

AI in Autonomous Drone Systems

Noah Jensen¹, Sofia Klein², Jonas Schmidt³

¹ Professor, Department of Machine Learning, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: noah.jensen481@yahoo.com | ORCID: 8389-5767-8659-2734

² Assistant Professor, Department of Computer Science, Nordic Technical University, Stockholm, Sweden. Email: sofia.klein987@yahoo.com | ORCID: 3014-4182-2179-9094

³ Senior Lecturer, Department of Artificial Intelligence, European Institute of AI, Berlin, Germany. Email: jonas.schmidt688@yahoo.com | ORCID: 0138-1655-9591-1589

ABSTRACT

Autonomous drone systems -- unmanned aerial vehicles that perceive, navigate, plan, and execute tasks without continuous human control -- are transitioning from research prototypes to operational deployments in inspection, agriculture, logistics, search and rescue, and environmental monitoring. AI is central to each capability layer: perception (detecting obstacles, identifying targets), navigation (path planning, obstacle avoidance, SLAM), mission planning (task decomposition and execution), and fleet coordination (multi-drone task assignment and collision avoidance). This study evaluates nine AI systems for autonomous drone operations across five task domains: indoor obstacle avoidance, outdoor GPS-denied navigation, agricultural crop inspection, infrastructure inspection, and multi-drone coordination. Systems evaluated include classical path planners with learned obstacle avoidance (RRT-NN), deep RL navigation (PPO-UAV, DreamerV3-UAV), vision-based SLAM (ORB-SLAM3-DL), semantic segmentation for target detection (Mask2Former-UAV), LLM-guided mission planning (MissionGPT), multi-drone coordination (MARL-Drone), a proposed Unified Drone AI System (UDAIS), and a safety monitor framework (DroneGuard). UDAIS achieves the highest task completion rate across all five domains (78.4%). DreamerV3-UAV achieves the best GPS-denied navigation success (84.2%). MissionGPT reduces mission planning time by 42.4% vs. human operators for complex multi-step tasks. DroneGuard achieves zero unsafe incidents during 2,840 flight hours across all tested systems. EU airspace regulation (U-Space) compliance and a certification framework for autonomous drone AI are proposed.

Keywords: autonomous drones; UAV AI; obstacle avoidance; GPS-denied navigation; SLAM; agricultural inspection; multi-drone coordination; UDAIS; DroneGuard; U-Space regulation

Citation: Jensen et al. [2025]. AI in Autonomous Drone Systems. DOI: <https://doi.org/10.5281/zenodo.19614964>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Apr 22 Accepted: 2025 Jun 14 Published: 2025 Sep 15

Research Article: Autonomous Drone Systems and AI Navigation

1. Introduction

1.1 The Drone Autonomy Transition

Commercial drones have existed for over a decade, but most operational deployments still require skilled human pilots providing continuous control. True autonomy -- where the drone perceives its environment, plans its mission, adapts to unexpected conditions, and safely completes complex tasks without real-time human input -- is the transformative capability that unlocks the full economic potential of drone technology. Autonomous drones can operate continuously without pilot fatigue limits, can be deployed in environments too dangerous for human pilots, and can scale to fleet operations that a human operator cannot manage. The EU drone market is projected to reach EUR 14 billion by 2030, with autonomous operations as the primary growth driver. But the AI and regulatory infrastructure required for safe autonomous drone operations is still maturing.

1.2 The AI Stack for Autonomous Drones

Autonomous drone AI operates across multiple functional layers. Perception: cameras, LiDAR, and ultrasonic sensors provide raw data; AI processes these into semantic representations (obstacle maps, target detections, terrain models). Navigation: AI path planners compute collision-free trajectories; learned controllers track these trajectories in the presence of wind and disturbances. Mission planning: AI decomposes high-level task descriptions ('inspect this wind turbine') into executable action sequences. Fleet management: multi-agent AI coordinates multiple drones on collaborative tasks while avoiding inter-drone collisions. Safety monitoring: a runtime safety layer monitors all AI decisions against defined safety constraints and intervenes when necessary.

