

الجمهورية الجزائرية الديمقراطية الشعبية
PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
وزارة التعليم العالي والبحث العلمي
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
جامعة عمّار ثليجي بالأغواط
UNIVERSITY OF LAGHOuat



FACULTY OF SCIENCES
DEPARTMENT OF COMPUTER SCIENCE

Field : Mathematics and Computer Science
Option : Computer Science
Specialization : Systems and Computer Science Networks

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE MASTER DEGREE IN COMPUTER SCIENCE

Submitted by: Affaf Khadroune & Safa Mahboubi

THEME

**A review of Multi-Agent Reinforcement Learning for
multi-UAV**

Examination Committee:

<i>Mr</i>	Mohamed Lahcen Bensaad	MC(B)	(University of Laghouat)	President
<i>Ms</i>	Zohra Abdelhafidi	MC(A)	(University of Laghouat)	Examiner
<i>Mr</i>	Tahar Bendouma	MC(A)	(University of Laghouat)	Supervisor

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة عمار ثلجي بالأغواط
UNIVERSITÉ AMAR TELIDJI LAGHOUAT



FACULTE DES SCIENCES
DEPARTEMENT D'INFORMATIQUE

Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes et Réseaux Informatiques

Mémoire de Master

Présenté par : Affaf Khadroune & Safa Mahboubi

THÈME

**Revue sur : Multi-Agent Reinforcement Learning
pour multi-UAV**

Soutenu publiquement devant le jury composé de :

Mr	Mohamed Lahcen Bensaad	MC(B)	(Université de Laghouat)	Président
M ^{lle}	Zohra Abdelhafidi	MC(A)	(Université de Laghouat)	Examineur
Mr	Tahar Bendouma	MC(A)	(Université de Laghouat)	Encadrant

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

DÉDICACE

Je dédie ce modeste travail :

À mon papa chéri, école de mon enfance, qui a été mon ombre durant toutes ces années d'études, et qui a veillé tout au long de ma vie à m'encourager, à m'aider et à me protéger.

À celle qui m'a donné la vie, qui s'est sacrée pour mon bonheur et ma réussite, à ma petite maman chérie que j'aime énormément.

À mes soeurs :Nour Elhouda, Amoula ,Meriem.

À mes deux fleurs de bonheur :Ghadir et Khadidja.

À mon cher frère : Mohamed.

À mon encadrant Mr Tahar Bendouma pour son excellent encadrement et ses conseils judicieux.

À mon très chère oncle Touati,votre affection et votre soutien m'ont été d'un grand secours au long de ma vie.

À mon grand-père qui nous a quitté en silence, tu me manques énormément, que le bon dieu t'accueille dans son vaste paradis.

À toute ma famille : Khadrone et Benboubakar.

À tous ceux qui m'ont aidé spécialement Omar, Saida, Rahil, Idris, Soraya, Yousra, Khadija et Ibrahim.

À Hadjer, merci pour ton amour. Tu es toujours ma meilleure.

Affaf

DÉDICACE

Je dédie ce travail à l'esprit de ma nounou et chère grand-mère à mon cœur, El hadja Zohra, qu'Allah ait pitié de son esprit pur.

Aussi je dédie ce travail à mon père mon trésor et ma mère ma vie, pour leur patience, leur amour, leur soutien et leurs encouragements à faire mieux.

Mes chers frères et sœur qui ont toujours été une source constante de soutien et d'encouragement tout au long de mon parcours universitaire. Pour toute ma famille votre soutien a été une fierté pour moi. A tous ceux que j'aime et ceux qui m'aiment.

Safa

REMERCIEMENT

Avant tout, Nous remercions ALLAH pour nous avoir donné la force et la volonté de faire ce modeste travail.

Ainsi , Nous tenons à exprimer notre sincère gratitude à notre superviseur M.Tahar Bendouma, pour son soutien ,disponibilité et ses précieux conseils qui ont été très utile pour l'achèvement de ce projet.

Bien sur,Nous tenons à exprimer notre sincère gratitude aux membres du jury, Dr Zohra Abdelhafidi et Dr Lahcen Bensaad, qui ont accepté d'évaluer et assisté à cette soutenance.

Enfin, nous ne saurons terminer ces remerciements sans de n'associer toute personne qui, de près ou de loin, nous a apporté son aide ou sa sympathie.

Safa & Affaf ,Juin 2022

ABSTRACT

In recent years, single-agent [Reinforcement Learning \(RL\)](#) has achieved tremendous success in solving various sequential decision-making problems in machine learning. Some of these problems require more than one agent to resolve, so the notion [Multi-agent Reinforcement Learning \(MARL\)](#) has come to light. [Unmanned Aerial Vehicle \(UAV\)](#) technology became an attractive way to undertake several challenging applications, especially when [Unmanned Aerial Vehicles \(UAVs\)](#) form a swarm . One of the areas that sees great use of [MARL](#) is multi-UAV systems. In this work, we reviewed a number of articles that used these two concepts together [MARL](#) and swarm of [UAVs](#). The objective is to minimize the search part for researchers who are more interested and educated in this field .

Keywords : [UAVs](#), [MARL](#), [RL](#), swarm of UAVs, [Machine learning \(ML\)](#), [Artificial intelligence \(AI\)](#).

RÉSUMÉ

Ces dernières années, l'apprentissage par renforcement [RL](#) pour un seul agent a enregistré un énorme succès dans la résolution de divers problèmes de prise de décision séquentielle dans l'apprentissage automatique. Certains de ces problèmes nécessitent pour leurs résolution plus qu'un agent, pour cela la notion [MARL](#) est venue à la lumière. La technologie des véhicules aériens sans pilote [UAV](#) devient un moyen intéressant d'entreprendre plusieurs applications difficiles, en particulier lorsque les [UAVs](#) forment un essaim (swarm of UAVs). L'un des domaines qui connaissent une grande utilisation en [MARL](#) c'est les systèmes multi-UAVs. Dans ce travail, nous avons passé en revue un certain nombre d'articles qui utilisaient ces deux concepts ensemble [MARL](#) et essaim des [UAVs](#) dont l'objectif est de minimiser la partie de recherche pour les chercheurs qui sont intéressés plus et fait ses études dans ce domaine .

Mots clés : [UAVs](#), [MARL](#), [RL](#), essaim des drones , [ML](#), [L'intelligence artificielle \(IA\)](#).

