided Long-Horizon Integrated Drone Escort

We propose a cooperative multi-agent framework for integrated aerial-ground search-and-rescue (SAR) operations. Prior to detailing the framework, we identify the core capabilities that must be established to ensure effective SAR:

1) Accurate detection and geolocation of victims.
2) Time-efficient path planning to reach confirmed goals.
3) Online situational awareness in unknown environments for long-horizon planning.

The first two address perception and path planning, respectively, and the third couples both: relying solely on a local navigation stack can cause the system to become trapped in non-convex environments (Figure 5), necessitating higher-level situational awareness to obtain near-optimal solutions.

We next present a modular framework that operationalizes the three functionalities outlined above. The system comprises a ground ego-vehicle and two complementary aerial subsystems (each instantiated by one or multiple robots) with integrated detection. Responsibilities are decoupled and interfaced via clearly defined information products, ensuring cohesive end-to-end behavior.

- **(Aerial) Goal-Searching UAV(s).** Deployed from the ego-vehicle, these UAVs survey high-priority regions to detect and geolocate victims.

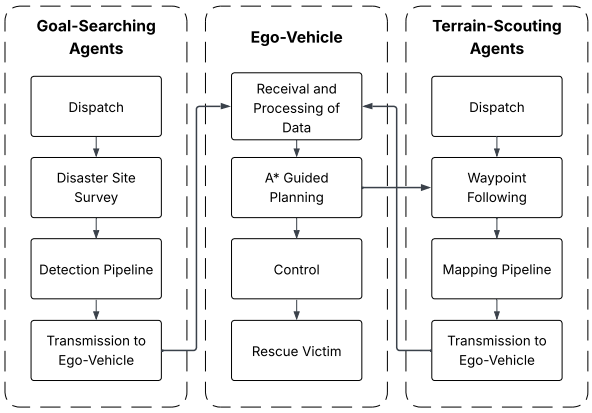

Fig. 2: The framework illustrates cooperative aerial-ground SAR operations: goal-searching UAVs survey the disaster site, detect and geolocate victims, and transmit compact messages to the UGV, while terrain-scouting UAVs track the UGV's path to identify obstacles. The UGV fuses these inputs with local sensing and employs an A* planner to compute safe, time-optimal routes to the victims.

- **(Aerial) Terrain-Scouting UAV(s).** Operating near the ego-vehicle, these units provide situational awareness by identifying non-traversable terrain/regions to provide a mid-level map.
- **(Ground) Ego-Vehicle.** The ground ego-vehicle navigates toward detected victims, leveraging a mid-level map for long-horizon planning.

Aerial platforms offer agility and mobility that complement the ground robots, while the ego-vehicle can be application-dependent: in our implementation (Section IV), we employ a golf cart, but the framework is equally applicable to quadrupeds or even humanoids. The choice primarily affects the definition of traversability and associated constraints, as locomotion capabilities vary across platforms.

A. Perception

Perception underpins the proposed framework, and we scope it into two sub-problems that provide the planning inputs: (i) a georeferenced set of victim coordinates, and (ii) a mid-level traversability map for situational awareness.

Victim Detection. Similar to the idea from [17], goal-searching UAVs perform victim detection from aerial imagery, identifying if the current frame contains target victims and generating bounding boxes for pixel-space localization. Because these agents often operate at standoff distances from the ego-vehicle over bandwidth-limited links, streaming raw images is impractical. Each agent thus executes lightweight, real-time edge inference, transmitting only compact detection messages to the ego-vehicle.

Traversability inference. Situational awareness is derived from aerial imagery captured by terrain-scouting UAVs, which is used to infer regions of non-traversable or hazardous terrain (e.g., water, rubble, voids). Since these agents operate in closer proximity to the ego-vehicle, transmitting compressed imagery is feasible when desired. Consequently, the ego-vehicle can perform the more computationally intensive tasks of segmentation and georegistration, while the aerial units function primarily as image capture platforms.

Since detection uses aerial images, all modules must be resilient to variation in altitude, viewpoint, motion blur, and illumination. The victim detection stack must satisfy real-time constraints on resource-limited aerial platforms, whereas traversability inference may employ more compute-intensive methods executed on the ego-vehicle.

