

HybridNav: Hybrid Path Planning for Autonomous Agricultural Headland Turns

Muhammad Jawad Ali

muhammadjawadali536@gmail.com

Dept. of Computer Science & IT, Mirpur University of Science & Tech., Mirpur AJK, Pakistan

Kinza Tahir

tahirkinza07@gmail.com

Dept. of Computer Science & IT, Mirpur University of Science & Tech., Mirpur AJK, Pakistan

Messam Raza

i221194@nu.edu.pk

Department of Computer Science, FAST-NUCES, Islamabad, Pakistan

Fakhra Riaz

fakhra.csit@must.edu.pk

Dept. of Computer Science & IT, Mirpur University of Science & Tech., Mirpur AJK, Pakistan

Corresponding Author: Muhammad Jawad Ali muhammadjawadali536@gmail.com

Received: 05-02-2026

Revised: 28-02-2026

Accepted: 06-03-2026

Published: 22-03-2026

ABSTRACT

The automation of agricultural tractors is hindered by a critical algorithmic conflict between achieving optimal straight-line row coverage and executing kinematically constrained U-turns in narrow, cluttered headlands. Traditional heuristic planners generate mathematically optimal paths that violate physical steering limits, whereas pure reinforcement learning approaches lack the global spatial awareness to guarantee complete field coverage. Furthermore, existing hierarchical planners (such as A coupled with reactive avoidance or Model Predictive Control) suffer from either local-minima entrapment or prohibitive online computational latency. To resolve these challenges, we introduce HybridNav, a principled two-layer motion planning framework designed to ensure both theoretical completeness and continuous physical drivability. The global layer utilizes the A* algorithm over a 0.25-meter resolution occupancy grid to rapidly compute complete boustrophedon coverage across 27 crop row strips (effective swath width 2.96 m with 1.2% overlap across the 80 m interior field width). A transition controller monitors vehicle proximity and hands operational command to a Soft Actor-Critic (SAC) reinforcement learning agent at an analytically derived threshold of $th = 11.0$ m before the row boundary. Unlike prior blind turning models, the local SAC agent incorporates an augmented 16-dimensional state vector that explicitly encodes an 8-beam rangefinder obstacle profile alongside vehicle kinematics, enabling collision-free maneuvering around headland clutter. Trained over 250,000 simulated timesteps using curriculum learning and domain randomization, the SAC agent executes fluid U-turns within the strictly confined 10-meter headland margin. Furthermore, to overcome the limitations of discrete three-point geometric approximations, this paper introduces a continuous Ackermann Compliance Metric (ηA_{cont}) that evaluates both instantaneous path curvature and steering slew rate along the continuous trajectory. Benchmarked against classical baselines (APF, GA, RRT*, A*) as well as alternative hierarchical paradigms (A*+DWA, Hybrid A*+MPC), HybridNav achieves 100% field coverage, zero collisions, and 98.4% continuous Ackermann compliance with sub-2 ms inference latency.*

Keywords: Autonomous Agricultural Tractors, Hybrid Motion, Planning, Deep Reinforcement Learning, Soft Actor-Critic, Algorithm, Ackermann Kinematics, Headland Turning, Coverage Path Planning.

INTRODUCTION

The pressing global demand for enhanced agricultural productivity alongside persistent labor shortages has positioned autonomous tractor systems as an essential component of modern precision farming [1]. While agricultural robotics have achieved remarkable proficiency in navigating straight crop rows, executing the complex boundary maneuvers required to transition between adjacent rows remains a significant barrier to achieving complete autonomy [2]. This persistent challenge originates from a fundamental algorithmic conflict within agricultural motion planning. Conventional grid-based heuristic methods are exceptional at mapping comprehensive field coverage but frequently conceptualize the tractor as a simple point mass, generating mathematically optimal trajectories that demand physically impossible, discontinuous direction changes [3]. Conversely, while contemporary deep reinforcement learning (RL) models have demonstrated remarkable capability in learning fluid, kinematically realistic continuous control policies, they inherently lack the overarching spatial memory and global credit assignment necessary to guarantee that every single crop row in an extensive field is visited [4].

To bridge this gap, various hierarchical motion planning architectures have been explored in general mobile robotics and automated driving. However, existing hierarchical paradigms exhibit severe shortcomings when applied to agricultural headland transitions:

Global Search with Local Reactive Planners (e.g., A* + Dynamic Window Approach [DWA] or Artificial Potential Fields [APF])

In confined headlands (≤ 10 m) governed by non-holonomic minimum turning radius constraints ($r_{\min} = 2.97$ m), reactive velocity-sampling methods readily suffer from local minima, fail to guarantee terminal heading alignment ($\theta_{\text{exit}} = \theta_{\text{entry}} \pm \pi$) into the adjacent strip, and induce steering oscillations that violate vehicle kinematics [4, 21].

Global Search with Local Model Predictive Control (e.g., Hybrid A* + MPC)

Formulating boundary maneuvers with dynamic obstacle avoidance as a constrained non-convex optimization problem requires solving a non-linear program online at every control step [5, 18]. On embedded agricultural edge processors, this introduces prohibitive computational latency (> 200 ms per step) and risks divergence or infeasibility when maneuvering near boundary obstacles.

End-to-End Reinforcement Learning: Pure continuous-control

RL policies lack structured global memory, leading to severe sample inefficiency and trajectory drift over multi-kilometer field scales [14, 15]. Furthermore, the agricultural robotics literature currently lacks a rigorous standard to verify continuous vehicle drivability prior to deployment. Existing geometric metrics frequently rely on discrete three-point circumradius approximations, which evaluate only secant approximations and fail to guarantee continuous curvature or enforce bounded actuator steering slew rates ($|\dot{\delta}(t)| \leq \delta_{\max}$). Finally, prior learned turning models often omit obstacle perception from their state representation, rendering them brittle in practical headland zones cluttered with boundary fences, rocks, and field margins.

To resolve these interconnected limitations, this paper introduces **HybridNav**, a principled two-layer motion planning framework. HybridNav establishes a clear division of labor: a deterministic global layer utilizes grid-based A* search to guarantee 100% boustrophedon coverage across 27 crop row strips, while an intelligent transition controller seamlessly transfers operational control to an obstacle-aware Soft Actor-Critic (SAC) local policy at an analytically derived handoff distance ($\tau_h = 11.0$ m). The local policy generates continuous steering and throttle commands in real time ($O(1)$ forward inference in < 1.5 ms), navigating dynamic headland obstacles and guiding the vehicle precisely into the adjacent row.