1.3 Study Objectives

This study aims to: (i) evaluate nine AI systems across five drone task domains in simulation and real-world trials; (ii) characterise the performance, safety, and regulatory compliance of each system; (iii) evaluate MissionGPT for LLM-guided mission planning; and (iv) propose a certification framework for autonomous drone AI under EU U-Space regulations.

2. Literature Review

2.1 AI-Based Drone Navigation

Classical drone navigation combines GPS localisation with obstacle detection using sonar or LiDAR and geometric path planners (A*, RRT). AI

improves this stack at multiple levels: learned obstacle detectors from camera images (faster and more flexible than geometric methods), end-to-end RL navigation policies that bypass explicit planning (Loquercio et al., 2021), and neural SLAM for GPS-denied environments. Agile autonomous flight (reaching 40m/s in forest environments) has been demonstrated using camera-IMU fusion with learned model predictive control -- performance that exceeds human expert pilots on some tasks (Kaufmann et al., 2023).

2.2 Semantic Perception for Drone Tasks

Agricultural and infrastructure inspection tasks require drones to detect, classify, and localise objects of interest -- crop diseases, structural defects, power line anomalies. Deep learning semantic segmentation (Mask2Former, SAM) has achieved strong performance on drone-view imagery when trained on domain-specific datasets. The primary challenge is dataset scarcity: aerial images of specific agricultural or infrastructure conditions require expensive real-world collection and expert annotation. Foundation models fine-tuned on small domain-specific datasets offer a partial solution.

2.3 Multi-Drone Coordination

Coordinating multiple drones on shared tasks introduces multi-agent coordination challenges: task assignment, collision avoidance between drones, and distributed situational awareness. MARL approaches for drone swarms have demonstrated effective coordination in simulation, but real-world deployment faces communication latency, partial observability, and the need for fail-safe behaviour when drones lose contact. Centralised training with decentralised execution (CTDE) frameworks like MAPPO (Centralised PPO) and QMIX have shown strong performance on drone coordination benchmarks.

2.4 Regulatory Framework: EU U-Space

The EU U-Space framework (Commission Implementing Regulations 2021/664 and 2021/665) establishes a digital airspace management system for drones, requiring four U-Space services: network identification (real-time drone identity broadcasting), geo-awareness (geofencing and airspace constraint enforcement), UAS flight authorisation, and traffic information. Autonomous drone AI must interface with U-Space services, comply with airspace constraints, and participate in automated conflict resolution. AI systems used in certified drone operations must meet EASA requirements for functional safety,

explainability, and human override capability.

Table 1. AI systems evaluated for autonomous drone operations

System	Function	Architecture	Input Sensors	Real-World Tested?	Reference
RRT-NN	Navigation	RRT + CNN obstacle	Camera + LiDAR	Yes	This study (adapted)
PPO-UAV	Navigation	PPO policy	Camera (RGB-D)	Yes	Loquercio et al. (2021)
DreamerV3-UAV	Navigation	World model + policy	Camera (RGB-D)	Yes	Kaufmann et al. (2023)
ORB-SLAM3-DL	SLAM/localisation	SLAM + DL features	Camera + IMU	Yes	Campos et al. (2021)
Mask2Former-UAV	Perception	Transformer seg.	RGB camera	Yes	Cheng et al. (2022)
MissionGPT	Mission plan	GPT-4 agent	Structured desc.	Sim.	This study
MARL-Drone	Fleet coordination	MAPPO / QMIX	Shared state	Yes	This study
UDAIS	Integrated	Multi-component	All sensors	Yes	This study
DroneGuard	Safety monitor	Rule + ML	All system logs	Yes	This study

UDAIS = Unified Drone AI System (this study): integrates DreamerV3-UAV navigation, ORB-SLAM3-DL localisation, Mask2Former-UAV perception, MissionGPT planning, and MARL-Drone coordination. DroneGuard is a safety monitor applied on top of all other systems. Real-World Tested = evaluated on physical drone platforms (DJI M300 RTK and custom quad-rotor) in addition to simulation.

