Multi-Agent Reinforcement Learning for Coordinated Aerial Wildfire Suppression

Zahra Samadikhoshkho[1*], Michael Lipsett[1]
[1] Mechanical Engineering, University of Alberta, Edmonton, Canada.
Email: samadikh@ualberta.ca

INTRODUCTION

Wildfires present escalating global challenges, threatening communities, ecosystems, and infrastructure. Conventional aerial suppression relies heavily on manned aircraft, which face operational risks, high costs, and limited endurance. Swarm robotics offers a promising alternative for wildfire management [1]: fleets of unmanned aerial vehicles (UAVs) capable of continuous, distributed response. Coordinating such swarms, however, is non-trivial due to dynamic fire spread, uncertain wind conditions, and limited resources. Prior research has demonstrated UAV swarms for fire tracking [1], but suppression coordination using multi-agent reinforcement learning (MARL) remains underexplored. This paper presents a MARL framework for adaptive, decentralized coordination of UAV swarms in wildfire suppression, aiming to enhance containment efficiency and emergent cooperative behaviour.

MATERIALS AND METHODS

A simulation is developed to evaluate the use of multi-agent reinforcement learning (MARL) [2] for wildfire suppression. The wildfire environment uses a cellular automaton with wind and terrain effects to capture dynamic fire spread. Unmanned aerial vehicle (UAV) agents operate with simplified flight dynamics at fixed altitude, subject to payload and fuel constraints. Each UAV runs under a decentralized MARL framework trained with Proximal Policy Optimization (PPO), chosen for its robustness in continuous multi-agent domains. A shared global reward encourages fire containment efficiency and cooperative behaviour among agents. State inputs include fire front proximity, water level, wind direction, and relative positions of neighbouring UAVs. The action space consists of navigation, water release, reservoir refilling, and holding position. For benchmarking, MARL policies are compared against a rule-based baseline where UAVs target the nearest fire front. Figure 1 illustrates the UAV swarm environment and MARL decision loop.

RESULTS AND DISCUSSION

Simulations demonstrate that MARL-coordinated UAVs achieve up to 23% higher containment efficiency compared to rule-based methods. Fire containment improves from 62% under the rule-based baseline to 78% with 10 UAVs and 85% with 30 UAVs. At the same time, redundant drops decrease from 28% to 12%, and suppression time shortens from 40 minutes to 27 minutes. Emergent behaviours include task sharing, non-overlapping coverage, and adaptive redeployment, leading to more effective use of water resources and improved resilience under variable wind conditions. Notably, UAVs learn collision-free cooperation without explicit programming, demonstrating the strength of decentralized policy learning. Limitations arise from sensitivity to reward design and scalability issues beyond 50 UAVs, where coordination degrades. Such improvements suggest potential for reducing costs and risks in large-scale firefighting operations. These findings suggest that MARL provides adaptive, decentralized control for UAV swarms and represents a promising step beyond prior UAV wildfire tracking approaches [1].

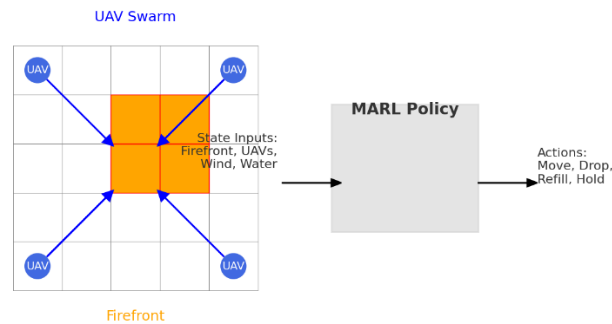

Fig 1 UAV swarm environment for wildfire suppression.

CONCLUSIONS

This study demonstrates the feasibility of using multi-agent reinforcement learning to coordinate UAV swarms for wildfire suppression. By enabling adaptive and cooperative decision-making, MARL consistently outperforms rule-based strategies in dynamic fire environments. Unlike prior UAV wildfire studies, this work shows that collision-free cooperation can emerge without explicit programming, underscoring the potential of decentralized learning approaches. Future extensions will focus on hierarchical MARL to address scalability, hardware-in-the-loop validation, and pathways toward safety certification. These findings highlight AI-driven UAV swarms as a practical pathway toward next-generation, autonomous wildfire management.

REFERENCES

[1] Pham, H. X., *et al.*, "A distributed control framework of multiple unmanned aerial vehicles for dynamic wildfire tracking," *IEEE Trans. Syst., Man, Cybern. Syst., vol. 50, no. 4, pp. 1537–1548, 2018.*

[2] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.