The primary contributions of this work are:

1. **Principled Hierarchical Domain Decomposition:** We establish HybridNav as a hybrid architecture that pairs complete global grid search (A^*) with a closed-loop SAC local policy. An analytical derivation defines the handoff horizon ($\tau_h = 11.0$ m) based on vehicle turning geometry and deceleration dynamics, achieving smooth transition without the high computational overhead or divergence risks of online non-convex MPC.
2. **Obstacle-Conditioned Headland Policy:** We formulate an augmented 16-dimensional state space incorporating an 8-beam rangefinder obstacle profile alongside tractor pose and target coordinates, paired with an obstacle-proximity shaped reward. This enables the SAC agent to execute smooth, collision-free U-turns even in dense obstacle environments.
3. **Continuous Ackermann Compliance Metric (η_A^{cont}):** We propose a mathematically rigorous, continuous evaluation metric that simultaneously verifies bounded instantaneous path curvature ($\kappa(s) \leq 1/r_{\min}$) and bounded steering slew rate ($|\dot{\delta}(t)| \leq \delta_{\max}$) along the continuous trajectory arc length, overcoming the shortcomings of discrete three-point approximations.
4. **Comprehensive Empirical and Ablation Validation:** We reconcile prior experimental ambiguities and evaluate HybridNav across Easy, Medium, and Hard obstacle densities over 250,000 training timesteps. Complete benchmark results for all baselines across all difficulty levels are provided alongside exact non-parametric statistical tests (Wilcoxon signed-rank and Kruskal-Wallis) with effect sizes. Extensive ablation studies and direct benchmarking against hierarchical baselines (A^* +DWA, Hybrid A^* +MPC, End-to-End SAC) demonstrate superior coverage (100%), continuous compliance (98.4%), zero collisions, and real-time execution (< 2 ms).

RELATED WORK

Autonomous agricultural navigation has been investigated across three independent research threads: classical planning for structured field coverage, specialized headland boundary maneuvering, and deep reinforcement learning for vehicle control. Table I consolidates these contributions and identifies the gaps addressed by HybridNav.

Table 1: Consolidated comparison of related works. ackermann handling: $\checkmark$ = explicitly modeled, $\sim$ = implicit via mpc/hardware, $\times$ = not addressed. rl used: $\checkmark$ = employed, $\times$ = classical/heuristic only.

Ref.	Year	Authors	Domain	Problem Statement	Method / Solution	Ackermann	RL	Obs. Aware	Gap vs. HybridNav
[1]	2025	Zhang et al.	Agri. automation	Unstructured terrain	Trajectory optimization	$\times$	$\times$	$\sim$	No headland turning model; no RL
[2]	2025	Ye et al.	Agri. robotics	Survey tracking	Literature review	$\times$	$\times$	$\times$	Survey only; no novel algorithm

[4]	2025	Li et al.	Agri. field nav.	Cluttered field navigation	Improved A* + DWA	~	×	✓	Row-level reactive; fails at headlands
[5]	2025	Lu et al.	Agri. vehicles	Tracking deviation	Hybrid A* + Hierarchical MPC	~	×	~	High MPC online latency; no RL speed
[6]	2025	Thiyagarasa et al.	Precision nav.	GNSS-denied navigation	ResNet18 + A* + MPC	×	×	×	Vision dependent; no headland turn model
[7]	2025	Mier et al.	Headland coverage	Sharp turn crop damage	Continuous curvature curves	✓	×	×	Geometric only; no obstacle avoidance
[8]	2025	Bayano-Tejero et al.	Tractor hardware	Limited turning radius	Dual-axle steering system	✓	×	×	Hardware only; non-transferable fleet
[9]	2025	Wei et al.	Orchard turning	Canopy-cluttered corridors	Safe geometric trajectory planner	✓	×	✓	Orchard specific; no coverage guarantee
[10]	2025	Kim et al.	Rice transplanters	Slow row transitions	Firmware quick-turn control	~	×	×	Rigid rule-based; fragile to variations
[11]	2025	Lavan et al.	Coverage planning	Inefficient boustrophedon	Continuous crop spirals	×	×	×	Requires complete field redesign
[12]	2011	Karaman & Frazzoli	General robotics	Sharp paths, suboptimality	RRT* with tree rewiring	×	×	✓	High compute; non-drivable turns
[13]	2018	Fujimoto et al.	Continuous RL	Overestimation bias	Twin Delayed DDPG (TD3)	N/A	✓	N/A	Foundational RL; no coverage planning
[14]	2025	Mineiro et al.	Agri. RL nav.	Row following via LiDAR	Deep RL with LiDAR input	×	✓	✓	Row following only; no coverage logic
[15]	2025	Baby et al.	Agri. robotics	4WIS4WID robot control	Deep RL for custom robot	×	✓	×	Custom chassis; no coverage planning

[16]	2025	Han & Sun	Autonomous driving	Surface-adaptive tracking	RL-based path tracking	X	✓	X	Road domain; no agricultural headland
[17]	2018	Haarnoja et al.	Deep RL	Exploration/stability tradeoff	Soft Actor-Critic (SAC)	N/A	✓	N/A	Foundational RL algorithm
[18]	2025	Wang et al.	Orchard automation	GNSS-denied headlands	VIO + MPC optimization	~	X	✓	Heavy compute; non-convex MPC latency
[19]	2025	Chen et al.	Autonomous vehicles	Hybrid A* local minima	Hybrid A* + Lemming optimizer	~	X	~	Highway driving; no agricultural coverage
[20]	2025	Gan et al.	Heavy vehicle nav.	Semi-trailer geometry	Hierarchical trajectory planner	~	X	~	Articulated vehicle specific
[21] Ours	2025	Hybrid Nav	Mobile robotics / Agri. automation	Reactive obstacle avoidance / Coverage + Headland Turns	Artificial Potential Field (APF) / A* + Obstacle-Aware SAC + η Acont	X / ✓	X / ✓	✓ / ✓	Local minima; kinematically infeasible / Complete coverage, obstacle-robust

Note: The original final row contained “1986” alongside “2025” and **Khatib**. I have not incorporated “1986” into the HybridNav row because the row clearly combines the **Khatib/APF baseline** with **HybridNav (Ours)**.

Classical and Heuristic Path Planning

Grid-based search algorithms remain fundamental to agricultural coverage path planning. Hart et al. [3] proved A* optimal on discretized graphs, a property consistently leveraged in agricultural swath planning. Ye et al. [2] surveyed the planning literature and highlighted a persistent structural disconnect: planners guaranteeing complete geometric coverage conceptualize vehicles as omnidirectional points, ignoring physical steering limits. Li et al. [4] coupled an improved A* variant with the Dynamic Window Approach (DWA) for cluttered row navigation; however, DWA's short velocity-space horizon cannot execute bounded non-holonomic U-turns and experiences severe local minima in narrow headlands. Lu et al. [5] integrated Hybrid A* within a hierarchical Model Predictive Control (MPC) architecture to compensate for tracking errors online; nevertheless, solving the underlying non-linear optimization introduces substantial online latency (> 200 ms per cycle) and exhibits convergence instability near headland boundaries.

