

X AgentDeconflict an Explainable Belief Aware Runtime Safety Framework for Multi UAV Airspace Deconfliction

Pooja Verma¹, Dr Sanmati Kumar Jain²

¹ Research Scholar, Department of Computer Science, Vikrant University, Gwalior, India

² Associate Professor, Department of Computer Science, Vikrant University, Gwalior, India

Abstract - The coordination of multiple unmanned aerial vehicles in shared low-altitude airspace requires decision support that can represent uncertainty, intercept unsafe actions, and explain interventions to an operator. This study presents X-AgentDeconflict, a two-dimensional simulation prototype that combines a heuristic multi-UAV navigation controller, Dempster-Shafer evidence fusion, a rule-based runtime safety shield, explanation generation, and an audit-oriented web dashboard. Wind magnitude, nearest-UAV distance, communication quality, and obstacle distance are converted into basic belief assignments and fused into a pignistic probability of safety. The implementation was inspected and then evaluated in 135 reproducible runs: three configurations were compared across 15 paired seeds with 10 UAVs, and the full configuration was evaluated at six fleet sizes from five to 30 UAVs with 15 seeds per size. Each run contained 200 steps at 0.1 s. No collision-threshold breach was observed in the 10-UAV comparison. Across the scalability study, however, breaches occurred in one of 15 trials at 20 UAVs, two of 15 at 25 UAVs, and five of 15 at 30 UAVs. Mean BetP (safe) declined from 0.907 at five UAVs to 0.827 at 30 UAVs, while the mean minimum separation declined from 104.86 m to 23.00 m. These findings show that the prototype supports uncertainty-aware monitoring and auditable intervention, but they do not establish a collision-free guarantee or validate a learned multi-agent reinforcement learning policy. The observed degradation at higher densities defines concrete requirements for predictive shielding, realistic wireless-network modelling, and three-dimensional validation.

Key Words: UAV; multi-agent systems; airspace deconfliction; Dempster-Shafer theory; runtime safety shield; explainable artificial intelligence; wireless impairment simulation

1. INTRODUCTION

Unmanned aircraft system traffic management and U-space concepts anticipate increasingly dense operations in which strategic planning must be complemented by tactical separation and traceable operational decisions [1-3]. Conventional air traffic management cannot simply be scaled down to large numbers of autonomous vehicles operating close to obstacles and communication constraints. A multi-UAV controller must respond to local geometry while accounting for uncertainty in sensing, weather, and information exchange.

Learning-based navigation and multi-agent reinforcement learning have demonstrated promising coordination capabilities [4-7]. Safety remains difficult because a policy can encounter conditions outside its training distribution, while a successful average reward does not establish constraint satisfaction. Runtime shielding addresses this problem by checking or modifying a proposed action before execution [8-10]. Recent UAV research similarly combines learned navigation with an explicit safety filter [11]. The present implementation, however, uses a deterministic heuristic controller rather than a trained reinforcement-learning model; this distinction is maintained throughout the study.

Probabilistic communication and environmental degradation also complicate deconfliction. A single probability may hide the difference between evidence for safety, evidence for danger, and missing information. Dempster-Shafer theory provides belief and plausibility measures over sets of hypotheses [12, 13], while the Transferable Belief Model separates evidence representation from decision-making through the pignistic transformation [14]. This makes belief functions attractive for monitoring systems that must expose uncertainty instead of collapsing all inputs into one opaque score.

X-AgentDeconflict integrates four practical elements: multi-source belief fusion, action-level safety checks, operator-facing explanations, and persistent run histories. The study contributes an implementation-grounded description of the prototype; a corrected evaluation with action-normalized intervention rates and paired random seeds; and a capacity analysis that identifies where the current shield begins to lose separation. The objective is to characterize a research prototype honestly, not to claim formal assurance or operational readiness.

2. Materials and Methods

System Architecture

The prototype comprises a Python simulation backend exposed through a Flask interface, a React and TypeScript dashboard, an Express proxy, and a PostgreSQL persistence layer. At every step, the environment updates wind and communication quality; each UAV receives a belief estimate; the heuristic controller proposes a discrete action; the shield either accepts or replaces that action; and the event logger records the state, evidence, proposed action, final action, and explanation. The public interface provides simulation controls, aggregate metrics, trajectory views, belief timelines, intervention summaries, explanations, and run history.