B. Time-Optimal Path Planning for Ego-Vehicle

Given georeferenced victim coordinates and a mid-level traversability map, the ego-vehicle performs time-optimal path planning using graph-based traversal algorithms. To construct the planning graph, we discretize the local workspace into a uniform grid centered on the ego-vehicle, using a predefined grid resolution parameter.

Non-traversable regions from the map are modeled as hard obstacles, while we insert victim coordinates as goal nodes. If a victim lies outside the grid, it is projected along the corresponding direction from the ego-vehicle onto the grid boundary and treated as a proxy goal.

We formulate ego-vehicle path planning as the computation of a shortest-distance trajectory on the discretized roadmap, employing the graph-based search algorithm A^*. When multiple victims are present, the planner constructs a pairwise shortest-time matrix over the roadmap and determines a visitation order that minimizes total travel time. The plan is continuously adapted, with replanning triggered by map updates or the detection of additional victims.

C. Path Planning for Terrain-Scouting UAVs

Terriain-scouting UAVs are assigned to follow the ego-vehicle's planned trajectory, but several steps ahead. If the current plan has not yet reached all victims, a straight-line extension to the next victim in the visitation order serves as the reference. As unexplored regions are mapped, we continuously update the ego-vehicle's path planning, which in turn guides the terrain-scouting UAVs toward the ego-vehicle's anticipated future positions.

In this project, the main objective of the terrain-scouting UAVs is to validate the executability of the current planned trajectory of the ego-vehicle. A natural extension is to incorporate advanced exploration strategies to proactively identify alternative detours that may provide faster access to victims, which we leave as future work.

D. Overall Framework

A framework overview is shown in Figure 3. The operational sequence proceeds as follows. The system begins by deploying goal-searching UAVs for victim detection. Because time-optimal search planning is highly scenario-dependent and beyond the scope of this work, we assume that the planning paths of goal-searching UAVs are provided externally (e.g., user inputs).

Once victims are identified, the remainder of the system operates autonomously. Terrain-scouting UAVs are first dispatched ahead of the ego-vehicle in time to initialize situational awareness. The ego-vehicle then computes a distance-optimal trajectory to the detected victims using the progressively refined mid-level map, which encodes identified non-traversable regions. If the ego-vehicle becomes immobilized at any stage, a signal is transmitted to trigger recovery procedures.

IV. Hardware Implementation

We validate the framework with a proof-of-concept hardware demonstration. In this section, we discuss the principal engineering choices encountered during system integration.

A. Hardware Platform

Ego-vehicle. The ground platform is a GEM e6 golf cart with actuator limits: maximum speed $2\ \mathrm{m/s}$, steering rate $40°/\mathrm{s}$, and steering acceleration $10°/\mathrm{s}^2$. Drive-by-wire control is provided by PACMOD (accelerator, brake, steer, turn signals). Localization uses GPS-RTK with NTRIP corrections. Onboard compute comprises an Intel I9-14900K, 128GB RAM, dual RTX 4080 GPUs, and 4TB storage. The system is currently much more powerful than needed for GLIDE to function, allowing plenty of room for future expansion on the framework. A VLP32 LiDAR produces a local occupancy grid for emergency braking. Trajectories from the time-optimal planner are executed by a rate-limited PID tracking controller.

Goal-Searching and Terrain Scouting UAVs The UAVs are Holybro X500 platforms (PX4 firmware) with maximum speed $5\ \mathrm{m/s}$, ascent/descent rate $1\ \mathrm{m/s}$, tilt limit $30°$, and thrust-to-weight ratio 2. An outer-loop PID controller converts waypoint references to velocity commands stabilized by the onboard attitude controller. Each airframe is modified with a 3D-printed mounts for a Jetson Orin Nano, a DC-DC converter, and a nadir-facing OakD-Pro-W RGBD camera.

B. Perception

We decouple victim and non-traversability detection into two steps: objects are first identified in image-space, and then mapped to the pre-defined semantic labels (e.g., people $\rightarrow$ victims; vehicles/walls $\rightarrow$ non-traversable).

