

## Autonomous Drone Control: A Reinforcement Learning Approach

**Auteur :** Hansen, Julien

**Promoteur(s) :** Ernst, Damien

**Faculté :** Faculté des Sciences appliquées

**Diplôme :** Master en ingénieur civil en informatique, à finalité spécialisée en "intelligent systems"

**Année académique :** 2024-2025

**URI/URL :** <http://hdl.handle.net/2268.2/23358>

---

### Avertissement à l'attention des usagers :

*Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.*

*Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.*

---

## Autonomous Drone Control : A Reinforcement Learning Approach

Drones have become an essential tools across a wide range of industries, from agriculture to surveillance, and are increasingly deployed in military contexts for detection, recognition, identification, exploration, and combat purposes. While most systems remain controlled by human, the shift toward autonomy is intensifying, driven by breakthroughs in artificial intelligence, notably in reinforcement learning and scalable simulation techniques.

This Master's thesis explores the potential of reinforcement learning for drone control within both single-agent and multi-agent frameworks. Two tasks are addressed : navigation in unknown terrains and adversarial drone combat. Our work focuses on designing simulation environments that model the learning process of agents as they interact with these tasks. Our navigation environment consists of multiple randomly spaced obstacles (spikes), a target, and a drone placed on opposite sides of the terrain. The drone is equipped with a sensor—either a LiDAR or a camera—which it uses to explore the environment and reach the target. In the adversarial scenario, the environment includes two drones : an attacker and a defender. The attacker attempts to reach a designated target, while the defender tries to intercept it by colliding with it.

Reinforcement learning is particularly well suited to these tasks due to its ability to learn complex, sequential decision-making policies from interaction with the environment. In scenarios such as drone navigation or combat, where the environment is often partially observable, highly dynamic, and difficult to model analytically, RL offers a flexible and data-driven approach to learning effective control strategies. Furthermore, Reinforcement learning naturally supports learning in multi-agent settings, where agents must coordinate or compete in real time.

To tackle these tasks, *policy gradient* methods such as *Proximal Policy Optimization*, its multi-agent extension *Independent Proximal Policy Optimization* and a variant inspired by self-play methods were explored. To train and evaluate our agents, IsaacLab environments were designed following the formalism of partially observable Markov decision process and stochastic games. Our work highlights the performance of trained agents regarding these tasks and show promising potential for future improvements regarding autonomous drone control.

**Key words :** Reinforcement Learning, drone, quadcopter, IsaacLab, IsaacSim, equilibrium, policy gradient method, self-play.

## Contrôle autonome de drones : Une approche par apprentissage par renforcement

Les drones sont devenus des outils essentiels dans de nombreux secteurs, de l'agriculture à la surveillance, et sont de plus en plus déployés dans des contextes militaires pour des missions de détection, reconnaissance, identification, exploration et combat. Bien que la plupart des systèmes restent contrôlés par des humains, la transition vers l'autonomie s'intensifie, portée par des avancées en intelligence artificielle, notamment dans l'apprentissage par renforcement et les techniques de simulation à grande échelle.

Ce mémoire de Master explore le potentiel de l'apprentissage par renforcement pour le contrôle de drones, dans des cadres à agent unique ainsi qu'à agents multiples. Deux tâches sont abordées : la navigation dans des terrains inconnus et le combat aérien adversarial. Notre travail se concentre sur la conception d'environnements de simulation modélisant le processus d'apprentissage des agents en interaction avec ces tâches. Notre environnement de navigation est composé de multiples obstacles (pics) placés aléatoirement, d'une cible et d'un drone positionnés de part et d'autre du terrain. Le drone est équipé d'un capteur, soit un LiDAR, soit une caméra qu'il utilise pour explorer l'environnement et atteindre la cible. Dans le scénario adversarial, l'environnement comprend deux drones : un attaquant et un défenseur. L'attaquant cherche à atteindre une cible désignée, tandis que le défenseur tente de l'intercepter en entrant en collision avec lui.

L'apprentissage par renforcement est particulièrement adapté à ces tâches grâce à sa capacité à apprendre des politiques complexes de prise de décision séquentielle via l'interaction avec l'environnement. Dans des scénarios tels que la navigation ou le combat de drones, où l'environnement est souvent partiellement observable, très dynamique et difficile à modéliser analytiquement, le RL offre une approche flexible et basée sur les données pour apprendre des stratégies de contrôle efficaces. De plus, le RL supporte naturellement l'apprentissage en contextes multi-agents, où les agents doivent coopérer ou s'affronter en temps réel.

Pour relever ces défis, des méthodes dites de *policy-gradient*, telles que *Proximal Policy Optimization*, son extension à plusieurs agents *Independent Proximal Policy Optimization*, ainsi qu'une variante inspirée du paradigme *self-play*, ont été explorées. Pour entraîner et évaluer nos agents, des environnements IsaacLab ont été conçus en suivant le formalisme des processus de décision markoviens et des jeux stochastiques. Notre travail met en lumière les performances des agents entraînés sur ces tâches et révèle un potentiel prometteur pour de futures améliorations en matière de contrôle autonome de drones.