ملخص

مؤخراً، حقق التعلم المعزز بوكيل واحد نجاحاً هائلاً في حل مشكلات اتخاذ القرار المتسلسلة المختلفة في التعلم الآلي. تتطلب بعض هذه المشكلات أكثر من وكيل واحد لحلها ، لذلك ظهرت فكرة التعلم المعزز متعدد الوكلاء. أصبحت تكنولوجيا الطائرات بدون طيار وسيلة جذابة للقيام بالعديد من التطبيقات الصعبة لا سيما عندما تشكل الطائرات بدون طيار سرباً. أحد المجالات التي تشهد استخداماً كبيراً ل التعلم المعزز متعدد الوكلاء هي هاته الأنظمة. في هذا العمل، قمنا بدراسة مجموعة من المقالات. تناولت المفهومين مع بعض. هدفنا هو تقليص دائرة البحث للعديد من الباحثين، ايضاً لتسهيل و توجيه الباحثين المهتمين الذين هم بصدد القيام بأبحاث في هذا المجال الجديد.

الكلمات المفتاحية: التعلم المعزز بوكيل واحد، التعلم المعزز متعدد الوكلاء، طائرات بدون طيار، سرب طائرات بدون طيار.

TABLE DES MATIÈRES

1	Introduction	1
1.1	Contexte	1
1.2	Problématique et motivation	2
1.3	Organisation du cette mémoire	2
2	Reinforcement Learning : Background	4
2.1	Introduction	5
2.2	L'intelligence artificielle	5
2.2.1	Domaines de l'intelligence artificielle	5
2.2.2	Les avantages et les inconvénients de l'intelligence artificielle	6
2.3	Machine Learning	7
2.3.1	Définition	7
2.3.2	Classification	8
2.3.3	Supervised Learning	9
2.3.4	Unsupervised Learning	12
2.4	Reinforcement Learning	14
2.4.1	Le processus de décision de Markov	14
2.4.2	La structure du Reinforcement Learning	15
2.4.3	Les domaines du Reinforcement Learning	17

2.5	La différence entre les trois types d'apprentissages	18
2.6	La classification des algorithmes du Reinforcement Learning .	19
2.6.1	Model-Free RL	21
2.6.2	Model-Based RL	23
2.7	Les principaux défis du Reinforcement Learning	24
2.8	Conclusion	25
3	Unmanned Aerial Vehicles (UAVs)	26
3.1	Introduction	27
3.2	Unmanned Aerial Vehicles	27
3.3	Composants des UAVs	27
3.4	Classification	29
3.5	Communication	31
3.5.1	Communication UAV à UAV :	32
3.5.2	Communication UAV-sol :	32
3.5.3	Communication par satellite :	33
3.6	Système multi-UAVs (swarm)	33
3.6.1	Architectures de communication multi-UAVs	34
3.7	Domaines d'applications	36
3.7.1	Collecte de données et détection	36
3.7.2	Vidéo et surveillance	37
3.7.3	Opérations de recherche et d'action	38
3.7.4	Couverture de zone (Area coverage)	38
3.7.5	Précision agriculture	39
3.7.6	Livraison (Delivery)	39
3.7.7	Aide à la connexion	40
3.7.8	Edge / fog / cloud computing	41
3.7.9	Missions spéciales	43
3.8	Les défis des UAV's	45
3.9	Conclusion	46

4	Multi-agent Reinforcement Learning	47
4.1	Introduction	48
4.2	Définition	48
4.3	Structure du Multi-agent Reinforcement Learning	48
4.4	L'interaction entre les agents	50
4.4.1	Coopération	50
4.4.2	Compétition	51
4.4.3	Hybride	51
4.5	Défis de Multi-agent Reinforcement Learning	51
4.6	Applications du MARL	52
4.7	Multi-agent Reinforcement Learning pour multi-UAVs	53
4.8	Conclusion	69
5	Conclusion et perspectives futures	70
5.1	Conclusion	70
5.2	Perspectives futures	71
	Bibliographie	72

TABLE DES FIGURES

2.1	Les types de Machine Learning	8
2.2	La structure du Supervised Learning	10
2.3	Exemple du Supervised Learning [1]	11
2.4	Fonctionnement de l'Unsupervised Learning	13
2.5	La relation entre l'intelligence artificielle, Machine Learning et Reinforcement Learning	14
2.6	La Structure du Reinforcement Learning	16
2.7	Les étapes d'apprentissages d'un agent	17
2.8	Les applications du Reinforcement Learning	18
2.9	Une comparaison entre les trois types d'apprentissages [2] . . .	19
2.10	Une hiérarchie des algorithmes du Reinforcement Learning . .	20
2.11	Les trois méthodes du Model-Free RL	21
2.12	Fonctionnement du méthode actor critic	23
2.13	Model-Based RL stratégie	24
3.1	Principaux composants d'UAV	28
3.2	Exemples de styles de corps d'UAV	30
3.3	Différentes méthodes pour les communications UAVs	32
3.4	Les architectures de communication	35
3.5	les domaines d'application des UAVs	36
3.6	Utilisation de systèmes UAV dans les opérations de livraison .	40

3.7	Utilisez UAV pour fournir une couverture sans fil aux utilisateurs extérieurs	41
3.8	Utilisez UAV pour fournir une couverture sans fil aux utilisateurs intérieurs	41
3.9	edge / fog /cloud présentation	42
3.10	l'utilisation des UAVs dans la pandémie de covid-19	45
4.1	La structure du Multi-agent Reinforcement Learning	50
4.2	Les applications du MARL	52

LISTE DES TABLEAUX

2.1	Les domaines de l'intelligence artificielle [3]	6
4.1	Classification (1) selon le domaine aide à la connexion	54
4.2	Classification (2) selon le domaine aide à la connexion	55
4.3	Classification (3) selon le domaine aide à la connexion	56
4.4	Classification (4) selon le domaine aide à la connexion	57
4.5	Classification (1) selon le domaine Fog /Edge /Cloud Computing	58
4.6	Classification (2) selon le domaine Fog /Edge /Cloud Computing	59
4.7	Classification (3) selon le domaine Fog /Edge /Cloud Computing	60
4.8	Classification (1) selon le domaine collection des données et détection	61
4.9	Classification (2) selon le domaine collection des données et détection	62
4.10	Classification (1) selon le domaine vidéo & surveillance	63
4.11	Classification (2) selon le domaine vidéo & surveillance	64
4.12	Classification (1) selon le domaine missions spéciales	65
4.13	Classification (2) selon le domaine missions spéciales	66
4.14	Classification (1) selon le domaine recherche et opération	67

LISTE DES ABRÉVIATIONS

A2C Advantage Actor Critic. [53](#), [62](#)

A3C Asynchronous Advantage Actor Critic. [53](#)

AC Actor Critic. [22](#), [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#)

AI Artificial intelligence. [v](#)

BS Base Station. [30](#)

C Concurrent. [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#)

DDPG Deep Deterministic Policy Gradient. [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#)

DPG Deterministic Policy Gradient. [22](#), [53](#)

DQN Deep Q-Network. [21](#), [56](#), [63](#), [65](#), [66](#), [67](#)

GBS Ground Base Station. [31](#), [32](#), [33](#)

GCS Ground Control Station. [29](#)

H Hybride. [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#)

HAUs High Altitude UAVs. [29](#)