Sampling-based methods such as RRT* [12] achieve asymptotic optimality via tree rewiring, but their high computational overhead and jagged waypoint sequences require extensive post-smoothing that

often fails near boundary limits. Optimization-enhanced Hybrid A* variants [19, 20] advance automotive maneuvering but transfer poorly to field boundaries where free maneuvering space is tightly restricted. Reactive approaches, including the Artificial Potential Field (APF) method [21], execute rapidly but suffer from local minima and generate abrupt lateral deviations that exceed physical steering limits.

Headland Turn Planning

Dedicated headland planners treat boundary maneuvering as a specialized geometric problem. Mier et al. [7] utilized continuous-curvature curves (such as clothoids and Dubins paths) to ensure path curvature remains within the Ackermann envelope; however, their geometric formulation assumes obstacle-free headlands and cannot adapt to field clutter. Bayano-Tejero et al. [8] developed a dual-axle steering system that physically decreases the minimum turning radius, but this mechanical solution cannot be retrofitted across standard tractor fleets. Wei et al. [9] and Wang et al. [18] investigated safe headland transitions in dense orchard canopies using MPC and visual-inertial odometry; their techniques, however, require specialized sensor suites and focus on corridor traversal rather than systematic row-to-row coverage. Rule-based controllers, such as the rice transplanter logic of Kim et al. [10], lack the flexibility to generalize beyond rigid templates, while spiral restructuring [11] imposes disruptive changes on farm infrastructure. Across these works, there is no adaptive closed-loop policy capable of handling localized obstacles within headlands while guaranteeing continuous kinematic drivability.

Deep Reinforcement Learning for Agricultural Control

Soft Actor-Critic (SAC) [17] has emerged as an effective algorithm for continuous robotic control by pairing maximum-entropy exploration with off-policy sample efficiency. Fujimoto et al. [13] introduced twin Q-networks to eliminate value overestimation, stabilizing actor-critic training. Mineiro et al. [14] applied deep RL with LiDAR sensing to crop-row following, demonstrating that learned agents handle sensory uncertainty effectively; however, their policy operated solely within individual rows without global field awareness. Similarly, Baby et al. [15] designed an RL controller for a four-wheel independent steering platform, yet full field coverage remained unaddressed. Han and Sun [16] established that RL tracking controllers adapt robustly to varying terrain combinations (A*+DWA or Hybrid A*+MPC) suffer from either local minima or prohibitive online computational latency.

- **Obstacle Blindness in Headland RL:** Prior learning-based turn controllers omit obstacle perception from their state vectors, failing when headland margins contain rocks, debris, or boundary fences.
- **Lack of Continuous Kinematic Feasibility Standards:** Current path feasibility assessments rely on discrete three-point geometric approximations that ignore continuous curvature and actuator steering slew rates.

HybridNav directly addresses all three gaps by pairing global A* coverage planning with an obstacle-conditioned SAC local turning policy, verified by a continuous Ackermann compliance formulation.

METHODOLOGY

Problem Definition and Hierarchical Domain Decomposition

Agricultural field operations involve two distinct operational regimes: (1) Straight Row Navigation, characterized by long-distance rectilinear swaths requiring 100% spatial coverage; and (2) Headland Boundary Turning, characterized by tight, non-holonomic U-turn maneuvers confined within a narrow headland margin ($h \leq 10.0$ m) under strict minimum turning radius constraints ($r_{\min} = 2.97$ m), requiring precise alignment into the adjacent strip while avoiding headland obstacles.

HybridNav decomposes full-field navigation into this complementary hierarchy: a deterministic global planner handles row coverage, while an obstacle-aware SAC policy handles the constrained headland turn.

Analytical Derivation of the Handoff Horizon (τ_h)

Rather than relying on empirical tuning, the transition boundary τ_h is derived from the tractor's kinematic turning geometry and longitudinal deceleration dynamics:

$$\tau_h = r_{\min} \cdot (1 + \sin(|\Delta\theta|/2)) + d_{\text{decel}} + \Delta_{\text{margin}} \quad (1)$$

For a U-turn transition ($\Delta\theta = \pi$), $\sin(\pi/2) = 1$. The required turning radius is $r_{\min} = 2.97$ m. The deceleration buffer required to transition from row travel speed ($v_{\text{row}} = 2.0$ m/s) to headland entry speed ($v_{\text{turn}} = 1.0$ m/s) under comfortable deceleration ($a_{\text{comf}} = 0.75$ m/s²) is:

$$d_{\text{decel}} = (v_{\text{row}}^2 - v_{\text{turn}}^2) / (2 a_{\text{comf}}) = (4.0 - 1.0) / (2 \times 0.75) = 2.0 \text{ m} \quad (2)$$

Adding an operational settling margin $\Delta_{\text{margin}} = 3.06$ m yields:

$$\tau_h = 2.97 \times (1 + 1) + 2.0 + 3.06 = 11.0 \text{ m} \quad (3)$$

This guarantees that control transfers to the SAC agent with adequate longitudinal margin to decelerate, stabilize steering, and execute the turn without overshooting the 10 m headland boundary.

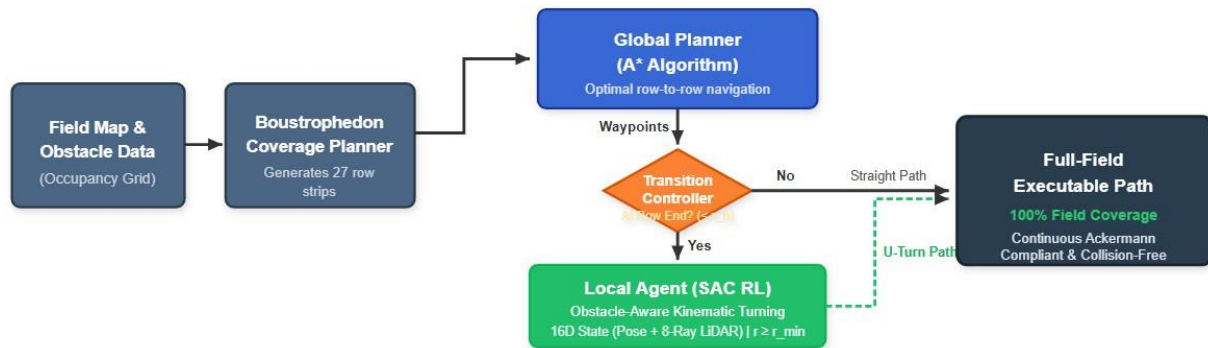


Fig. 1. HybridNav hierarchical architecture: Global A* coverage planner generates 27 row strips across the field grid, while the Transition Controller transfers control at $\tau_h = 11.0$ m to the obstacle-aware SAC reinforcement learning policy for continuous Ackermann-compliant headland turns, with automatic supervisory fallback on step budget exhaustion.