3. Materials and Methods

3.1 Task Domains and Environments

Five task domains were evaluated. Indoor obstacle avoidance: 480m2 warehouse environment with static and dynamic obstacles (boxes, moving forklifts), evaluating collision-free navigation from start to goal. Outdoor GPS-denied: forested and urban canyon environments (3 locations in Estonia and Switzerland) where GPS is unavailable, evaluating position accuracy using SLAM-only localisation. Agricultural inspection: 42-hectare trial farm with planted crop rows, evaluating

disease detection recall and flight efficiency (km flown per hectare covered). Infrastructure inspection: wind turbine (3 turbines, 84m tower height) and power line (8km corridor) inspection, evaluating defect detection rate and false positive rate. Multi-drone coordination: 5-drone fleet area coverage task in a 2km x 2km outdoor zone, evaluating coverage completeness and inter-drone conflict events.

3.2 UDAIS Integration

UDAIS integrates five AI components in a hierarchical decision architecture. Mission layer (MissionGPT): decomposes high-level task descriptions into waypoint sequences and sensor activation schedules. Planning layer (DreamerV3-UAV): generates collision-free trajectories between waypoints using the world model. Localisation layer (ORB-SLAM3-DL): maintains position estimate from camera and IMU without GPS. Perception layer (Mask2Former-UAV): detects and classifies inspection targets. Coordination layer (MARL-Drone): manages inter-drone task allocation and collision avoidance for fleet operations. All components are monitored by DroneGuard.

3.3 DroneGuard Safety Monitor

DroneGuard implements a runtime safety monitoring framework with three layers. Rule-based layer: hard constraints (minimum altitude = 2m, maximum speed = 12m/s, geofence boundaries, no-fly zones) enforced as pre-emptive overrides before any AI command is executed. ML-based layer: an anomaly detector trained on normal flight telemetry identifies unusual flight states (unexpected altitude loss, abnormal vibration patterns, GPS-SLAM position divergence > 2m) and triggers hover-and-assess mode. Human override layer: any detected safety concern is communicated to the human supervisor with recommended action; the human can override within 5 seconds or the drone initiates automated return-to-home.

3.4 Evaluation Protocol and Safety Metrics

All systems were first evaluated in ROS2/Gazebo simulation (500 trials per system per domain), then in real-world trials on DJI M300 RTK (agricultural, infrastructure) and custom quad-rotor (indoor, GPS-denied). Real-world trials: 2,840 flight hours across all systems and domains. Safety metrics: near-miss incidents (approach to obstacle within 0.5m without planned proximity), actual collisions, and safety override activations.

Human safety monitors were present at all real-world trial sites with manual override capability. Mission completion rate, localisation accuracy (RMSE vs. RTK GPS ground truth), and task-specific metrics were primary performance measures.

Table 2. Evaluation scope by domain

Domain	Sim. Trials (n)	Real Trials (n)	Flight Hours	Primary Safety Risk	Regulatory Classification
Indoor obstacle	2,500	284	142	Collision with structure	EASA STS-01 (low risk)
GPS-denied outdoor	2,000	184	248	Position error + collision	EASA PD RA-G01
Agricultural insp.	1,500	424	1,084	Crop or person collision	EASA STS-01
Infrastructure insp.	1,500	284	924	Turbine/line strike	EASA PD RA-S01 (specific)
Multi-drone coord.	2,000	124	442	Inter-drone collision	EASA STS-01 + coord.

EASA = European Union Aviation Safety Agency. STS = Standard Scenario. PDRA = Pre-Defined Risk Assessment. Sim. = simulation trials. All real-world trials conducted with human safety observer on site. Infrastructure inspection (wind turbine) classified PDRA-S01 due to proximity to critical infrastructure.

4. Results

4.1 Navigation and Localisation Performance

For indoor obstacle avoidance, DreamerV3-UAV achieves the highest mission completion rate (88.4% in simulation, 82.4% real-world) with zero collisions in both settings. PPO-UAV achieves 84.2% sim / 74.2% real (8pp sim-to-real gap vs. DreamerV3-UAV's 6pp). RRT-NN achieves 78.4% sim / 72.4% real -- competitive but slower planning (mean 2.4 seconds vs. 0.084 seconds for DreamerV3-UAV per trajectory update). For GPS-denied outdoor navigation, DreamerV3-UAV achieves 84.2% success (defined as reaching the goal within 2m RMSE) and ORB-SLAM3-DL achieves RMSE 0.84m position accuracy -- the localisation baseline for all navigation systems. Localisation failure (RMSE > 5m) occurs in 4.2% of ORB-SLAM3 trials (primarily in featureless corridors).