- IA** L'intelligence artificielle. [vi](#), [1](#), [2](#), [5](#), [6](#), [41](#)
- IOT** Internet Of Things. [41](#)
- LAUs** Low Altitude UAVs. [30](#)
- LoS** Line-of-Sight. [32](#), [33](#)
- MARL** Multi-agent Reinforcement Learning. [v](#), [vi](#), [2](#), [48](#), [49](#), [50](#), [51](#), [52](#), [69](#), [70](#), [71](#)
- MAUs** Medium Altitude UAVs. [30](#)
- MDP** Markov Decision Process. [14](#)
- MEC** Mobile Edge Computing. [42](#)
- ML** Machine learning. [v](#), [vi](#), [1](#), [2](#), [8](#), [70](#), [71](#)
- P** Coopératif. [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#)
- PB** Policy-Based. [53](#), [55](#)
- PPO** Proximal Policy Optimization. [22](#), [55](#), [57](#), [63](#)
- Q-Learning** Quality Learning. [21](#), [53](#), [54](#), [55](#), [56](#), [57](#), [60](#), [61](#), [62](#), [63](#), [66](#)
- RL** Reinforcement Learning. [v](#), [vi](#), [1](#), [2](#), [14](#), [15](#), [16](#), [17](#), [19](#), [21](#), [24](#), [48](#), [49](#), [51](#), [53](#), [54](#), [55](#), [56](#), [57](#), [58](#), [59](#), [60](#), [61](#), [62](#), [63](#), [64](#), [65](#), [66](#), [67](#), [71](#)
- SAC** Soft Actor Critic. [56](#), [63](#)
- SAR** Search and rescu. [40](#)
- SARSA** State-Action-Reward-State-Action. [53](#), [54](#)
- SATCOM** Satellite Communication. [31](#), [33](#), [53](#), [55](#), [56](#), [59](#), [61](#), [65](#), [66](#)
- TRPO** Trust Region Policy Optimization. [22](#), [57](#)

U2G UAV-to-Ground. 31, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67

U2U UAV-to-UAV. 31, 32, 53, 54, 55, 58, 59, 60, 61, 64, 65, 66, 67

UAS Unmanned Aerial System. 27

UAV Unmanned Aerial Vehicle. v, vi, 1, 2, 27, 28, 29, 31, 33, 34, 35, 36, 38, 39, 41, 42, 43, 44, 45, 46, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67

UAVs Unmanned Aerial Vehicles. v, vi, 1, 2, 27, 28, 29, 32, 33, 36, 37, 38, 41, 42, 43, 44, 46, 48, 53, 69, 70, 71

VB Value-Based. 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 66, 67

CHAPITRE 1

INTRODUCTION

Sommaire

1.1	Contexte	1
1.2	Problématique et motivation	2
1.3	Organisation du cette mémoire	2

1.1 Contexte

Aujourd’hui le développement technologique dans le monde a touché tous les domaines qui servent à faciliter la vie quotidienne de l’être humain. Nous mentionnons notamment la nouvelle technologie des véhicules aériennes sans pilote (**UAVs**) qui ont attiré une large attention de la communauté industrielle, scientifique et universitaire en tant que nouveau paradigme en raison de leurs diverses caractéristiques et de leur large éventail de domaines d’applications notamment militaires et civiles. Récemment , l’intelligence artificielle (**IA**), le Machine Learning (**ML**) et le Reinforcement Learning (**RL**) sont présents dans tous les secteurs et à tous les niveaux. Ce nouveau concept **UAV** est utilisé largement avec le Machine Learning précisément avec Reinforcement Learning en raison de sa capacité à résoudre un large éventail de tâche en utilisant une architecture simple et sans avoir besoin de connaissances préalables sur la dynamique du problème à résoudre.

1.2 Problématique et motivation

D'une part et à mesure que la technologie des (UAVs) se développe et que leur coût diminue, ils deviennent un moyen intéressant d'entreprendre plusieurs applications difficiles, en particulier lorsque les UAVs forment un essaim (swarm of UAVs).

D'autre part, malgré le succès de l'approche du Reinforcement Learning pour un seul agent (RL), un grand nombre de problèmes du monde réel ne peuvent être entièrement résolus par un seul agent alors que la solution de ce problème était le système multi-agents. D'ici a apparu la notion Multi-agent Reinforcement Learning (MARL).

Il existe un grand nombre d'articles qui sont implémentés MARL sur un swarm d'UAVs dans divers domaines. Tandis que ce thème est nouveau dans la communauté scientifique un nombreux chercheurs doivent avoir des problèmes ou des obstacles lors ses recherches pour cela dans notre travail on a essayé de minimiser quelques difficultés dans cette phase par une proposition d'une revue qui organise et classifie les travaux récents sur le Multi-agent Reinforcement Learning pour multi-UAV ,selon plusieurs critères.

1.3 Organisation du cette mémoire

Ce document contient cinq chapitre, qui ont de contenus :
Tout d'abord, le premier chapitre on le considère comme une introduction générale pour cette mémoire. Le deuxième chapitre présente une petite introduction sur l'[IA](#), et les concepts de base du [ML](#) avec les trois types d'apprentissage. Ensuite, on a détailler plus l' apprentissage par renforcement [RL](#), on a discuter de leurs classifications, applications et les défis. En plus dans le troisième chapitre on a définie l'[UAV](#), classer et citer ses domaines d'applications. On a aussi discuter les essaims des [UAVs](#) et ses défis.
Pour le quatrième chapitre, il contient deux parties. L'une présente la structure du [MARL](#), ses composants et ses défis.

L'autre partie contient des tableaux qui résument les critères principales de 61 articles.

Enfin, nous terminons le travail par une conclusion et des perspectives futures.

CHAPITRE 2

REINFORCEMENT LEARNING : BACKGROUND

Sommaire

2.1	Introduction	5
2.2	L'intelligence artificielle	5
2.2.1	Domaines de l'intelligence artificielle	5
2.2.2	Les avantages et les inconvénients de l'intelligence artificielle	6
2.3	Machine Learning	7
2.3.1	Définition	7
2.3.2	Classification	8
2.3.3	Supervised Learning	9
2.3.4	Unsupervised Learning	12
2.4	Reinforcement Learning	14
2.4.1	Le processus de décision de Markov	14
2.4.2	La structure du Reinforcement Learning	15
2.4.3	Les domaines du Reinforcement Learning	17
2.5	La différence entre les trois types d'apprentissages	18
2.6	La classification des algorithmes du Reinforcement Learning	19
2.6.1	Model-Free RL	21
2.6.2	Model-Based RL	23
2.7	Les principaux défis du Reinforcement Learning	24
2.8	Conclusion	25

2.1 Introduction

L'intelligence artificielle (IA) et le Machine Learning (ML) sont deux concepts visant à améliorer les compétences d'un système informatique.