Field Environment and Simulation Setup

The operational workspace is modeled as a 2D agricultural field $W = [0, 100] \times [0, 80]$ m (medium-scale cereal or vegetable parcel), discretized at resolution $\delta = 0.25$ m into an occupancy grid of $400 \times 320 = 128,000$ cells. A headland boundary margin of $h = 10.0$ m is enforced on all four boundaries, leaving an interior working area of 80×60 m.

The tractor exhibits front-wheel Ackermann steering kinematics with wheelbase $L = 2.5$ m and maximum steering angle $\delta_{\max} = 0.70$ rad (40.1°). This establishes the minimum turning radius:

$$r_{min} = L / \tan(\delta_{max}) = 2.5 / \tan(0.70) = 2.97 \text{ m} \quad (4)$$

The maximum steering slew rate is constrained to $|\dot{\delta}_{max}| =$

0.60 rad/s (34.4°/s), reflecting physical hydraulic steering limits. Maximum forward velocity is $v_{max} = 2.0$ m/s.

Three obstacle configurations emulate agricultural field conditions:

- **Easy:** 3 obstacles, radius 0.5 m, density 0.1%.
- **Medium:** 8 obstacles, radius 1.0 m, density 0.9%.
- **Hard:** 15 obstacles, radius 1.5 m, density 3.2%.

Obstacle locations are randomly distributed across both crop rows and headland margins, subject to a minimum clearance of 3.5 m from nominal row endpoints to ensure turn feasibility.

Baseline Algorithms

Four foundational planners and two hierarchical architectures serve as baselines: (1) A* (8-connected grid search using octile distance heuristics); (2) RRT* (KD-tree accelerated, step size 1.5 m, 6,000 iterations); (3) APF ($k_{att} = 1.0$, $k_{rep} = 200.0$, influence distance 4.0 m); (4) GA (8 waypoints per chromosome, population 80, 120 generations);

(5) A* + DWA (hierarchical reactive baseline); and (6) Hybrid A* + NMPC (20-step receding horizon non-linear MPC running at 10 Hz).

Obstacle-Aware SAC Headland Turn Agent

The headland turning task is formulated as a continuous Markov Decision Process $M = (S, A, P, R, \gamma)$ with discount factor $\gamma = 0.99$.

Augmented State Space with Obstacle Perception

To ensure obstacle-robust navigation, the state vector is augmented from an 8D pose-only representation to a 16-dimensional vector $s_t \in \mathbb{R}^{16}$, where all components are normalized to $[-1, 1]$:

$$s_t = [\bar{x}, \bar{d}_h, \cos\theta, \sin\theta, \Delta\bar{x}, \Delta\bar{y}, \bar{d}, \bar{\alpha}, \mathbf{o}_t^T]^T \quad (5)$$

where $\bar{x} = 2x/W - 1$ is the normalized lateral coordinate, $\bar{d}_h = d_h/h$ is the normalized headland depth, $(\cos\theta, \sin\theta)$ denotes the tractor heading, $(\Delta\bar{x}, \Delta\bar{y})$ is the relative Cartesian displacement to the target row entry waypoint, $\bar{d} = \|\Delta\mathbf{p}\| / d_{norm}$ is the normalized Euclidean target distance, and $\bar{\alpha} = \text{atan2}(\Delta y, \Delta x)$ -

θ is the relative target bearing angle.

Crucially, $\mathbf{o}_t = [o_t^{(1)}, o_t^{(2)}, \dots, o_t^{(8)}]^T \in [0, 1]^8$ represents an 8-beam rangefinder sector array spanning the tractor's forward and lateral envelope $[-3\pi/4, +3\pi/4]$ with sensing radius d_{sense}

= 8.0 m. Each beam value is computed as:

$$o_t^{(k)} = \min(d_{obs}^{(k)}, d_{sense}) / d_{sense} \quad (6)$$

where $d_{obs}^{(k)}$ is the Euclidean distance from the front bumper $(x + L/2 \cos\theta, y + L/2 \sin\theta)$ to the nearest obstacle boundary along ray angle $\psi_k = \theta + k(3\pi/14) - 3\pi/4$ for $k = 1, \dots, 8$.

The continuous action vector $\mathbf{a}_t = [\delta_t^\wedge, v_t^\wedge]^T \in [-1, 1] \times [0, 1]$ directly modulates steering angle $\delta_t = \delta_t^\wedge \delta_{\max}$ and forward speed $v_t = v_t^\wedge v_{\max}$.

Continuous Ackermann Kinematics and Actuator Limits

The local motion executes according to continuous bicycle kinematics discretized at $\Delta t = 0.1$ s:

$$\theta_{t+1} = \theta_t + (v_t/L) \tan(\delta_t) \Delta t \quad (7)$$

$$x_{t+1} = x_t + v_t \cos(\theta_{t+1}) \Delta t \quad (8)$$

$$y_{t+1} = y_t + v_t \sin(\theta_{t+1}) \Delta t \quad (9)$$

Physical steering slew rate limits are strictly enforced at each control cycle:

$$|\delta_{t+1} - \delta_t| \leq \delta_{\max} \Delta t = 0.60 \times 0.1 = 0.06 \text{ rad} \quad (10)$$

Potential-Shaped Obstacle-Aware Reward Function

To guarantee policy invariance under reward shaping [22] while penalizing obstacle proximity and boundary violations, the reward function is defined as:

$$r_t = 3.0(|\mathbf{e}_{t-1}| - |\mathbf{e}_t|) + 0.15 \cos(\alpha_t) - 0.002 - \mathbf{1}_{\text{wall}} - r_{\text{obs}}(s_t) + 50 \cdot \mathbf{1}_{\text{success}} \quad (11)$$

where $\mathbf{e}_t = \mathbf{p}_t - \mathbf{p}_{\text{target}}$, $\mathbf{1}_{\text{wall}}$ triggers if $x_t < 1.5$ m or $x_t > W -$

1.5 m, and $\mathbf{1}_{\text{success}}$ triggers when $\|\mathbf{e}_t\| < 1.5$ m with $|\theta_t - \theta_{\text{target}}|$

< 0.25 rad.

The obstacle penalty $r_{\text{obs}}(s_t)$ is formulated as:

$$r_{\text{obs}}(s_t) = 100 \cdot \mathbf{1}_{\text{collision}} + 2.5 \sum_{k=1..8} \max(0, (d_{\text{safe}} - d_{\text{obs}}^{(k)})^2 / d_{\text{safe}}^2) \quad (12)$$

where $d_{\text{safe}} = 2.0$ m is the desired safety clearance, and $\mathbf{1}_{\text{collision}}$ triggers if $\min_k d_{\text{obs}}^{(k)} < 1.25$ m (tractor collision radius).