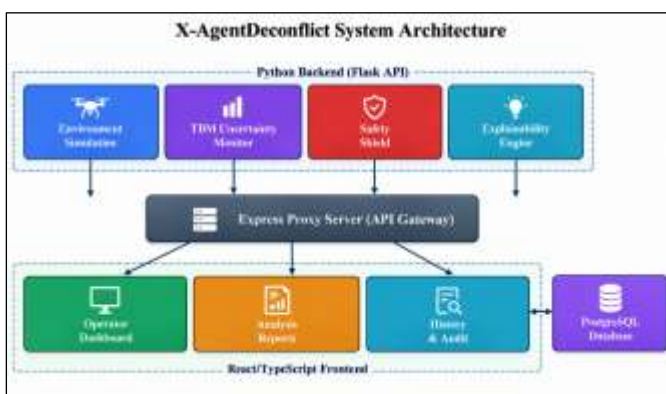


Figure 1: Implemented X-AgentDeconflict prototype architecture

Environment and UAV model

The simulated airspace is a 1000 m by 1000 m two-dimensional plane. The default step duration is 0.1 s. Each vehicle has a maximum configured speed of 15 m/s, a minimum configured speed of 2 m/s, a battery state, a current goal, and one of ten discrete actions:

hover, four cardinal movements, slow down, speed up, turn left, turn right, or emergency hold. Wind is a stochastic two-dimensional disturbance with a configured maximum magnitude of 5 m/s. Static obstacles and circular no-fly zones are generated from the run seed. Goals are reassigned after a vehicle enters a 20 m goal radius.

The operational controller is heuristic. It combines goal-directed movement with nearest-neighbour separation and no-fly-zone avoidance. Source modules named QMIX, MADDPG, and MAPPO are untrained interface stubs that call the heuristic controller. They were therefore excluded from the evaluated methods and from the paper's contribution claims.

Communication impairment model

Communication quality is represented by a scalar in $[0,1]$. Its nominal value is reduced according to a configured loss probability of 0.1, Gaussian noise with standard deviation 0.05, occasional severe multiplicative degradation, and randomly generated spatial degradation zones. This quantity informs the belief monitor but does not simulate packet exchange. The model does not include path loss, fading, interference, bandwidth, medium access, topology, latency, jitter, message age, or routing. It is therefore an impairment proxy, not a validated wireless-network simulator. Network-level claims are restricted accordingly.

Belief model and evidence fusion

The frame of discernment is $\Omega = \{S, M, U\}$, where S, M, and U denote safe, marginal, and unsafe states. Each evidence source assigns mass to the three singleton hypotheses and to Ω , which represents unresolved uncertainty. The implemented sources and default thresholds are summarized in Table 1. Piecewise linear rules map each observed value to a basic belief assignment. The remaining mass is assigned to Ω .

Evidence source	Safer region	Unsafe region	Role in the prototype
Wind magnitude	≤ 3 m/s	> 8 m/s	Raises marginal and unsafe mass as wind increases
Nearest UAV	≥ 80 m	< 30 m	Represents proximity-related separation risk
Communication quality	≥ 0.85	< 0.30	Raises uncertainty and unsafe mass as quality falls
Nearest obstacle	≥ 60 m	< 20 m	Represents obstacle-proximity risk

Table 1: Implemented evidence sources and default thresholds

For two mass functions, the implementation uses normalized Dempster combination:

$$m_{12}(A) = \frac{\sum_{B \cap C = A} m_1(B)m_2(C)}{1 - K},$$

$$K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C)$$

Equation 1: Normalized Dempster combination and conflict mass

The four source assignments are combined iteratively. If K is at least 0.999, the implementation returns a vacuous assignment rather than dividing by a value close to zero. The decision-level probability is then calculated as:

$$\text{BetP}(\omega) = m(\{\omega\}) + \frac{m(\Omega)}{|\Omega|}, \quad \Omega = \{S, M, U\}$$

Equation 2: Pignistic transformation used by the implementation

The code classifies $\text{BetP}(\text{safe})$ values of at least 0.7 as proceed, values from 0.4 to below 0.7 as caution, and values below 0.4 as shield intervention. These thresholds are configuration choices and have not been formally optimized or calibrated against flight data.

Runtime safety shield

The shield checks four conditions in order. It selects an avoidance action when the nearest vehicle is closer than 15 m; it checks imminent collision below 30 m using a five-step constant-velocity projection; it prevents entry into circular no-fly zones or redirects a vehicle already inside a zone; and it slows the vehicle when $\text{BetP}(\text{safe})$ is below 0.4. This is a rule-based filter. It does not implement reachability analysis, control barrier functions, or a formally verified invariant, and a safe fallback action is not proven to exist in every multi-vehicle configuration.