A cet effet, dans ce chapitre nous allons présenter une petite introduction sur l'intelligence artificielle puis les concepts de base du Machine Learning et ses types d'apprentissages. Ensuite, on va détailler le Reinforcement Learning pour un seul agent, ses applications et leurs défis.

2.2 L'intelligence artificielle

IA a été inventée par John McCarthy lors de la célèbre conférence de Dartmouth en été 1956. Cette conférence a été considérée comme le point de départ de **IA**. L'**IA** est l'une des innovations technologiques qui sont produites, pour remplacer le travail manuel qui est fait par l'homme dans divers domaines.

On la considère comme une branche de la science et technologie qui crée des machines intelligentes et des programmes informatiques pour effectuer diverses tâches qui nécessitent une intelligence humaine et imite des fonctions qu'un être humain peut faire. L'**IA** a changé la société d'une manière que personne n'aurait pu prévoir dont l'objectif est de simuler l'intelligence humaine à l'aide d'ordinateurs.

2.2.1 Domaines de l'intelligence artificielle

L'intelligence artificielle se manifeste sous plusieurs formes dans divers domaines on les résume dans le tableau suivant [2.1](#)

TABLE 2.1 : Les domaines de l'intelligence artificielle [3]

L'intelligence Artificielle (IA)	Applications Cognitives	Système expert Système d'apprentissage Logic Floue Algorithmes des Génétiques Réseaux de Neurones Les agents intelligents
	Applications Robotiques	Perception Visuelle Tactilité Dextérité Locomotion Navigation
	Applications des Interfaces Naturelles	Langage Naturelle Reconnaissance de parole Interfaces multi-sensoriel La réalité virtuelle

2.2.2 Les avantages et les inconvénients de l'intelligence artificielle

1. **Avantages** : Il existe plusieurs pour l'IA, parmi ces avantages on peut citer :

- La tâche terminée plus vite qu'un être humain avec un taux de réussite élevé.
- Le travail stressant et complexe peut réaliser facilement, en peu de temps, avec la possibilité d'effectuer à la fois diverses fonctions.
- Augmenter l'efficacité par la réduction des erreurs et les défauts dans le travail.
- Calcul de situations longues et complexes et découvrez des choses inexplorées.

2. **Inconvénients** : Certains des principaux inconvénients de l'intelligence artificielle dans notre vie quotidienne sont les suivants :

- Parfois, il peut être mal utilisé entraînant une catastrophe à grande échelle.
- Emplois humains touchés qui augmente le problème du chômage et réduit les contacts humains.
- La créativité dépend du programmeur.
- L'utilisation ou le choix d'une technique adéquate pour un domaine précis nécessite une expérience .

2.3 Machine Learning

2.3.1 Définition

L'apprentissage automatique est une branche de l'intelligence artificielle capable d'apprendre des modèles à partir de données brutes ou d'apprendre à partir d'un échantillonnage d'observation de l'environnement, permettant aux systèmes informatiques d'acquérir des connaissances. De nombreuses études ont été faites sur la façon de faire apprendre les machines elles-mêmes sans être explicitement programmées.

Il s'agit d'une science moderne permettant de découvrir des patterns et d'effectuer des prédictions à partir des données en se basant sur des statistiques, sur la fouille de données, sur la reconnaissance de patterns et sur les analyses prédictives.

2.3.2 Classification

Les types de Machine Learning peuvent être classés en trois catégories principales qui sont basé sur différents algorithmes pour résoudre des problèmes. Le type d'algorithme utilisé dépend du type de problème que vous souhaitez résoudre, du nombre de variables, du type de modèle qui lui conviendrait le mieux. La figure suivante montre une hiérarchie du [ML 2.1](#) .

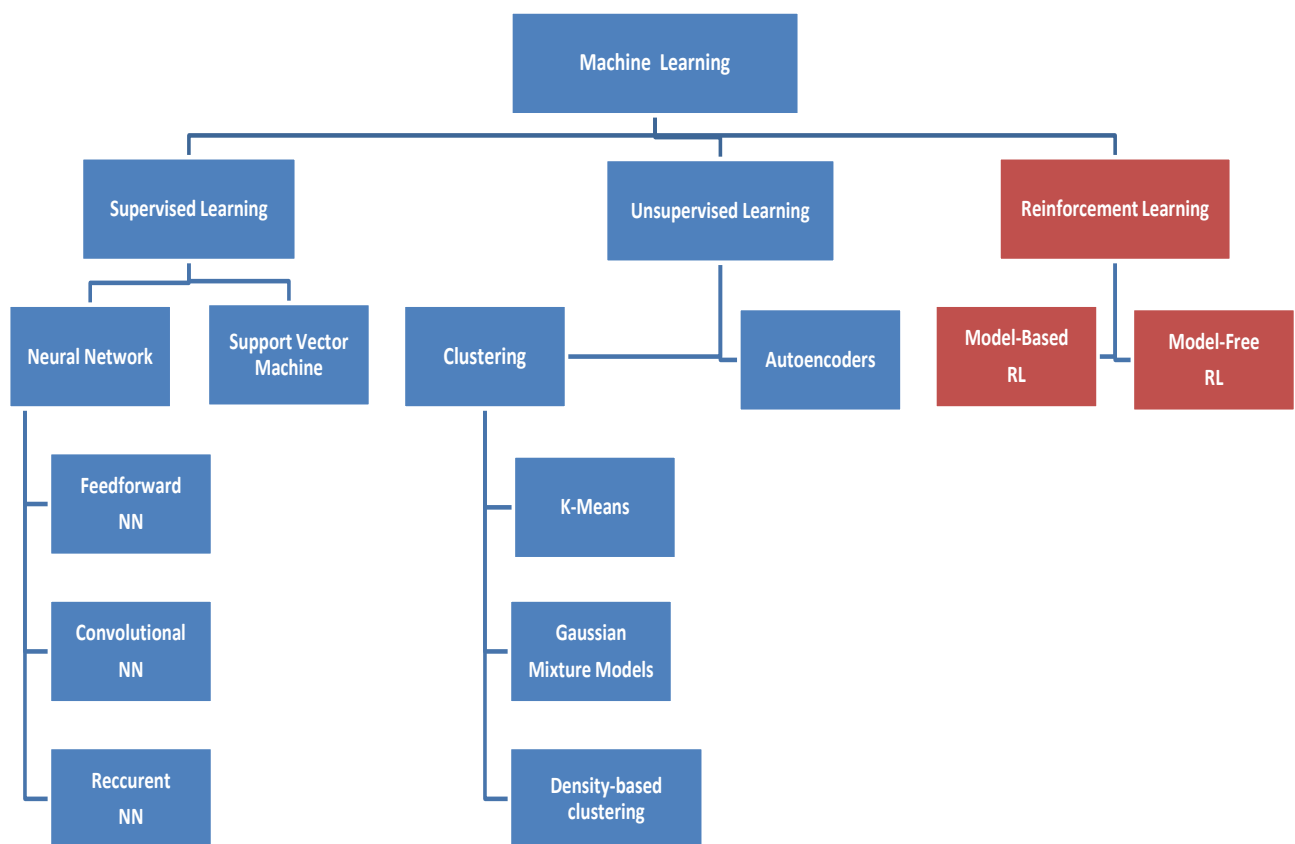


FIGURE 2.1 : Les types de Machine Learning

Ces trois catégories sont :