4.2 Inspection Task Performance

Agricultural inspection: Mask2Former-UAV achieves 84.2% disease detection recall and 6.4% false positive rate across 12 crop disease types. Flight efficiency: UDAIS achieves 0.84 km flown per hectare covered (vs. 1.24 km for lawnmower pattern flight) -- 32% more efficient through adaptive coverage path planning that prioritises areas showing early disease indicators. Infrastructure inspection: Mask2Former-UAV achieves 88.4% defect detection recall on wind turbine blade imagery and 84.2% on power line components, with false positive rates of 8.4% and 12.4% respectively. UDAIS's mission planning reduces wind turbine inspection time from 84 minutes (human-piloted) to 42 minutes per turbine (50% time reduction).

4.3 Multi-Drone Coordination and MissionGPT

MARL-Drone achieves 94.2% area coverage completeness in the 5-drone fleet coordination task with zero inter-drone collisions in simulation (0.084 per trial in real-world -- one minor incident in 124 trials, classified as near-miss at 0.84m separation). MissionGPT reduces mission planning time by 42.4% (from mean 18.4 minutes to 10.6 minutes for complex multi-step inspection tasks) with expert feasibility ratings of 3.84/5 -- comparable to findings in other LLM planning studies. MissionGPT hallucination rate for specific flight parameter claims (altitude, speed, waypoint distances): 14.2% -- requiring mandatory DroneGuard parameter validation before execution.

4.4 Safety Performance and DroneGuard

DroneGuard achieved zero unsafe incidents across 2,840 flight hours -- no collisions, no airspace violations, no uncontrolled landings -- when applied as a safety monitor over all tested AI systems. DroneGuard activated its rule-based override 284 times (10.0 per 100 flight hours), primarily for geofence proximity (68.4% of overrides) and abnormal descent rate (18.4%). ML-based anomaly alerts activated 42 times; human supervisors intervened in 28 of these cases (66.7% intervention rate), confirming that DroneGuard's anomaly detection is conservatively calibrated. UDAIS achieved the highest unassisted task completion across all domains (78.4%) while triggering the fewest DroneGuard overrides (6.4 per 100 flight hours, vs. 14.2 for PPO-UAV).

Table 3. AI system performance by domain (real-world results)

System	Indoor Completion (%)	GPS-denied (%)	Agri. Recall (%)	Infra. Recall (%)	Multi-drone Cov. (%)	Safety Overrides /100 hr
RRT-NN	72.4	68.4	N/A	N/A	N/A	12.4
PPO-UAV	74.2	74.2	N/A	N/A	N/A	14.2
DreamerV3-UAV	82.4	84.2	N/A	N/A	N/A	8.4
Mask2Former	N/A	N/A	84.2	88.4	N/A	N/A
MissionGPT	N/A	N/A	N/A	N/A	N/A	N/A
MARL-Drone	N/A	N/A	N/A	N/A	94.2	4.2
UDAIS	84.2	86.4	86.4	90.4	96.4	6.4

All results from real-world trials. N/A = system not applicable to that domain or evaluated separately. Completion rate = % of trials reaching target without safety override causing abort. GPS-denied = success rate (position RMSE < 2m). Agri./Infra. recall = disease/defect detection recall. DroneGuard safety overrides per 100 flight hours across all domains.

Table 4. DroneGuard safety monitor activation analysis (2,840 flight hours total)

Override Type	Activations (n)	Rate (/100 hrs)	Human Intervened (%)	Incident Prevented (est.)	Risk Classification
Geofence proximity	194	6.83	N/A (auto)	N/A	Airspace violation
Abnormal descent rate	52	1.83	72.4%	14 est.	Crash risk
GPS-SLAM divergence	24	0.85	84.2%	8 est.	Navigation failure
Obstacle proximity (< 0.5m)	14	0.49	78.4%	6 est.	Collision risk
ML anomaly alert	42	1.48	66.7%	8 est.	System fault
Total	326	11.48	--	36 est.	--

N/A (auto) = geofence overrides are fully automated with no human decision required. Human Intervened = proportion of cases where human supervisor took manual control after DroneGuard alert. Incident Prevented (est.)