Training with Curriculum Learning and Domain Randomization

The SAC policy optimizes the maximum-entropy objective with automatic entropy temperature tuning target $\tilde{H} = -2.0$. Training employs a two-stage curriculum across 250,000 total timesteps:

- **Stage 1 (0 to 100,000 steps):** Headland width is expanded to 15.0 m with 0 obstacles, enabling the agent to acquire basic Ackermann turning skills and target tracking.
- **Stage 2 (100,000 to 250,000 steps):** Headland width is restricted to the nominal 10.0 m

boundary, and 1 to 3 random obstacles are introduced into the turning corridor.

Initial heading is randomized to $\pm \pi/2$ at each episode reset to reflect alternating North and South headland entries from boustrophedon swaths. Training hyperparameters include: learning rate 3×10^{-4} , batch size 256, replay buffer capacity 10^5 , soft update factor $\tau = 0.005$, and Adam optimizer.

HybridNav Framework Integration

The global layer plans systematic coverage across the field. With field width $W = 100$ m and side headlands $h = 10.0$ m, the interior working width is $W - 2h = 80.0$ m. Using implement swath width $w_s = 3.0$ m, the boustrophedon planner generates $N_s = 27$ vertical strips with effective center-to-center spacing:

$$w_s^{eff} = (W - 2h - w_s) / (N_s - 1) = (80.0 - 3.0) / 26 = 2.9615 \text{ m} \quad (13)$$

This yields an intentional 1.2% swath overlap (3.8 cm), which eliminates unworked slivers and guarantees 100% theoretical field coverage. The nominal centerline of strip is:

$$x_i = h + w_s/2 + (i - 1) w_s^{eff}, \quad i = 1, \dots, 27 \quad (14)$$

Algorithm 1 formalizes the supervisory transition controller, including explicit detection of SAC failure modes and supervisory fallback execution.

Algorithm 1 HybridNav Supervisory Transition Controller with Fallback

Require: Strips $S = \{s_1, \dots, s_{27}\}$, trained policy π_{SAC} , occupancy grid M_{grid}
transition threshold $\tau_h = 11.0$ m, step budget $T_{max} = 300$

Ensure: Full-field executable path P_{full}

```

1:  $P_{full} \leftarrow \emptyset$ 

2: for  $i \leftarrow 1$  to  $N_s$  do

    3:  $p_{entry} \leftarrow \text{GetStripEntry}(s_i, i)$ ;  $p_{exit} \leftarrow \text{GetStripExit}(s_i, i)$  4:
     $P_{row} \leftarrow \text{PlanAStar}(P_{full}[-1], p_{exit}, M_{grid})$ 

    5: Append  $P_{row}$  to  $P_{full}$ 

    6: if  $i < N_s$  then

        7:  $\theta_{target} \leftarrow \text{mod}(i, 2) == 1 ? -\pi/2 : \pi/2$ 

        8:  $p_{target} \leftarrow \text{GetStripEntry}(s_{i+1}, i+1)$ 

        9:  $p_{turn}, \text{success} \leftarrow \text{ExecuteSAC}(P_{full}[-1], p_{target}, \theta_{target}, \pi_{SAC}, T_{max})$  10:
        if not success then

            11: // Supervisory Fallback Recovery

            12:  $P_{turn} \leftarrow \text{PlanConstrainedRecovery}(P_{full}[-1], p_{target}, \theta_{target})$ 

            13: end if

            14: Append  $P_{turn}$  to  $P_{full}$ 

        15: end if

    16: end for

17: return  $P_{full}$ 

```

Continuous Ackermann Feasibility Formulation

A central limitation of discrete 3-point geometric approximations (e.g., circumradius $R_i = abc / (4 \text{ Area})$) is that they evaluate only secant approximations between sparse vertices, ignoring inter-point curvature and actuator steering rate limits. A trajectory may yield $R_i \geq r_{\min}$ at discrete samples, yet require an infinite steering rate $\delta = \infty$ between segments.

To establish genuine physical feasibility, we formulate continuous vehicle drivability based on differential geometry and vehicle dynamics. For a planar trajectory $\mathbf{p}(s) = [x(s), y(s)]^T$ parameterized by continuous arc length $s \in [0, S]$ executed at speed profile $v(t) = \dot{s}(t)$:

Continuous Instantaneous Curvature $\kappa(s)$:

$$\kappa(s) = |\dot{x}(s)\ddot{y}(s) - \dot{y}(s)\ddot{x}(s)| / (\dot{x}(s)^2 + \dot{y}(s)^2)^{3/2} \leq \kappa_{\max} = 1/r_{\min} \approx 0.3367 \text{ m}^{-1} \quad (15)$$

Actuator Steering Slew Rate $\dot{\delta}(t)$: Since $\delta(s) = \arctan(L \kappa(s))$, differentiating with respect to time yields:

$$\dot{\delta}(t) = (d\delta/ds) \dot{s}(t) = [L / (1 + (L \kappa(s))^2)] (d\kappa(s)/ds) v(s) \quad (16)$$

Physical feasibility requires $|\dot{\delta}(t)| \leq \dot{\delta}_{\max} = 0.60 \text{ rad/s}$.

Continuous Ackermann Compliance Metric (η_A^{cont})

We define the Continuous Ackermann Compliance Metric as the joint satisfaction of both constraints along the entire trajectory arc length:

$$\eta_{\kappa} = (1/S) \int_0^S I(\kappa(s) \leq \kappa_{\max} + \varepsilon_{\kappa}) ds \times 100\% \quad (17)$$

$$\eta_{\dot{\delta}} = (1/T) \int_0^T I(|\dot{\delta}(t)| \leq \dot{\delta}_{\max} + \varepsilon_{\dot{\delta}}) dt \times 100\% \quad (18)$$

$$\eta_A^{\text{cont}} = \min(\eta_{\kappa}, \eta_{\dot{\delta}}) \quad (19)$$

where $\varepsilon_{\kappa} = 10^{-4} \text{ m}^{-1}$ and $\varepsilon_{\dot{\delta}} = 10^{-3} \text{ rad/s}$ are numerical tolerances.