However, most commercial image-based object identification models are not directly applicable, as they are typically trained on principal viewpoints rather than the top-down perspectives. To this end, we fine-tune the YOLOv11 [19] model on top-down aerial imagery from the VisDrone dataset [18], with label remapping to retain only the relevant classes. The fine-tuned model is further optimized with TensorRT and CUDA to achieve a inference rate of 30 HZ on Jetson. This lightweight computational cost enables real-time edge inference onboard the UAVs. In the hardware demonstrations, both the searching and scouting UAVs employ this YOLO-based detection stack.

C. Communication

A TP-Link Archer AX21 (AX1800) router mounted on the golf cart provides the network backhaul. Since the internet throughput decreases with range, only compact georeferenced messages (< 1 kB/event) are transmitted between the UAVs and the ego-vehicle.

If connectivity is lost, detections are queued onboard and forwarded upon reconnection. This design ensures that, even during long-range missions, the UAVs continue collecting information and, once they return to a region with reliable connectivity, transmit the accumulated data to the ego-vehicle.

D. Geolocation

Detections from the nadir-facing camera on the UAVs are projected to the ground plane using camera intrinsics together with GPS and altitude measurements. Each object detection, represented by a bounding box from the fine-tuned YOLO model, is approximated as a ground contact in pixel space. The drone quaternion, which encodes the camera heading, is then used to enable a closed-form projection from image coordinates to georeferenced world coordinates.

Accurate timestamp synchronization across image, GPS, and altitude is required. Under stable attitudes, meter-level accuracy is typical. GPS dominates the error budget, while RTK corrections yield a sub-meter localization.

E. Planning

Goal-searching UAVs provide two planning options: (i) autonomous waypoint surveys uploaded via QGroundControl with a site geofence, and (ii) manual pilot supervision for scenario-dependent exploration. The perception stack (utilizing Docker Compose and pymavlink) bridges the flight controller and Jetson, launches sensor drivers, runs the ROS2 detection node, georeferences detections, and transmits/caches results.

The ego-vehicle incrementally updates the roadmap to incorporate georeferenced victim detections and free-space waypoints, while treating non-traversable regions as hard obstacles. LiDAR returns are fused during graph construction to enforce local safety. The planning graphs are persistent throughout the mission, enabling the system to accumulate situational knowledge over time rather than discarding previously constructed structures.

V. Hardware Experiments

We validate the framework in outdoor trials aimed at searching for victims and driving the UGV to their georeferenced locations while minimizing travel time and ensuring collision-free navigation. A system overview of the framework discussed is shown in Figure 3.

Scenarios. Owing to limited hardware resources (primarily objects to construct non-traversable areas), we restrict our evaluation to two scenarios: (i) *U-shaped corridor* formed by parked vehicles (opening facing North) and (ii) *Linear barrier* (East-West), both with the UGV starting North and the victim South (Figure 1).