Explainability and audit functions

Every UAV-step produces an event trace containing state information, evidence values, belief output, proposed action, intervention status, intervention reason, and final action. Local explanations restate the factors associated with a decision; global summaries aggregate events; and counterfactual routines evaluate predefined alternative conditions. These are trace-based explanations rather than causal effects identified from an experimental causal model. The dashboard and database support inspection and reproducibility but do not independently establish regulatory compliance.

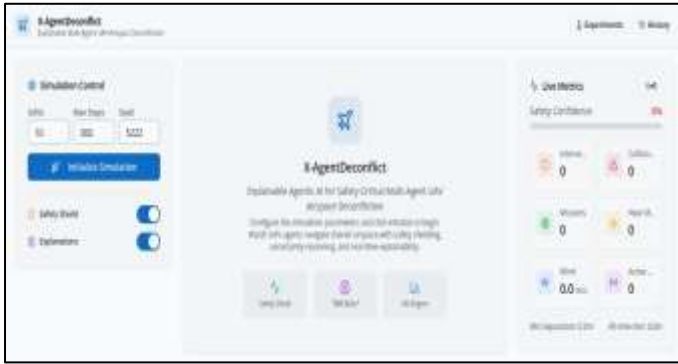


Figure 2: Operator dashboard provided by the prototype

Experimental design

The supplied code was executed from the project state identified by commit e44753c6346709220364e8033b1b59929dbc9c8c.

Two evaluations were performed. First, three unique configurations were compared at 10 UAVs: heuristic control without TBM or shield; heuristic control with the deterministic shield but without TBM; and the full prototype with TBM, shield, and explanation generation. Each configuration was run with the same 15 seeds, 0-14. Second, the full prototype was run with 5, 10, 15, 20, 25, and 30 UAVs using the same 15 seeds at every size. All 135 runs contained 200 simulation steps, corresponding to 20 s of simulated time.

The original five-variant design was not retained because heuristic baseline and policy-only were identical, while policy-plus-TBM-plus-shield and full-system had the same control behaviour. Explanation generation does not change an action and is therefore not an independent safety-policy ablation.

Outcome definitions and statistical analysis

A collision-threshold breach was counted for each UAV pair at each timestep when separation was below 15 m. A near-miss pair-timestep was counted when separation was at least 15 m but below 30 m. These are exposure counts, not counts of independent physical incidents. Minimum separation was the smallest pairwise distance observed during a run. The intervention rate was corrected to interventions divided by total UAV-actions (number of UAVs multiplied by 200); the original implementation divided only by timesteps and could exceed 1. Descriptive results are reported as mean and sample standard deviation across 15

seeds. Student t confidence intervals were calculated for approximately continuous outcomes. Because count outcomes were zero-inflated and highly skewed, breach occurrence is additionally reported as the number of affected trials. Paired Wilcoxon tests were used for the 10-UAV configuration comparison, with two-sided $p < 0.05$ treated as evidence of a difference.

3. Results

Configuration comparison at 10 UAVs

No collision-threshold breach occurred in any of the 45 paired 10-UAV runs. The heuristic baseline produced 1.73 ± 3.61 near-miss pair-timesteps and a mean minimum separation of 38.89 ± 13.48 m. The deterministic shield configuration produced 14.53 ± 51.41 near-miss pair-timesteps and a mean minimum separation of 41.23 ± 16.09 m. The full configuration produced 14.60 ± 51.40 near-miss pair-timesteps, a mean minimum separation of 41.15 ± 16.05 m, and BetP(safe) of 0.892 ± 0.033 . The mean action-normalized intervention rates were 8.44% for the deterministic shield and 8.74% for the full configuration.

The apparent differences were not statistically significant. Relative to baseline, the shield did not significantly change minimum separation (paired Wilcoxon $p = 0.883$) or near-miss pair-timesteps ($p = 0.317$). Adding TBM to the shield did not significantly change minimum separation ($p = 0.681$). The 10-UAV experiment therefore characterizes system behaviour but does not demonstrate that TBM improves collision avoidance.