2.3.3 Supervised Learning

Dans l'apprentissage supervisé, les données fournies sont étiquetées, en d'autres termes, nous fournissons pour chaque entrée de données leur étiquette réelle afin que l'algorithme utilise ces valeurs pour apprendre à mettre une décision pour une nouvelle entrée non étiquetée. Le but principal de cette technique est de trouver une fonction qui manipule les entrées et les sorties s'appelle un modèle de prédiction. Les données d'entrée peuvent être des valeurs numériques, alphanumériques, des images. Les données de sortie (variable à prédire) peuvent être l'une des deux catégories suivantes :

- **Variable discrète** : est une variable qui peut prendre uniquement certaines valeurs d'un intervalle de nombres réels. Généralement, les valeurs admissibles ne sont que les nombres entiers [4]. Exemple : L'âge d'une personne arrondi à l'année ne peut admettre que des nombres entiers comme valeur.
- **variable continue** : est une variable qui peut prendre toutes les valeurs possibles d'un intervalle de nombres réels [4]. Par exemple, prédire le prix d'un drone à partir de ses caractéristiques. Dans cet exemple, vous devez fournir à l'algorithme un ensemble de données d'entraînement contenant les caractéristiques de chaque drone et son étiquette associée (le prix).

L'ensemble de données est généralement divisé en deux ensembles, ensemble d'apprentissage "**Training Set**" et un ensemble de test "**Test Set**". Tel que l'ensemble d'apprentissage est utilisé pour apprendre la relation entre l'entrée et la sortie et l'ensemble de test est utilisé pour valider le modèle en mesurant sa précision.

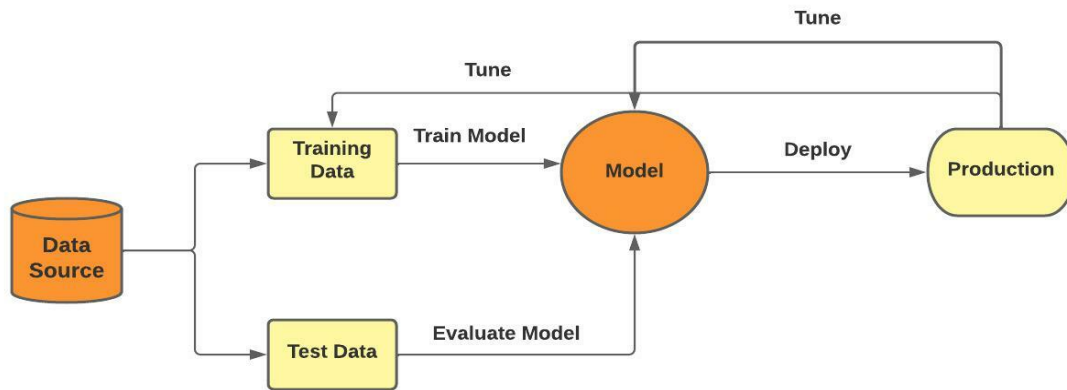


FIGURE 2.2 : La structure du Supervised Learning

L'objectif du Supervised Learning : L'apprentissage supervisé principalement est utilisé pour deux objectifs :

- **Classification :** Quand la variable à prédire prend une valeur discrète, on parle d'un problème de classification. Quand l'ensemble des classes dépasse deux classes, on parle de classification multi-classes. Exemple : classer les maladies cancéreuses bénignes ou malignes
- **Régression :** Quand la variable à prédire est une valeur spécifique, on parle d'un problème de régression, un algorithme de régression permet de trouver un modèle (une fonction mathématique) en fonction des données d'entraînement (données d'entrées). Le modèle calculé permettra de donner une estimation sur une nouvelle donnée non encore vue par l'algorithme (qui ne faisait pas partie des données d'entraînement). Les algorithmes de régression peuvent prendre plusieurs formes en fonction du modèle qu'on souhaite construire. La régression linéaire est le modèle le plus simple, il consiste à trouver la meilleure droite qui s'approche le plus de données d'entrée. Parmi les algorithmes de régression, on retrouve : régression linéaire, régression polynomiale, régression crête... ex : prévoir un prix

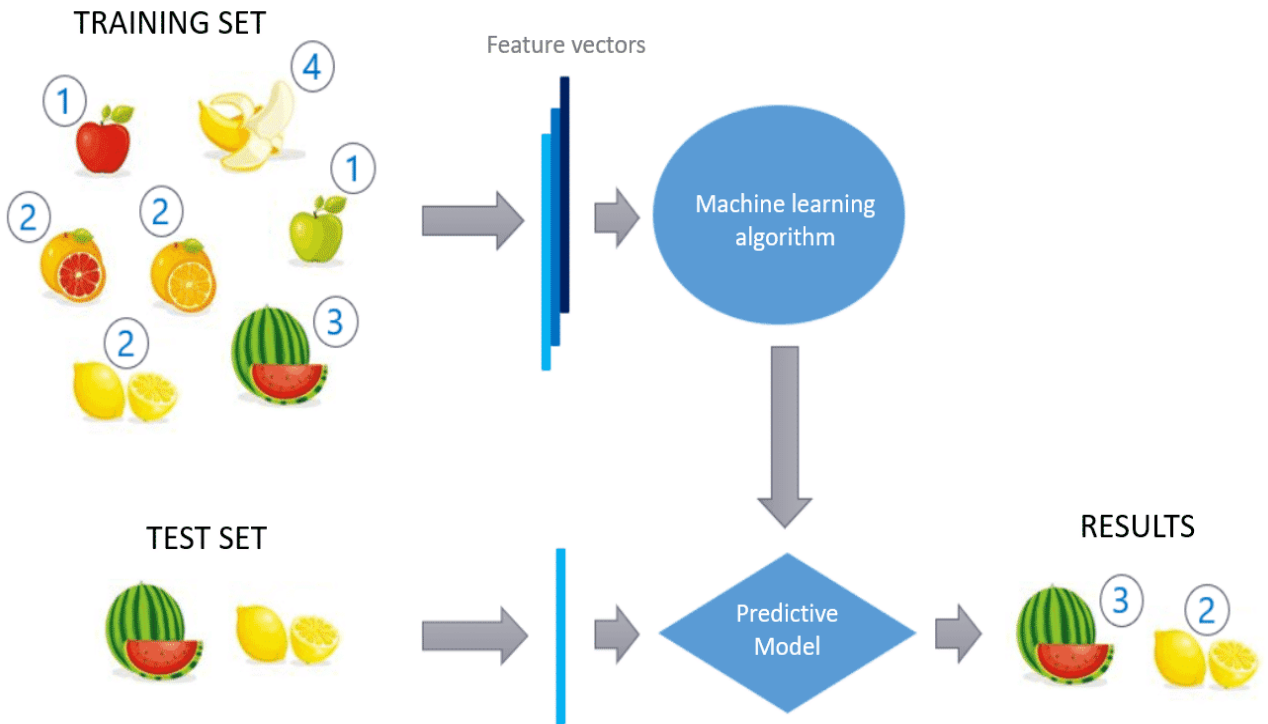


FIGURE 2.3 : Exemple du Supervised Learning [1]