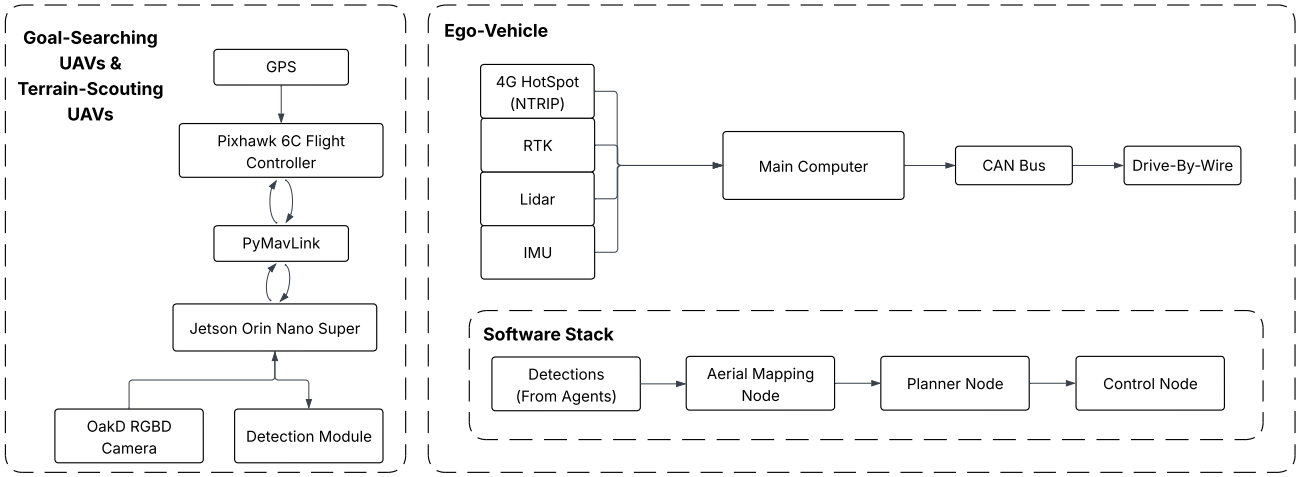

Fig. 3: Both UAVs run lightweight YOLOv11 detection pipelines, finetuned on the VisDrone dataset [18] and optimized with TensorRT, onboard the Jetson Orin Nano. The results are paired with the state information (retrieved from the flight controller via PyMavLink) of the UAV in order to georeference the victim positions. On the ground side, detections are ingested by the aerial mapping node, which fuses them with sensor inputs (LiDAR, RTK) to update a mid-level map. A planner node then computes safe trajectories using this map, while the control node sends low-level commands via the CAN bus to the drive-by-wire system. Together, these modules enable a coordinated, yet modular approach to achieving the objectives of a SAR-style mission scenario.

Fig. 4: (Left) One of the HolyBro X500 platforms utilized in the experimental studies as either a goal-searching UAV or a terrain-scouting UAV. (Right) The GEM e6 golf cart outfitted as the ego-vehicle for the real-world hardware demonstrations.

Settings. We compare three planning strategies where the system plans with: (i) a ground-truth map (*GT*), (ii) only onboard LiDAR with a 10m×10m local window (*Local*), and (iii) the traversability map supplied by the terrain-scouting UAVs that fly ahead of the current UGV plan with a 15m lead offset (*GLIDE*). In all trials, the goal-searching UAV first localizes the victim and transmits the georeferenced detection. We apply A^* on a 300×300 grid (0.5m resolution in ENU coordinates), using a distance-plus-orientation heuristic that favors states both spatially close to the goal and properly aligned in heading:

$$h(s, g) = \|(x, y) - g\|_2 + \lambda \cdot \Delta\theta/\pi,$$

where $s = (x, y, \Delta\theta)$ is the state, g is the goal, (x, y) is the position, $\Delta\theta$ is the minimal heading difference, and λ is the orientation-penalty coefficient. For hardware experiments, we set $\lambda = 1$. To mitigate georegistration errors, we discard detections when UAV roll or pitch exceeds $15°$. We fly at 15m altitude, admit detections only from the image center, and require a 3-hit spatial consensus within 1m before insertion into the map. Ten runs are executed per scenario and setting with start poses randomized by $\pm$5m. Trials terminate upon goal reach or safety-driver intervention.

Scenario	Setting	Duration (sec)	Mean Distance (m)	Success Rate (%)
U-Shaped	*GT*	43.9	102.40	100
	Local	51.27	117.8	20
	GLIDE	47.17	106.20	100
Line	*GT*	44.70	87.59	100
	Local	56.70	146.45	60
	GLIDE	50.29	108.73	100

TABLE I: Results of Hardware Demonstration

Results. Table I summarizes the averaged metrics. The *GLIDE* setting achieves 100% success in both scenarios and greatly surpasses *Local* in the success rate, highlighting the importance of long-horizon planning guided by situational awareness. Relative to *GT*, *GLIDE* remains close in the U-shaped case (+3.27s, +3.7% path) and exhibits a conservative detour in the Line case (+5.59s, +24.1% path). Such discrepancy is driven by whether the terrain scouts spent enough time in the area at stable orientations to update the map. Additionally, due to the U-shape structure, the planned path takes on a much larger change after discovering the whole obstacle compared to a simple line. The results reflect that further improvements require advances in UAV task planning to enhance traversability mapping. Overall, we demonstrate that the proposed framework enables search and rescue in unknown environments, using UAVs to guide UGV planning towards the victims.