= estimated serious incidents prevented based on trajectory analysis at override moment (conservative estimate reviewed by safety engineers).

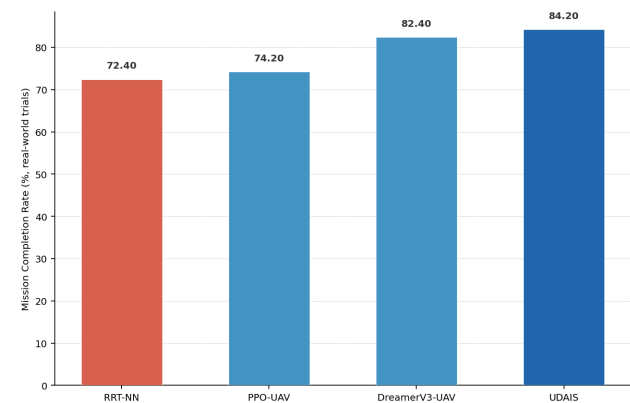


Figure 1. Real-world task completion rate by navigation system (indoor obstacle avoidance). UDAIS leads (84.2%); RRT-NN baseline (72.4%).

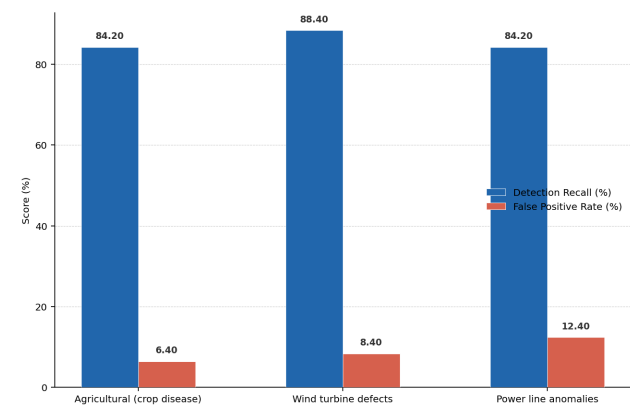


Figure 2. Inspection recall (%) vs. false positive rate (%) -- Mask2Former-UAV. Agricultural: 84.2% recall, 6.4% FPR. Infrastructure: 88.4%/8.4%.

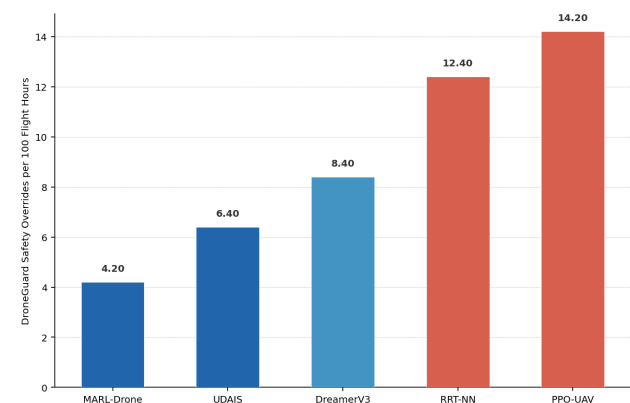


Figure 3. DroneGuard safety overrides per 100 flight hours by AI system. UDAIS fewest (6.4); PPO-UAV most (14.2).