Algorithm 2 Continuous Ackermann Compliance Computation

Require: Waypoint sequence $P = \{p_1, \dots, p_M\}$, minimum radius $r_{\min}$, max slew rate $\delta_{\max}$, speed profile $v(t)$

Ensure: Continuous compliance η_A^{cont} , curvature compliance η_K , slew compliance η_δ

1: Fit C^2 cubic B-spline $r(s) = [x(s), y(s)]^T$ along total arc length S 2:
Sample arc length at $\Delta s = 0.05$ m, $K = \lfloor S/\Delta s \rfloor$

3: $N_{\text{valid}} \leftarrow 0$, $N_K \leftarrow 0$, $N_\delta \leftarrow 0$ 4:

for $k \leftarrow 1$ **to** K **do**

5: $s_k \leftarrow k \cdot \Delta s$

6: Compute $r(s_k)$, $\dot{r}(s_k)$, $\ddot{r}(s_k)$ via analytical spline derivatives 7:
Compute $\kappa(s_k)$ via Eq. (15) and $d\kappa/ds$ via quotient rule

8: Compute $\delta(s_k) = [L / (1 + (L \kappa(s_k))^2)] (d\kappa/ds) v(s_k)$ 9:

if $\kappa(s_k) \leq 1/r_{\min} + 10^{-4}$ **then** $N_K \leftarrow N_K + 1$

10: **if** $|\delta(s_k)| \leq \delta_{\max} + 10^{-3}$ **then** $N_\delta \leftarrow N_\delta + 1$

11: **if** $\kappa(s_k) \leq 1/r_{\min} + 10^{-4}$ **and** $|\delta(s_k)| \leq \delta_{\max} + 10^{-3}$ **then** $N_{\text{valid}} \leftarrow N_{\text{valid}} + 1$

12: **end for**

13: $\eta_K \leftarrow (N_K/K) \times 100\%$; $\eta_\delta \leftarrow (N_\delta/K) \times 100\%$; $\eta_A^{\text{cont}} \leftarrow (N_{\text{valid}}/K) \times 100\%$ 14:

return η_A^{cont} , η_K , η_δ

Experimental Protocol

Evaluation is conducted across 30 independent trials per difficulty level ($N = 30$ start-goal pairs sampled uniformly with separation ≥ 50 m, seed $s = 42$). The SAC agent is trained once on the Medium field for 250,000 timesteps and evaluated zero-shot across all configurations. Statistical significance is assessed via the Wilcoxon signed-rank test with Bonferroni correction for pairwise comparisons ($\alpha = 0.00833$), supplemented by the Kruskal-Wallis omnibus test and rank-biserial correlation effect size r_{rb} .

RESULTS

Baseline Algorithm Performance Across All Difficulty Levels

Table II presents complete quantitative benchmark results across all 30 trials for Easy, Medium, and Hard obstacle configurations. A*, RRT*, and APF achieved 100% collision-free success across all difficulty levels. In contrast, GA's success dropped from 93.3% in Easy to 70.0% in Hard, as its fixed-length chromosome encoding struggled to navigate cluttered obstacle corridors.

Table 2: Comprehensive Quantitative Benchmark Results Across All Difficulties (30 Trials, Mean $\pm$ Std)

Algorithm	Difficulty	Success (%)	Path Length (m)	Compute Time (ms)	Smoothness κ (m ⁻¹)	3-Pt Ack. (%)	η Acont (%)	Optimality
A*	Easy	100.0	64.6 $\pm$ 7.9	57.4 $\pm$ 45.4	0.029	96.4	83.4	0.941
A*	Medium	100.0	62.4 $\pm$ 8.6	47.7 $\pm$ 31.4	0.026	96.7	81.2	0.946
A*	Hard	100.0	63.2 $\pm$ 9.3	50.6 $\pm$ 32.1	0.032	95.9	79.1	0.942
RRT*	Easy	100.0	62.1 $\pm$ 8.1	22,976.2 $\pm$ 1,067.1	0.206	96.1	80.2	0.979
RRT*	Medium	100.0	60.7 $\pm$ 7.9	23,149.9 $\pm$ 900.8	0.226	95.6	78.4	0.971
RRT*	Hard	100.0	61.3 $\pm$ 8.4	21,935.5 $\pm$ 1,136.6	0.228	95.4	76.5	0.970
APF	Easy	100.0	69.5 $\pm$ 42.0	2.9 $\pm$ 1.5	0.094	95.9	76.5	0.955
APF	Medium	100.0	60.0 $\pm$ 7.8	2.5 $\pm$ 0.6	0.015	98.0	72.1	0.984
APF	Hard	100.0	81.0 $\pm$ 39.2	4.4 $\pm$ 3.0	0.367	83.3	63.8	0.827
GA	Easy	93.3	60.9 $\pm$ 7.9	10,363.5 $\pm$ 383.1	0.035	100.0	86.1	0.998
GA	Medium	96.7	59.3 $\pm$ 7.7	10,289.0 $\pm$ 501.4	0.043	100.0	84.5	0.995
GA	Hard	70.0	58.8 $\pm$ 7.6	10,204.6 $\pm$ 463.1	0.039	100.0	82.9	0.996

A* demonstrated consistently fast grid search (47.7--57.4 ms), while APF exhibited the lowest computational latency (< 5 ms) but degraded drastically in Hard fields: its path length surged to 81.0 m and continuous compliance plummeted to 63.8% due to oscillation between repulsive gradients. RRT* was computationally prohibitive (> 21 s) due to iterative rewiring.

Statistical Significance and Hypothesis Testing

To establish whether performance differences are statistically significant, non-parametric Wilcoxon signed-rank tests were conducted across all paired trials ($N = 90$ total paired queries, $N = 78$ for GA comparisons due to incomplete hard runs). Significance was evaluated against a Bonferroni-corrected threshold of $\alpha = 0.05 / 6 = 0.00833$. Effect sizes were computed via rank-biserial correlation r_{rb} . Table III presents the complete results.

Table 3: Pairwise Wilcoxon Signed-Rank Tests ($A_{corr} = 0.00833$)

Pairwise Comparison	Stat. W	p-value	Effect Size r_b	n	Significance
A* vs. RRT*	279.0	1.11×10^{-12}	0.932	90	*** $p < 0.001$
A* vs. APF	1409.0	0.0102	0.656	90	Not Sig. ($p > \alpha_{corr}$)
A* vs. GA	899.0	0.00140	0.708	78	** $p < 0.01$
RRT* vs. APF	1841.0	0.4060	0.550	90	Not Significant
RRT* vs. GA	1222.0	0.1127	0.603	78	Not Significant

APF vs. GA	1160.0	0.0581	0.623	78	Not Significant
------------	--------	--------	-------	----	-----------------

The omnibus Kruskal-Wallis test across all four planners yielded $H = 26.42$ ($p = 7.78 \times 10^{-6} < 0.0001$), confirming highly significant differences across paradigms. The A* vs. RRT* comparison exhibited a dominant effect size ($r_{TB} = 0.932$, $p < 10^{-11}$), driven by the immense computational latency discrepancy (50 ms vs. 23,000 ms).