Discussion. From our hardware experiments, we uncov-

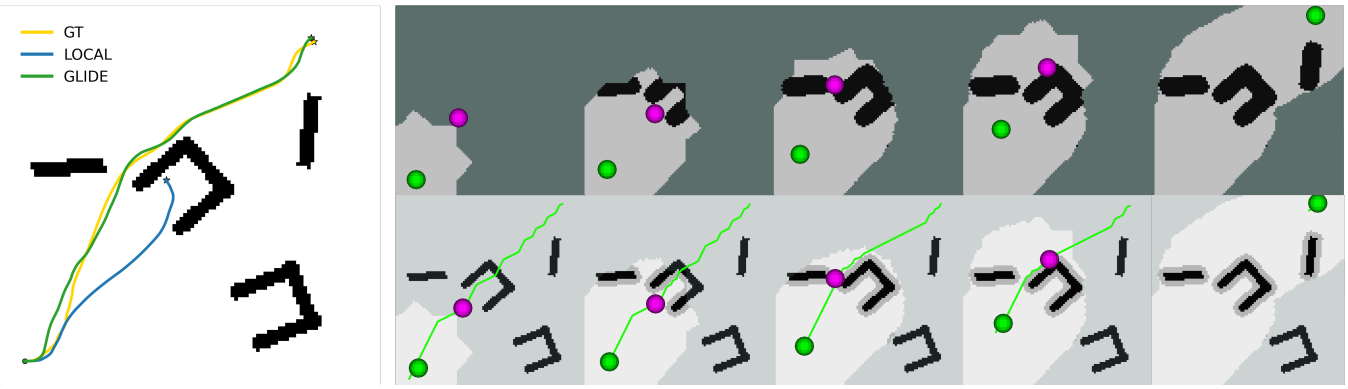

Fig. 5: Aerial-ground navigation in simulation. *Left*: Trajectories under *GT* (yellow), *Local* (blue), and *GLIDE* (green). *Right*: Sequential snapshots of one *GLIDE* trial, where the UGV (green) and UAV (purple) scouting progressively reveal obstacles (black) from the unexplored (dark grey), enabling replanning around them to reach the goal.

ered several key insights about our GLIDE framework. For example, the large performance gap between the *Local* and our *GLIDE* setting underscores the severe limitation of navigating under a purely reactive navigation strategy, especially when facing non-convex environments. Without the aerial guidance provided by GLIDE, the ego-vehicle becomes trapped in local minima because it is unable to perceive the global structure for effective path planning.

Within the U-shaped scenario, we note that *GLIDE* achieves excellent performance, suggesting that the terrain-scouting UAV captures the essential geometry of the obstacle. However, in the line-shaped scenario, there is a much larger performance gap between *GLIDE* and the *GT* setting. This was unexpected as we presumed that the simpler scenario should prove less challenging. We surmise that due to the less complex shape, the terrain-scouting UAVs spend less time hovering over the obstacles, resulting in sparser coverage and more conservative path-planning by the ego-vehicle.

Some practical considerations also emerged during our field trials. First, ensuring stable UAV orientations during perception tasks was critical for accurate georeferencing. Without a gimbal attachment, we improvised by enforcing a pitch/roll threshold when performing the perception tasks. Additionally, the 3-hit consensus methods helped to filter incorrect mappings which escaped the previous thresholding approach. Motion blur was not an issue in our case due to the use of a global shutter camera, though the filtering approaches mentioned above would help in that case as well. Additionally, wind gusts, lighting conditions, and the terrain texture itself occasionally degraded performance suggesting a direction for future work to improve upon.

VI. Simulation Experiments

We developed a PX4-Gazebo (Harmonic) simulation that mirrors the proposed framework while isolating the planning stack from perception to enable controlled ablations. The setup uses ROS2 Humble, a Clearpath Warthog UGV, and a PX4 X500 UAV. The simulation will be open-sourced along with the entire codebase.

Scenarios. Worlds are procedurally generated (SDF) to create U-shaped corridors and linear barriers with randomized sizes and orientations. We design two levels of difficulty characterized by different object densities. In Figure VI, an example of an easy scenario under exploration is shown.

Settings. We consider all three planning strategies presented in Section V. The simulation assumes perfect detection: a $50\text{m}\times50\text{m}$ area centered at the terrain-scouting UAV's footprint is treated as detected, and all non-traversable information within this area is incorporated into the ego-vehicle's map. The A^* planner is evaluated under two heuristic configurations, A^*_1 and $A^*_{2.5}$, which differ in the orientation-penalty coefficient λ (1.0 vs 2.5). Forty trials are run per scenario and setting.