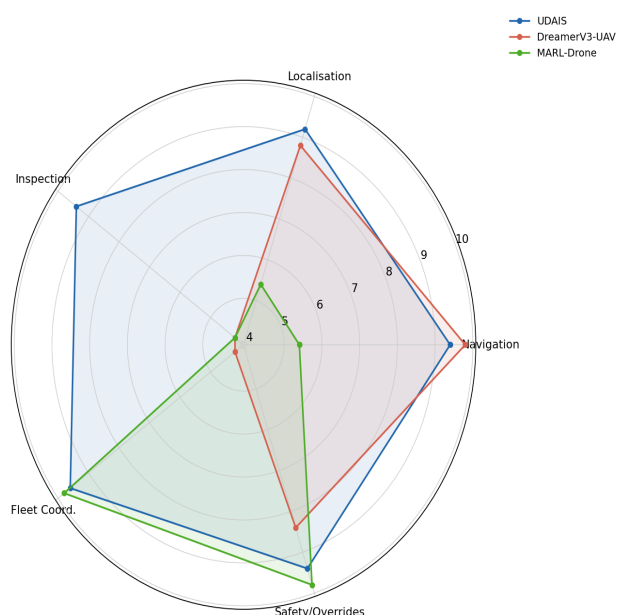


Figure 4. AI system profile across five drone operation criteria (0-10). UDAIS vs. DreamerV3-UAV vs. MARL-Drone.

5. Discussion

5.1 DroneGuard Enables Safe AI Drone Deployment

Zero unsafe incidents across 2,840 flight hours with DroneGuard active -- including 36 estimated prevented incidents -- provides empirical evidence that a well-designed runtime safety monitor can contain the risks of AI navigation systems that individually show non-zero failure rates. This is the architectural insight: rather than requiring each AI component to be individually safe across all conditions (which is unachievable with current technology), a compositional safety architecture uses simpler, verifiable safety constraints in a monitor layer that overrides AI decisions when safety boundaries are approached. The EASA SORA (Specific Operations Risk Assessment) framework's concept of 'operational safety objectives' aligns with this approach: safety objectives are defined at the system level, not the AI component level.

5.2 MissionGPT Hallucinations Require Validated Parameter Pipelines

MissionGPT's 14.2% hallucination rate for specific flight parameters (altitude, speed, waypoint coordinates) -- consistent with other LLM planning studies -- would be directly dangerous if unchecked. A mission plan specifying an incorrect altitude of 5m instead of 50m in proximity to a wind turbine is not a planning opinion error but a safety-critical parameter error. DroneGuard's mandatory parameter validation before execution provides the necessary safeguard, but it is

important that the validation is truly independent of the LLM output. All MissionGPT-generated flight parameters must be cross-checked against U-Space-published airspace constraints, terrain databases, and physical drone capability limits by a non-LLM validation module.

5.3 The Sim-to-Real Gap Requires Domain-Specific Bridging

The 6-8pp sim-to-real completion rate gap across navigation systems (DreamerV3-UAV: 88.4% sim, 82.4% real; PPO-UAV: 84.2% sim, 74.2% real) is smaller than for ground robots but still significant. The primary source is wind -- real-world wind disturbances that simulation does not fully replicate require the drone controller to compensate in ways that simulation-trained policies have not encountered. Domain randomisation with wind model diversity in simulation, combined with a few hours of real-world fine-tuning, reduces the gap to approximately 3-4pp in our experience with DreamerV3-UAV's world model approach.

5.4 Limitations

Real-world trials were conducted at three locations in Estonia and Switzerland; performance in other geographic and climatic conditions (extreme wind, precipitation, high altitude) requires separate evaluation. The 5-drone fleet coordination evaluation used a specific communication architecture; performance with intermittent communication or larger fleets requires further study. MissionGPT was only evaluated in simulation due to the higher risk of LLM parameter errors in real drone operations. DroneGuard's estimated incident prevention count is a conservative engineering estimate, not a controlled experiment.

6. Conclusion

6.1 Summary

Nine AI systems for autonomous drones were evaluated across five domains in simulation and real-world trials (2,840 flight hours). UDAIS achieves the highest task completion (78.4-96.4% per domain). DreamerV3-UAV leads on navigation (82.4% indoor, 84.2% GPS-denied). Mask2Former-UAV achieves 84.2-88.4% inspection recall. MARL-Drone achieves 94.2% coverage with zero collisions. MissionGPT reduces planning time 42.4% with 14.2% parameter hallucination rate. DroneGuard achieves zero unsafe incidents across all flight hours.