2.3.4 Unsupervised Learning

L'apprentissage non supervisé consiste à ne disposer que des données d'entrée qui sont **non étiquetés** et pas de variables de sortie correspondantes. L'apprentissage par la machine se fait de manière entièrement indépendante.

L'objectif d'Unsupervised Learning : L'apprentissage non supervisé principalement est utilisé pour :

- **Regroupement (Clustering)** : La mise en groupe (cluster) consiste à séparer ou à diviser un ensemble de données dans un certain nombre de groupes, fonction de regroupement base sur grouper les données en fonction de leurs similitudes. Parmi les algorithmes de regroupement, on trouve : décalage moyen, DBSCAN, K-Means.
- **Recherche de motifs** : Il est aussi connu comme reconnaissance de formes, cette technique est définie comme un processus pour identifier les régularités dans les données. Ce processus peut se faire physiquement, mathématiquement ou par l'utilisation d'algorithmes. La reconnaissance des formes est largement utilisée dans les domaines techniques comme la vision par ordinateur, la reconnaissance vocale, la reconnaissance faciale, etc.
- **Réduction de dimension** : Cette méthode permettant de projeter des données d'un espace de grande dimension dans un espace de plus petite dimension. La réduction de dimension permet de réduire la complexité d'un problème à plusieurs niveaux :
 - **Au niveau théorique** : Cela entraîne automatiquement une amélioration de propriété de stabilité et de robustesse des algorithmes.

- **Au niveau pratique :** Cela permet de limiter le nombre de possibilités à tester, en réduisant l'espace des solutions, ce qui permet de traiter les données plus rapidement.

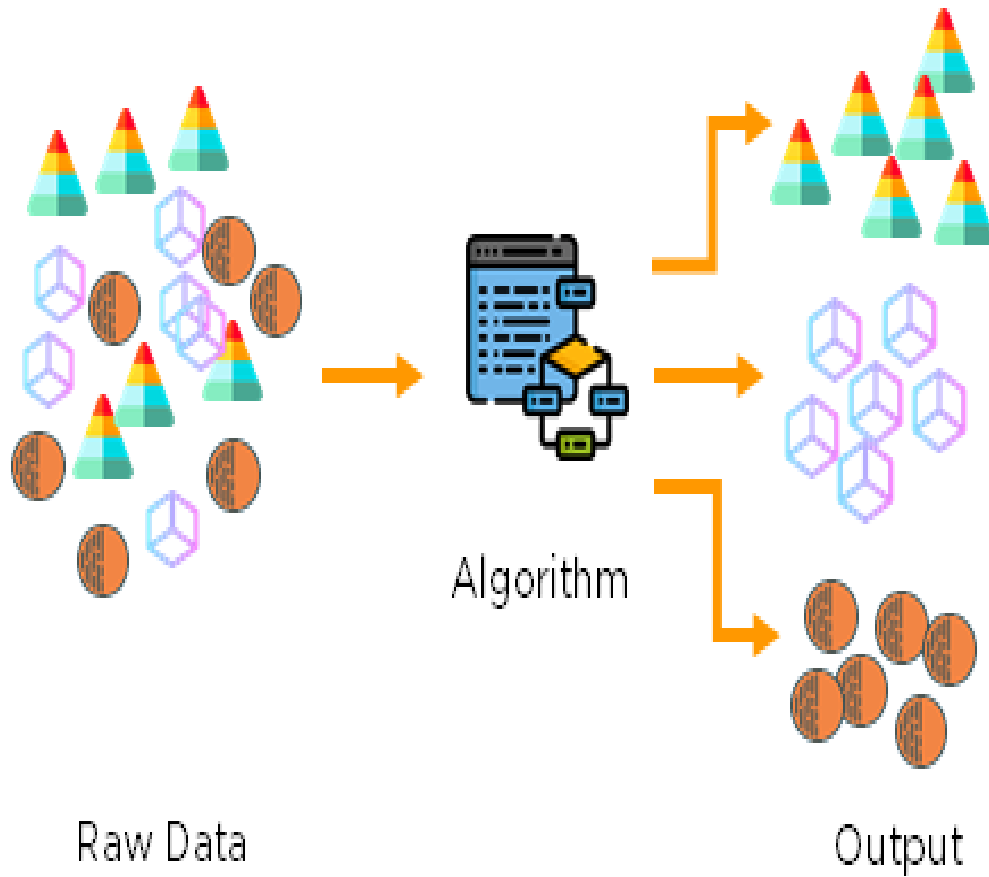


FIGURE 2.4 : Fonctionnement de l'Unsupervised Learning

2.4 Reinforcement Learning

Contrairement aux deux autres cadres d'apprentissages, qui fonctionnent à l'aide d'un ensemble de données statiques, **RL** fonctionne avec des données provenant d'un environnement dynamique. **RL** n'a pas de but de regrouper les données ou d'étiqueter les données, mais de trouver la meilleure séquence d'actions qui génèrent le résultat optimal. **RL** résout ce problème en permettant à un composant appelé agent **d'explorer, d'interagir** avec et **apprendre** de l'environnement. Le problème est généralement formulé comme un processus de décision de Markov (MDP).