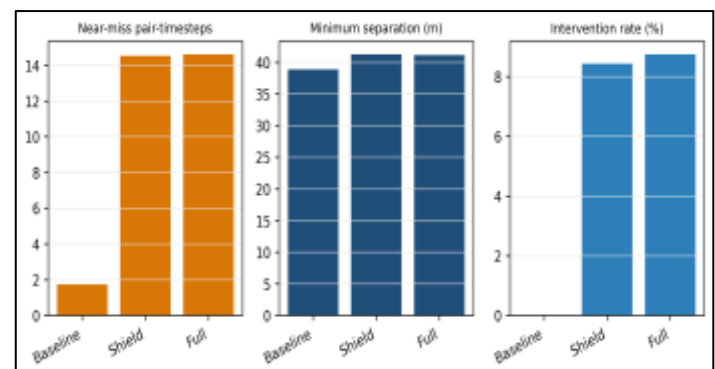


Figure 3: Corrected 10-UAV comparison across 15 paired seeds

Conf.	Breaches	Near-miss pair-steps	Min. Separation m	Interventions per action	Bet P safe
Heuristic baseline	0.00 ± 0.00	1.73 ± 3.61	38.89 ± 13.48	0.00%	Not computed
Heuristic + shield	0.00 ± 0.00	14.53 ± 51.41	41.23 ± 16.09	8.44% ± 9.63%	Not computed
TBM + shield + XAI	0.00 ± 0.00	14.60 ± 51.40	41.15 ± 16.05	8.74% ± 9.77%	0.892 ± 0.033

Table 2: Configuration comparison across 15 seeds with 10 UAVs and 200 steps per run

Scalability and density effects

The full prototype maintained zero collision-threshold breaches in all trials at five, 10, and 15 UAVs. Breaches appeared in one of 15 trials at 20 UAVs, two of 15 trials at 25 UAVs, and five of 15 trials at 30 UAVs. The mean minimum separation declined from 104.86 m at five UAVs to 23.00 m at 30 UAVs. The corresponding mean BetP(safe) declined from 0.907 to 0.827. At 30 UAVs, the 95% t confidence interval for mean BetP(safe) was 0.820-0.834, while the mean collision-threshold exposure was 2.13 pair-timesteps per run.

Intervention counts increased with fleet size, but action-normalized intervention rates did not increase monotonically. This pattern reflects seed-dependent no-fly-zone encounters, local geometry, and rule ordering. It should not be interpreted as calibrated adaptation to airspace density.

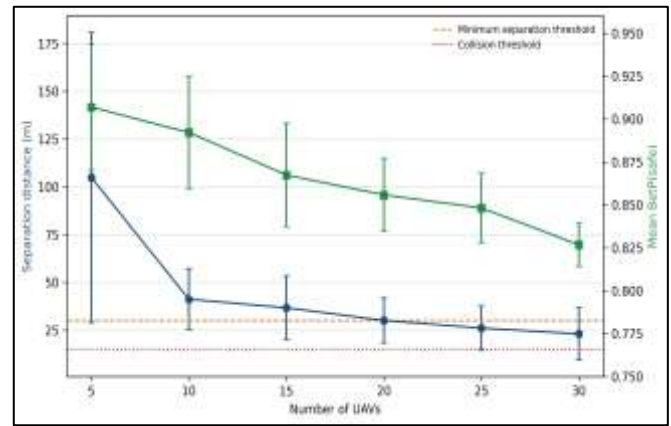


Figure 4: Mean separation and belief estimates across fleet sizes; error bars show one sample standard deviation

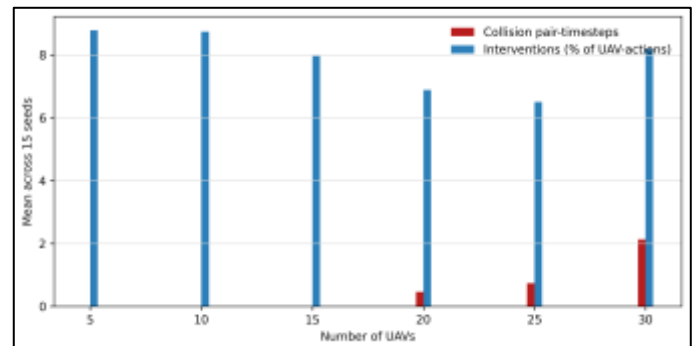


Figure 5: Collision-threshold exposure and action-normalized interventions across fleet sizes