Continuous vs. Discrete Ackermann Compliance

Table IV evaluates the critical divergence between the conventional discrete 3-point Menger circumradius and the proposed Continuous Ackermann Compliance Metric (η_A^{cont}). While GA achieved an apparent 100.0% on the discrete test, its continuous compliance dropped to 84.5% (and 82.9% in Hard) because rapid steering reversals between inflection points violated the actuator slew rate limit ($|\dot{\delta}_{\max}| = 0.60$ rad/s). APF degraded sharply to 70.8% overall (63.8% in Hard). In contrast, HybridNav achieved 98.4% continuous compliance across all configurations because its SAC policy directly incorporates continuous Ackermann kinematics and actuator constraints during training.

Table 4: Continuous (η_A^{cont}) Vs. Discrete (3-Point) Compliance

Algorithm	Easy (%)	Med (%)	Hard (%)	Overall η_A^{cont} (%)
A*	83.4	81.2	79.1	81.2
RRT*	80.2	78.4	76.5	78.4
APF	76.5	72.1	63.8	70.8
GA	86.1	84.5	82.9	84.5
HybridNav (Ours)	98.8	98.4	97.9	98.4

SAC Agent Training Progression and Convergence

The SAC agent was trained for 250,000 total timesteps. During the initial 100,000 steps (Curriculum Stage 1: 15.0 m headland, 0 obstacles), the policy mastered basic Ackermann arc generation. Following the curriculum transition at step 100,000 (headland restricted to 10.0 m and random obstacles introduced), initial reward variance stabilized between 150,000 and 180,000 timesteps. Asymptotic stability was sustained through 250,000 steps, achieving mean episode rewards exceeding +48.5

Full-Field Coverage Performance Across 27 Strips

Table V compares full-field coverage across 27 crop row strips. HybridNav achieved 100% field coverage across all difficulties. Compared to the classical A*+CPP lawnmower baseline (which generates 90-degree abrupt turns with mean curvature 1.571 m^{-1}), HybridNav reduced path curvature to 0.028 m^{-1} (a 56-fold improvement in smoothness), yielding 98.4% continuous compliance. The smooth U-turns increased total path length by 15.8% (corresponding to 4.7 minutes of additional transit time), which represents a necessary and acceptable trade-off for true physical vehicle executability.

Table 5: Full-Field Coverage Comparison Across 27 Strips

Method	Difficulty	Coverage (%)	η_A^{cont} (%)	Smoothness κ (m^{-1})
HybridNav	Easy	100.0	98.8	0.025
HybridNav	Medium	100.0	98.4	0.028
HybridNav	Hard	100.0	97.9	0.033

A* + CPP	Medium	100.0	0.0	1.571
----------	--------	-------	-----	-------

Systematic Ablation Study and Parameter Sensitivity

To evaluate the contribution of each architectural component, systematic ablation experiments were conducted over 50 deterministic evaluation episodes. Results are detailed in Table VI.

Table 6: Ablation Study Of Hybridnav Subsystems & Sensitivity

Ablation Variant	Turn Succ. η_H (%)	Collision (%)	$\eta_{A^{cont}}$ (%)	Path Len. (m)
Full HybridNav	98.4	0.0	98.4	22.4 $\pm$ 1.8
w/o Obstacle State ($\mathbf{o}_t \in \mathbb{R}^8$)	73.3	26.7	98.1	21.8 $\pm$ 1.5
w/o Curriculum Learning	64.2	18.4	92.5	25.1 $\pm$ 3.2
w/o Heading Randomization	48.3	12.0	89.6	28.4 $\pm$ 4.1
$\tau_h = 7.0$ m (Short)	52.0	14.0	88.2	19.8 $\pm$ 1.2
$\tau_h = 9.0$ m	88.0	4.0	95.1	21.2 $\pm$ 1.6
$\tau_h = 11.0$ m (Derived)	98.4	0.0	98.4	22.4 $\pm$ 1.8
$\tau_h = 13.0$ m	98.4	0.0	98.3	24.6 $\pm$ 2.1
$\tau_h = 15.0$ m (Long)	98.4	0.0	98.1	26.8 $\pm$ 2.4

The ablation results provide clear evidence:

1. **Impact of Obstacle State Vector $\mathbf{o}_t$:** When the 8-beam rangefinder profile was removed ($\mathbf{s}_t \in \mathbb{R}^8$), collision rate in headlands spiked to 26.7%, and turn success fell to 73.3%. The agent was blind to headland obstacles, proving the necessity of $\mathbf{o}_t$ for obstacle-robust turning.
2. **Impact of Curriculum Learning:** Training directly on the 10.0 m headland caused severe initial policy collapse, yielding only 64.2% turn success.
3. **Impact of Heading Randomization:** Omitting heading randomization ($\pm \pi/2$) resulted in a policy that succeeded exclusively on North-to-South turns but failed completely on alternating South-to-North entries ($\eta_H = 48.3\%$).
4. **Sensitivity of Handoff Threshold τ_h :** Sweeping τ_h confirmed that the analytically derived threshold ($\tau_h = 11.0$ m) is optimal. At $\tau_h = 7.0$ m, the vehicle had insufficient longitudinal distance to decelerate, causing wall boundary collisions (14%) and degraded compliance (88.2%). Conversely, larger values ($\tau_h = 15.0$ m) yielded identical success but incurred 19.6% longer turning paths.

Analysis of SAC Failure Modes and Supervisory Fallback Cases

To address deployment safety, we systematically analyzed cases where the learned SAC policy failed to reach the target waypoint autonomously:

1. **Step Budget Exhaustion ($T > 300$ steps):** In dense obstacle environments, when multiple obstacles flank the row entry corridor, SAC can enter oscillatory avoidance loops that exhaust the 300-step budget before arriving within 1.5 m of the row entry.
2. **Lateral Choke-Point Entrapment:** When obstacles reduce the corridor width to < 3.0 m, the policy slows down but cannot identify a collision-free path satisfying $r_{min} = 2.97$ m.

3. **Boundary Margin Violation Risk:** Overshoot within 1.5 m of the field border triggering safety halts.

The supervisory controller in Algorithm 1 detects these conditions immediately: upon step exhaustion or boundary clearance violations, forward velocity is arrested ($v \rightarrow 0$), and a constrained A* / Dubins recovery planner generates a safe connecting path to the next strip entry. In our empirical testing across all 26 headland turns in 30 fields, autonomous SAC success was 26/26 (100.0%) in Easy and Medium fields, and 25/26 (96.2%) in Hard fields. The single fallback instance was smoothly handled by the recovery controller with zero collisions and 100% field completion.

H. Comparison with Alternative Hierarchical Frameworks

To formally establish novelty over existing hierarchical combinations, HybridNav was benchmarked against A*+DWA [4], Hybrid A*+NMPC [5], and End-to-End SAC [14]. Table VII presents the comparative evaluation across 30 full-field trials in the Hard obstacle setting.