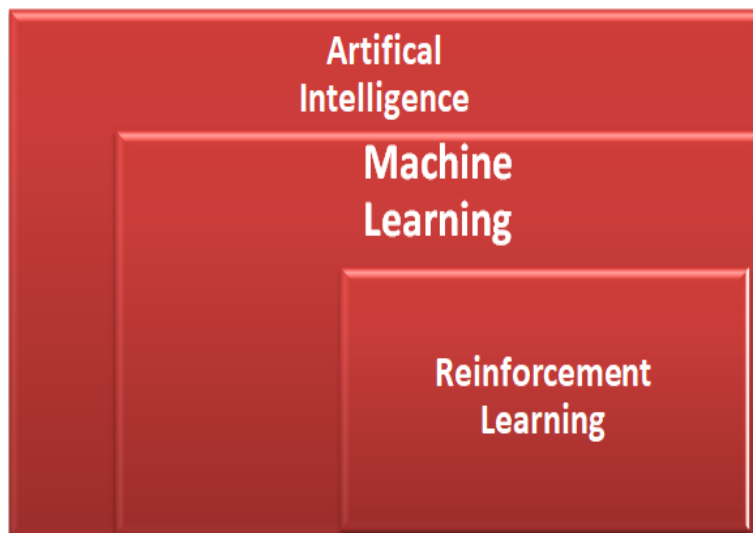


FIGURE 2.5 : La relation entre l'intelligence artificielle, Machine Learning et Reinforcement Learning

2.4.1 Le processus de décision de Markov

Markov Decision Process (MDP) est défini par un tuple (S, A, P, R, γ) où S et A désignent les ensembles d'états et d'actions respectivement ; $P : S \times A \rightarrow \Delta(S)$ désigne la transition Probabilité de tout état $s \in S$ à tout état $\acute{s} \in S$ pour toute action donnée $a \in A$; $R : S \times A \times S \rightarrow R$ est la fonction de récompense qui détermine la récompense immédiate reçue par l'agent pour une transition de (s, a) à $\acute{s}$; $\gamma \in [0, 1]$ est le facteur d'actualisation

qui compense l'instantané et futures récompenses.

2.4.2 La structure du Reinforcement Learning

Pour une meilleure compréhension de fonctionnement du RL. On définit d'abord les éléments du RL puis ses étapes.

- Les éléments du RL : un Renforcement problème a toujours cinq éléments principaux comme montre la Figure 2.6 :

1. **L'agent** : Une entité qui peut effectuer une action et reçoit une récompense R_t en conséquence.
2. **L'environnement** : Une représentation du monde réel dans lequel l'agent exploite les observations.
3. **La politique(Policy)** : C'est l'application pour chaque état S_t à une action A_t . prend en compte les observations d'état (les entrées) et les associe aux actions (les sorties). Cette fonction décide l'action à entreprendre selon l'ensemble d'observations.
4. **La récompense (Reward)** : C'est une valeur qui peut être positive ou négative, l'agent reçoit la récompense après avoir effectué une action. On désigne généralement La récompense par R_t .
5. **La fonction de valeur (value-function)** : Elle représente la qualité d'un état, il s'agit donc du total des récompenses futures attendues à partir d'un état donné.

Une fonction de valeur est généralement notée $V(s)$ où s est l'état qui nous intéresse. Mathématiquement, c'est formulé comme suit : $V(s) = E(G_t)$, où G_t est la somme actualisée des récompenses futures $G_t = \sum \gamma^{t-1} R_t$ où $\gamma \in [0, 1]$

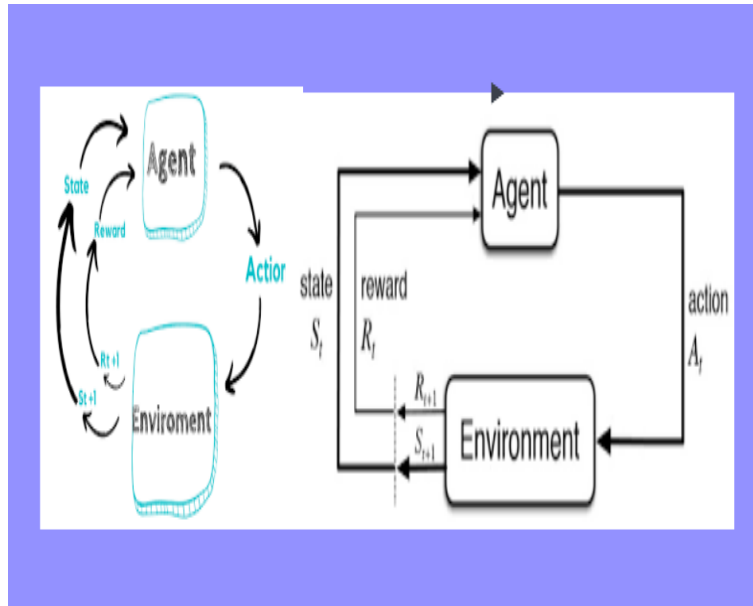


FIGURE 2.6 : La Structure du Reinforcement Learning

- **Les étapes du RL** : Pour atteindre la solution optimale, l'agent doit appliquer les étapes suivantes dans l'environnement :
 1. L'agent est capable d'observer l'état actuel de l'environnement, à partir de l'état observé, il décide l'action à entreprendre. Il observe les effets de ses actions, cette étape est appelée **l'exploration**.
 2. Après ces étapes si la performance de l'agent est correct il reçoit une récompense et si le contraire il reçoit une pénalité. L'environnement change d'état et produit une récompense pour cette action. Les deux sont reçus par l'agent. C'est l'étape de **l'exploitation**.
 3. Il améliore ses actions futures, c'est l'étape d'apprentissage. Le cycle observation-action-récompense continue jusqu'à ce que l'apprentissage soit terminé.
 4. Grâce à la récompense l'agent peut déterminer si cette action a été bonne et doit être répétée, ou si elle est mauvaise et doit être évitée.

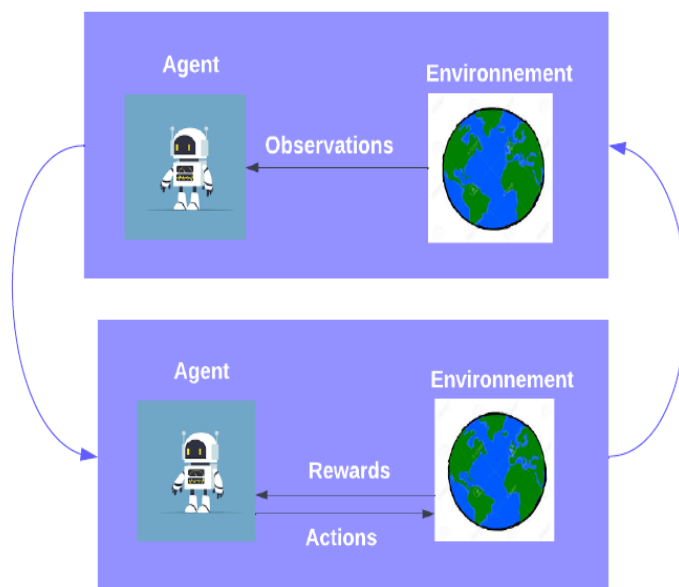


FIGURE 2.7 : Les étapes d'apprentissages d'un agent

2.4.3 Les domaines du Reinforcement Learning

Le Reinforcement Learning est largement connu pour ses exploits en jeux vidéos, il existe une multitude de domaines d'application auquel le RL est utilisé.