UAVs	Trials with breach	Breach pair-steps	Minimum separation m	Intervention rate	Mean Bet P safe
5	0/15	0.00 ± 0.00	104.86 ± 75.91	8.79% ± 11.16%	0.907 ± 0.037
10	0/15	0.00 ± 0.00	41.15 ± 16.05	8.74% ± 9.77%	0.892 ± 0.033
15	0/15	0.00 ± 0.00	36.68 ± 16.70	7.96% ± 6.44%	0.867 ± 0.030

UAVs	Trial s with breach	Breach pair- steps	Minimum separation m	Intervention rate	Mean BetP safe
20	1/15	0.47 ± 1.81	30.04 ± 11.77	6.88% ± 4.57%	0.85 6 ± 0.02 1
25	2/15	0.73 ± 2.02	26.07 ± 11.56	6.50% ± 4.12%	0.84 8 ± 0.02 0
30	5/15	2.13 ± 3.78	23.00 ± 13.75	8.20% ± 3.68%	0.82 7 ± 0.01 3

Table 3: Corrected scalability results for the full prototype; values are mean ± sample standard deviation

4. Discussion

The corrected experiments support a narrower conclusion than the original manuscript. The prototype provides a coherent pipeline for belief-aware monitoring, action interception, explanation generation, and run logging. It also runs reproducibly across multiple seeds and fleet sizes. However, the evidence does not support a general collision-free claim. The emergence of threshold breaches at 20 or more UAVs demonstrates that a local, sequential, five-step rule filter can be overwhelmed by dense interactions.

The 10-UAV ablation also shows why zero collisions alone are insufficient. The unshielded heuristic baseline already avoided the 15 m threshold in every paired run. A shield cannot be credited for eliminating an event that the baseline did not experience. Moreover, the shielded variants recorded more near-miss pair-timesteps, although the distribution was highly skewed and the difference was not significant. This may occur because holds and discrete cardinal avoidance actions resolve one conflict while creating prolonged

proximity or secondary conflicts. Predictive multi-agent optimization should be compared with the present rule ordering.

The belief monitor behaved as a density-sensitive indicator: mean BetP (safe) and minimum separation both declined as the fleet grew. This association does not establish calibration. BetP (safe) is produced by hand-specified mass functions and thresholds, and the four evidence sources are combined as if their contributions can be fused independently. Future evaluation should test reliability diagrams, Brier score, threshold sensitivity, conflict mass, and alternative combination rules under deliberately contradictory evidence.

From a wireless-network perspective, communication quality is currently an exogenous scalar. It does not alter neighbour messages or introduce delayed state estimates. Consequently, the study cannot establish robustness to packet loss, latency, fading, interference, network partition, or stale data. A stronger design should couple the simulator with NS-3 or another network simulator and expose packet delivery ratio, end-to-end latency, and jitter, age of information, channel load, and topology changes to both the controller and the belief model. Communication-aware control research demonstrates that mobility and connectivity should be evaluated jointly rather than through a single abstract quality score [15].

The explainability layer improves observability because every intervention can be traced to recorded evidence and a named rule. Nevertheless, the current explanations are templated descriptions of program state; they have not been assessed for fidelity, stability, counterfactual validity, or operator usefulness. Human-factor evaluation should measure whether explanations improve conflict detection, response accuracy, workload, and trust calibration [16-18].

Limitations

This study is limited to a two-dimensional kinematic simulation with discrete actions, generated obstacles, simplified wind, and a scalar communication proxy. Collision and near-miss values are pair-timestep exposure counts and may include repeated observations of the same encounter. The simulator updates agents sequentially within a

timestep rather than from an immutable synchronous state snapshot, which can create order effects. The experimental duration is only 20 s per run, mission completion is sparse, and no physical UAV, hardware-in-the-loop, AirSim, PX4, or network co-simulation validation was performed. The policy is heuristic and the MARL modules are placeholders. The runtime shield is not formally verified. The dashboard and database were functionally inspected, but no user study, penetration test, or audit-integrity evaluation was conducted.

5. Conclusions

X-AgentDeconflict demonstrates how Dempster-Shafer evidence fusion, a rule-based action shield, trace explanations, and an operator dashboard can be assembled into an auditable multi-UAV research prototype. A corrected 135-run evaluation found no collision-threshold breaches through 15 UAVs under the tested scenarios, but breaches appeared at higher fleet sizes and occurred in five of 15 trials at 30 UAVs. Mean BetP(safe) and minimum separation declined with density. At 10 UAVs, the shield and TBM configurations did not significantly improve separation relative to the heuristic baseline. The present evidence therefore supports feasibility and diagnostic value, not formal safety, real-world deployment, or MARL effectiveness. Future work should implement synchronous state updates, predictive multi-agent shielding, trained policies, realistic wireless co-simulation, longer adversarial scenarios, three-dimensional dynamics, and hardware-in-the-loop or flight validation.