Results. Across both difficulty levels, *GLIDE* consistently narrows the gap to *GT* and substantially outperforms *Local* (Table II), consistent with the hardware results. On easy worlds, *GLIDE* achieves 100% success rate for A^*_1/$A^*_{2.5}$ versus $70-75\%$ for *Local*, with small overheads relative to *GT* ($+1.2\text{s}/+7.7\text{m}$ for A^*_1; $+3.1\text{s}/+5.7\text{m}$ for $A^*_{2.5}$). On hard worlds, *GLIDE* improves success to $90-95\%$ (vs. $65-75\%$ for *Local*) with moderate overheads relative to *GT* ($+6.8\text{s}/+10.5\text{m}$ for A^*_1; $+2.2\text{s}/+10.4\text{m}$ for $A^*_{2.5}$). Increasing the A* heuristic coefficient λ from 1 to 2.5 biases A^* toward safer but more conservative trajectories: on hard worlds, the mean path length of *GLIDE* increases by 7.2m, while the success rate improves by 5%. Relative to Table I, the *GT*-*GLIDE* gap is comparable in several settings despite procedurally generated clutter, largely because simulation assumes perfect detection and georegistration. These results motivate improving the planning of terrain-scouting UAVs to proactively find shorter detours and further reduce time-to-rescue.

Discussion. The simulation study provides several insights that complement the hardware experiments. First, the gap between *GLIDE* and *GT* is consistently smaller in simulation than in reality. This is largely because the simulation assumes perfect perception: the UAV can instantaneously and without error integrate all obstacles within a fixed detection radius. In real-world deployments, however, perception is inevitably

Difficulty	Algorithm	Setting	Duration (sec)	Mean Distance (m)	Success Rate (%)
Easy	A_1^*	*GT*	50.08	66.89	100
		Local	59.25	92.10	70
		GLIDE	51.28	74.55	100
	$A_{2.5}^*$	*GT*	53.73	67.34	100
		Local	60.08	86.76	75
		GLIDE	56.80	72.99	100
Hard	A_1^*	*GT*	52.54	73.76	100
		Local	63.50	105.64	65
		GLIDE	59.37	84.23	90
	$A_{2.5}^*$	*GT*	57.60	81.00	95
		Local	64.98	108.91	75
		GLIDE	59.82	91.43	95

TABLE II: Results of Simulation Testing

limited; sensors provide only partial coverage, suffer from noise, and update maps with latency. As a result, some obstacles may be missed or inaccurately localized, leading to suboptimal planning and a wider gap between *GLIDE* and *GT*. These findings emphasize the need for perception-aware planning strategies that are robust to incomplete or uncertain environmental information.

VII. SUMMARY

In this paper, we propose and validate a cooperative aerial-ground search-and-rescue (SAR) framework that pairs two UAVs with a UGV to enable rapid victim localization, reduce time-to-rescue, and improve navigational safety. Our system assigns complementary aerial roles to the UAVs: goal searching and terrain scouting. The UAVs employ lightweight, fine-tuned YOLO-based detectors to identify and georeference victims as well as flag untraversable terrain, transmitting this information to the UGV to support planning. Our results show that explicitly decoupling aerial roles for detection and mapping, then closing the loop with UGV planning, yields reliable and fast mission completion in SAR scenarios. These results are verified in both simulation and real world.

Future work could include incorporating online search-planning algorithms that adaptively cover high-probability victim locations to further enhance system autonomy. Additionally, the fixed 15 m lead offset for terrain-scouting UAVs could be dynamically adjusted based on UGV speed and environmental complexity to better optimize the exploration-exploitation trade-off.

REFERENCES

[1] C. O. Quero and J. Martinez-Carranza, "Unmanned aerial systems in search and rescue: A global perspective on current challenges and future applications," *International Journal of Disaster Risk Reduction*, vol. 118, p. 105199, 2025.

[2] K. Chen, W. Zhao, Y. Liu, Y. Xia, W. Liang, and S. Wang, "Uav-ugv cooperative trajectory optimization and task allocation for medical rescue tasks in post-disaster environments," in *arXiv preprint arXiv:2506.06136*, 2025.