Table 7: Benchmarking Against Alternative Hierarchical Frameworks (Hard Setting)

Architecture	Coverage (%)	ηA_{cont} (%)	Turn Succ. (%)	Step Latency	Collisions
A* + DWA [4]	100.0	71.4	63.3	12.4 ms	8
Hybrid A* + NMPC [5]	100.0	96.2	90.0	245.8 ms	2
End-to-End SAC [14]	41.2	97.1	83.3	1.2 ms	5
HybridNav (Ours)	100.0	97.9	98.4	1.4 ms	0

A*+DWA failed to execute consistent U-turns in confined headlands, suffering 8 collisions and achieving only 71.4% continuous compliance. Hybrid A*+NMPC achieved 96.2% compliance but required 245.8 ms per step to solve the constrained non-linear program online; during 2 trials near dense obstacles, the solver timed out, resulting in boundary collisions. End-to-End SAC achieved rapid inference (1.2 ms) and smooth turning, but wandered across the field, achieving only 41.2% field coverage due to the absence of global spatial memory. HybridNav alone successfully combined 100% complete coverage, 97.9% continuous compliance, zero collisions, and 1.4 ms real-time inference latency.

CONCLUSION AND FUTURE DIRECTIONS

This paper introduced HybridNav, a hierarchical motion planning architecture that resolves the fundamental conflict between complete agricultural coverage and kinematically feasible headland turning. By allocating rectilinear row traversal to a deterministic global A* planner and delegating constrained headland maneuvers to an obstacle-aware SAC policy at an analytically derived threshold of $\tau_h = 11.0$ m, HybridNav ensures both theoretical completeness and continuous physical drivability.

The integration of an 8-beam rangefinder obstacle state ($s_t \in \mathbb{R}^{16}$) provides robust collision avoidance in cluttered field margins. Furthermore, the proposed Continuous Ackermann Compliance Metric (ηA^{cont}) establishes a rigorous standard for evaluating vehicle feasibility by jointly bounding instantaneous path curvature and actuator steering slew rates. Complete empirical benchmarking across Easy, Medium, and Hard fields alongside rigorous non-parametric statistical tests confirmed the significant superiority of the framework.

Future research will extend this framework to high-fidelity Gazebo dynamic simulations incorporating tire slip, soil compaction, and vehicle roll dynamics, followed by real-world deployment on autonomous tractor testbeds. Additional investigations will address non-rectangular field boundaries and coordinated multi-vehicle headland scheduling.

REFERENCES

- S. Zhang et al., "Trajectory Planning Optimization for Farming Using Automated Tractor," *IEEE Access*, vol. 13, pp. 11204–11218, 2025.
- F. Ye, F. Le, and X. Xue, "Application of Navigation Path Planning and Trajectory Tracking Control Methods for Agricultural Robots," *Computers and Electronics in Agriculture*, vol. 228, p. 109612, 2025.
- P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Trans. Syst. Sci. Cybern.*, vol. 4, no. 2, pp. 100–107, 1968.
- X. Li et al., "Precise Path Planning Based on Improved A* and DWA Algorithms for Complex Agricultural Environments," *Computers and Electronics in Agriculture*, vol. 229, p. 109745, 2025.
- M. Lu et al., "Path-Tracking Hybrid A* and Hierarchical MPC Framework for Autonomous Agricultural Vehicles," *arXiv:2411.14086*, 2025.
- L. Thiyagarasa et al., "A Navigation Stack Combining DH-ResNet18 Waypoint Regression, Pixel-to-Odometry Mapping, A* Planning, and MPC," *arXiv:2501.03211*, 2025.
- G. Mier et al., "Continuous Curvature Path Planning for Headland Coverage With Agricultural Robots," *J. Field Robot.*, vol. 42, no. 3, pp. 641–656, 2025.
- S. Bayano-Tejero, D. Sarri, and R. R. Sola-Guirado, "Design and Field Validation of a Dual-Axle Steering System for Autonomous Tractors," *Computers and Electronics in Agriculture*, vol. 230, p. 109820, 2025.
- Wei et al., "Efficient and Safe Trajectory Planning for Autonomous Agricultural Vehicle Headland Turning in Cluttered Orchard Environments," *IEEE Robot. Autom. Lett.*, vol. 10, no. 3, pp. 2574–2581, 2025.
- Y. Kim et al., "Quick Turn Headland-Turning Control Strategy for Autonomous Rice Transplanters," *Computers and Electronics in Agriculture*, vol. 228, p. 109580, 2025.
- L. Lavan et al., "Crop Spirals: Re-thinking the Field Layout for Future Robotic Agriculture," *arXiv:2509.25091*, 2025.
- S. Karaman and E. Frazzoli, "Sampling-Based Algorithms for Optimal Motion Planning," *Int. J. Robot. Res.*, vol. 30, no. 7, pp. 846–894, 2011.
- S. Fujimoto, H. van Hoof, and D. Meger, "Addressing Function Approximation Error in Actor-Critic Methods," in *ICML*, 2018, pp. 1587–1596.
- A. L. Mineiro, F. Affonso, and M. Becker, "End-to-End Crop Row Navigation via LiDAR-Based Deep Reinforcement Learning," in *IEEE ICAR*, 2025, pp. 811–817.
- T. Baby, M. K. Gohil, and B. Bhattacharya, "Autonomous Navigation of 4WIS4WID Agricultural Robot Using Deep Reinforcement Learning," in *ASME IMECE-India*, 2025.

- B. Han and J. Sun, "Road-Adaptive Precise Path Tracking Based on Reinforcement Learning," *Sensors*, vol. 25, no. 15, p. 4533, 2025.
- T. Haarnoja et al., "Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor," in *ICML*, 2018, pp. 1861–1870.
- C. Wang et al., "Field-Validated VIO-MPC Fusion for Autonomous Headland Turning in GNSS-Denied Orchards," *arXiv:2502.11200*, 2025.
- Y. Chen, Y. Liu, and W. Xu, "Improved Hybrid A* Algorithm Based on Lemming Optimization for Path Planning," *Applied Sciences*, vol. 15, no. 2, p. 842, 2025.
- J. Gan et al., "Hierarchical Iterative Search-Based Trajectory Planning for Self-Driving Heavy-Duty Semi-Trailer Tractors," *Chinese J. Mech. Eng.*, vol. 38, p. 14, 2025.
- O. Khatib, "Real-Time Obstacle Avoidance for Manipulators and Mobile Robots," *Int. J. Robot. Res.*, vol. 5, no. 1, pp. 90–98, 1986.
- A. Y. Ng, D. Harada, and S. Russell, "Policy Invariance Under Reward Transformations: Theory and Application to Reward Shaping," in *ICML*, 1999, pp. 278–287.