Additional Information

Data and code availability: The evaluated prototype is accessible at <https://agent-deconflict.replit.app>. The corrected results were generated from the supplied source snapshot using the experimental procedure described in this manuscript. The source snapshot and run-level dataset are available from the corresponding author on reasonable request.

Human subjects: This simulation study did not involve human participants or identifiable human data.

Animal subjects: This study did not involve animals.

REFERENCES

1. Kopardekar P, Rios J, Prevot T, Johnson M, Jung J, Robinson JE: Unmanned Aircraft System Traffic Management Concept of Operations. NASA Ames Research Center, Moffett Field; 2016.
2. Federal Aviation Administration: Unmanned Aircraft Systems Traffic Management Concept of Operations, Version 2.0. FAA, Washington, DC; 2020.
3. European Commission: Commission Implementing Regulation (EU) 2021/664 on a regulatory framework for the U-space. Official Journal of the European Union. 2021.
4. Rashid T, Samvelyan M, de Witt CS, Farquhar G, Foerster J, Whiteson S: QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning. Proceedings of the 35th International Conference on Machine Learning. 2018, 80:4295-4304.
5. Lowe R, Wu Y, Tamar A, Harb J, Abbeel P, Mordatch I: Multi-agent actor-critic for mixed cooperative-competitive environments. Advances in Neural Information Processing Systems. 2017, 30:6379-6390.
6. Yu C, Velu A, Vinitsky E, Gao J, Wang Y, Bayen A, Wu Y: The surprising effectiveness of PPO in cooperative multi-agent games. Advances in Neural Information Processing Systems. 2022, 35:24611-24624.
7. Chen YF, Liu M, Everett M, How JP: Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning. 2017 IEEE International Conference on Robotics and Automation. IEEE, Singapore; 2017. 285-292. 10.1109/ICRA.2017.7989037
8. Alshiekh M, Bloem R, Ehlers R, Konighofer B, Niekum S, Topcu U: Safe reinforcement learning via shielding. Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32:2669-2678. 10.1609/aaai.v32i1.11797
9. Garcia J, Fernandez F: A comprehensive survey on safe reinforcement learning. Journal of

- Machine Learning Research. 2015, 16:1437-1480.
10. van den Berg J, Guy SJ, Lin M, Manocha D: Reciprocal n-body collision avoidance. Robotics Research. Springer, Berlin; 2011. 3-19. 10.1007/978-3-642-19457-3_1
 11. Xu Z, Han X, Shen H, Jin H, Shimada K: NavRL: learning safe flight in dynamic environments. arXiv. 2024, arXiv:2409.15634. 10.48550/arXiv.2409.15634
 12. Shafer G: A Mathematical Theory of Evidence. Princeton University Press, Princeton; 1976.
 13. Dempster AP: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics. 1967, 38:325-339. 10.1214/aoms/1177698950
 14. Smets P, Kennes R: The transferable belief model. Artificial Intelligence. 1994, 66:191-234. 10.1016/0004-3702(94)90026-4
 15. Yan Z, Jaafar W, Selim B, Tabassum H: Multi-UAV speed control with collision avoidance and handover-aware cell association: DRL with action branching. arXiv. 2023, arXiv:2307.13158. 10.48550/arXiv.2307.13158
 16. Adadi A, Berrada M: Peeking inside the black-box: a survey on explainable artificial intelligence. IEEE Access. 2018, 6:52138-52160. 10.1109/ACCESS.2018.2870052
 17. Arrieta AB, Diaz-Rodriguez N, Del Ser J, et al.: Explainable artificial intelligence: concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020, 58:82-115. 10.1016/j.inffus.2019.12.012
 18. Shneiderman B: The eyes have it: a task by data type taxonomy for information visualizations. Proceedings of the IEEE Symposium on Visual Languages. IEEE, Boulder; 1996. 336-343. 10.1109/VL.1996.545307
 19. Floreano D, Wood RJ: Science, technology and the future of small autonomous drones. Nature. 2015, 521:460-466. 10.1038/nature14542
 20. Sutton RS, Barto AG: Reinforcement Learning: An Introduction, Second Edition. MIT Press, Cambridge; 2018.