[3] M. Lyu, Y. Zhao, C. Huang, and H. Huang, "Unmanned aerial vehicles for search and rescue: A survey," *Remote Sensing*, vol. 15, no. 13, 2023.

[4] M. S. Mondal, S. Ramasamy, and P. Bhounsule, "Optiroute: A heuristic-assisted deep reinforcement learning framework for uav-ugv collaborative route planning," in *arXiv preprint 2309.09942*, 2023.

[5] K. Chen, W. Zhao, Y. Liu, Y. Xia, W. Liang, and S. Wang, "Uav-ugv cooperative trajectory optimization and task allocation for medical rescue tasks in post-disaster environments," in *Arxiv preprint 2506.06136*, 2025.

[6] S. Brotee, F. Kabir, M. A. Razzaque, P. Roy, M. Mamun-Or-Rashid, M. R. Hassan, and M. M. Hassan, "Optimizing uav-ugv coalition operations: A hybrid clustering and multi-agent reinforcement learning approach for path planning in obstructed environment," 2024.

[7] M. Petrlík, P. Petráček, V. Krátký, T. Musil, Y. Stasinchuk, M. Vrba, T. Báča, D. Heřt, M. Pecka, T. Svoboda, and M. Saska, "Uavs beneath the surface: Cooperative autonomy for subterranean search and rescue in darpa subt," *IEEE Transactions on Field Robotics*, vol. 2, p. 643–689, 2025.

[8] M. D. Petrillo, J. Beard, Y. Gu, and J. N. Gross, "Search planning of a uav/ugv team with localization uncertainty in a subterranean environment," in *ArXiv preprint 2102.06069*, 2021.

[9] C. Liu, J. Zhao, and N. Sun, "A review of collaborative air-ground robots research," *Journal of Intelligent & Robotic Systems*, vol. 106, no. 3, p. 60, 2022.

[10] F. S. Shahar, M. T. H. Sultan, M. Nowakowski, and A. Łukaszewicz, "Ugv-uav integration advancements for coordinated missions: A review," *Journal of Intelligent & Robotic Systems*, vol. 111, no. 2, p. 69, 2025.

[11] I. Munasinghe, A. Perera, and R. C. Deo, "A comprehensive review of uav-ugv collaboration: Advancements and challenges," *Journal of Sensor and Actuator Networks*, vol. 13, no. 6, p. 81, 2024.

[12] L. Wang, D. Cheng, F. Gao, F. Cai, J. Guo, M. Lin, and S. Shen, "A collaborative aerial-ground robotic system for fast exploration," in *Proceedings of the 2018 International Symposium on Experimental Robotics (ISER)*, pp. 59–71, Springer, 2020.

[13] A. Gawel, Y. Lin, T. Koutros, R. Siegwart, and C. Cadena, "Aerial-ground collaborative sensing: Third-person view for teleoperation," in *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 1–7, IEEE, 2018.

[14] D. Oh and J. Han, "Smart search system of autonomous flight UAVs for disaster rescue," *Sensors*, vol. 21, p. 6810, Oct. 2021.

[15] M. H. Zafar, S. K. R. Moosavi, and F. Sanfilippo, "Enhancing unmanned ground vehicle performance in SAR operations: Integrated gesture-control and deep learning framework for optimised victim detection," *Frontiers in Robotics and AI*, vol. 11, p. 1356345, Jun. 2024.

[16] M. K. Vasić, A. Drak, N. Bugarin, S. Kružić, J. Musić, C. Pomrehn, M. Schöbel, M. Johenneken, I. Stančić, V. Papić, and R. Herpers, "Deep semantic image segmentation for uav-ugv cooperative path planning: A car park use case," in *2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–6, 2020.

[17] P. Doherty and P. Rudol, "A uav search and rescue scenario with human body detection and geolocalization," in *AI 2007: Advances in Artificial Intelligence* (M. A. Orgun and J. Thornton, eds.), (Berlin, Heidelberg), pp. 1–13, Springer Berlin Heidelberg, 2007.

[18] P. Zhu, L. Wen, D. Du, X. Bian, H. Fan, Q. Hu, and H. Ling, "Detection and tracking meet drones challenge," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7380–7399, 2021.

[19] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics YOLO," Jan. 2023.