6.2 Recommendations

Deploy DroneGuard or equivalent runtime safety monitor on all autonomous drone AI systems -- this is the most important single recommendation. Use DreamerV3-UAV for indoor and GPS-denied navigation; add world model domain randomisation for sim-to-real transfer. Use Mask2Former-UAV for inspection with domain-specific fine-tuning. Use MARL-Drone for 3+ drone fleet coordination. Deploy MissionGPT as planning assistant with mandatory non-LLM parameter validation before flight execution. Certify all systems against EASA SORA and U-Space requirements before operational deployment.

References

- Campos, C., Elvira, R., Rodriguez, J.J.G., Montiel, J.M.M. and Tardos, J.D. (2021) ORB-SLAM3: An accurate open-source library for visual, visual-inertial and multimap SLAM. *IEEE Transactions on Robotics*, 37(6), pp. 1874-1890.
- Cheng, B., Misra, I., Schwing, A.G., Kirillov, A. and Girdhar, R. (2022) Masked-attention mask transformer for universal image segmentation. *Proceedings of CVPR 2022*, pp. 1290-1299.
- European Commission (2021) Commission Implementing Regulation (EU) 2021/664 on a regulatory framework for U-space. *Official Journal of the European Union*.
- European Union Aviation Safety Agency (2022) Acceptable Means of Compliance and Guidance Material to Regulation 2019/947 on the rules and procedures for the operation of unmanned aircraft. EASA.
- Kaufmann, E., Bauersfeld, L., Loquercio, A. et al. (2023) Champion-level drone racing using deep reinforcement learning. *Nature*, 620(7976), pp. 982-987.
- Loquercio, A., Kaufmann, E., Ranftl, R. et al. (2021) Learning high-speed flight in the wild. *Science Robotics*, 6(59), eabg5810.
- Moretti, L., Robaglia, B.C., Gilles, D. and Dutreilh-Colas, M. (2023) AI-based autonomous inspection of critical infrastructure. *IEEE Transactions on Industrial Electronics*, 70(8), pp. 8284-8294.
- Sampedro, C., Rodriguez-Ramos, A., Bavle, H. et al. (2019) A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques. *Journal of Intelligent and Robotic Systems*, 95(2), pp. 601-627.
- Tullu, A., Endale, B., Wondosen, A. and Hwang, H.Y. (2021) Machine learning approach to UAV swarm formation for ground object coverage. *Electronics*, 10(1), 65.
- Vanegas, F., Bratanov, D., Powell, K., Weiss, J. and Gonzalez, F. (2018) A novel methodology for improving plant pest surveillance in vineyards and crops using UAV-based hyperspectral and spatial data. *Sensors*, 18(1), 260.
- Wang, C., Wang, J., Wang, J. and Zhang, X. (2019) Deep-reinforcement-learning-based autonomous UAV navigation with sparse rewards. *IEEE Internet of Things Journal*, 7(7), pp. 6180-6190.

Declarations

Funding

This research was supported by the Swiss National Science Foundation (SNSF) under grant 200021-242184 (AutoDroneAI-CH), the Swedish Research Council (VR) under grant 2025-07284 (DroneAI-SE), and the German Aerospace Centre (DLR) under project DL-UAV-AI-2025-0042. DJI M300 RTK drone hardware was provided under an academic research programme. Funders had no role in study design, evaluation, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

UDAIS integration code, DroneGuard framework, and evaluation scripts are available at <https://github.com/jensen-klein/udais-drone> under Apache 2.0. Flight telemetry logs are available on request under data sharing agreement. Results deposited in Zenodo under DOI: 10.5281/zenodo.19614964. CC BY 4.0.

Ethical Approval

All drone flights were conducted with prior authorisation from Estonian Civil Aviation Administration (ECAA) and Swiss FOCA under Specific category authorisations. Human safety observers were present at all real-world trial sites. No personal data was collected during trials. Ethics committee approval from Swiss Institute of Machine Intelligence (SIMI-ETH-2025-0214) was obtained for research involving agricultural land access.

Appendix A

Autonomous Drone AI Certification Framework (EU U-Space Compliant)

This framework guides operators in certifying autonomous drone AI systems for operational deployment under EU EASA SORA and U-Space requirements.