- **La robotique** : En robotique et en automatisation industrielle, le Reinforcement Learning est utilisé pour permettre au robot de créer un système de contrôle adaptatif efficace pour lui-même qui apprend de sa propre expérience et de son comportement [2].
- **La fouille de texte** : Les domaines d'applications du Reinforcement Learning comprennent également la fouille de texte ou Text mining [2].
- **La finance** : Si les grandes entreprises du secteur financier utilisent déjà les algorithmes de Machine Learning pour améliorer leurs activités de trading et d'investissement en capitaux propres, certaines d'entre elles ont déjà jeté leur dévolu sur le Reinforcement learning [2].

- **La santé** : Des articles récents suggèrent de multiples applications du Reinforcement Learning dans le secteur des soins de santé. On peut citer le dosage des médicaments [2].

En plus de ces domaines il existe d'autre comme transport, éducation, économie, énergie, gestion d'entreprise, business, chimie , contrôle, ingénierie scientifique de l'art, systèmes informatiques et la fabrication.



FIGURE 2.8 : Les applications du Reinforcement Learning

2.5 La différence entre les trois types d'apprentissages

on résume la différence dans le tableau suivant :

	APPRENTISSAGE SUPERVISÉ	APPRENTISSAGE NON-SUPERVISÉ	APPRENTISSAGE PAR RENFORCEMENT
DÉFINITION	L'algorithme apprend à partir de données labellisées	L'algorithme est entraîné à partir de données non labellisées sans indications particulières	L'algorithme interagit avec son environnement en réalisant des actions et en apprenant de ses erreurs et succès
TYPE DE PROBLÈMES	Régression et classification	Association et Clustering	Basés sur un système de récompense
TYPE DE DONNÉES	Données labellisées	Données non labellisées	Pas de données fournies au préalable
APPROCHE	Étudie les relations sous-jacentes qui lient les données en entrée aux labels	Découvre les motifs communs au sein des données d'entrée	Apprend une stratégie de comportement en fonction d'expériences passées et des récompenses perçues

FIGURE 2.9 : Une comparaison entre les trois types d'apprentissages [2]

Enfin, il est nécessaire de souligner qu'il existe d'autres types d'apprentissages tels que l'apprentissage semi-supervisé qui consiste en un "mix" entre apprentissage supervisé et apprentissage non supervisé.

2.6 La classification des algorithmes du Reinforcement Learning

Les algorithmes du [RL](#) peuvent être divisés en deux principales catégories qui sont subdivisées en sous-catégories on montre l'hierarchie dans la figure [2.10](#) :

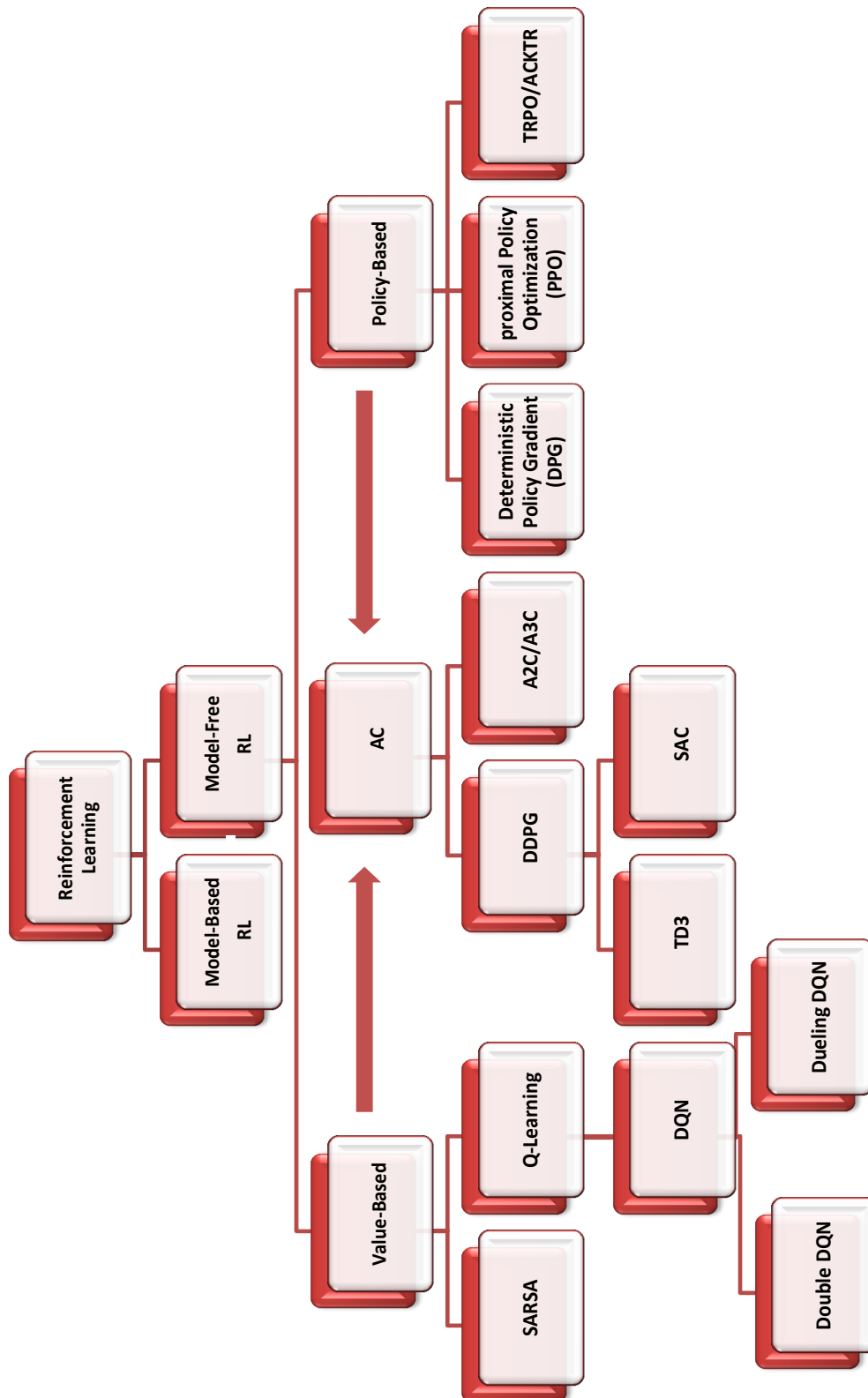


FIGURE 2.10 : Une hiérarchie des algorithmes du Reinforcement Learning

2.6.1 Model-Free RL

L'une des raisons pour lesquelles Reinforcement Learning est puissant c'est que l'agent n'a pas besoin de savoir leur environnement précédemment. Il reste encore apprendre à interagir avec lui. Avec [RL](#) sans modèle, vous pouvez placer un agent équipé de RL dans n'importe quel système et l'agent pourra apprendre la politique optimale alors ici les actions sont basées sur la stratégie des "essais et erreurs".

Pour ce modèle il existe trois classifications des méthodes comme montre la figure 2.11 :

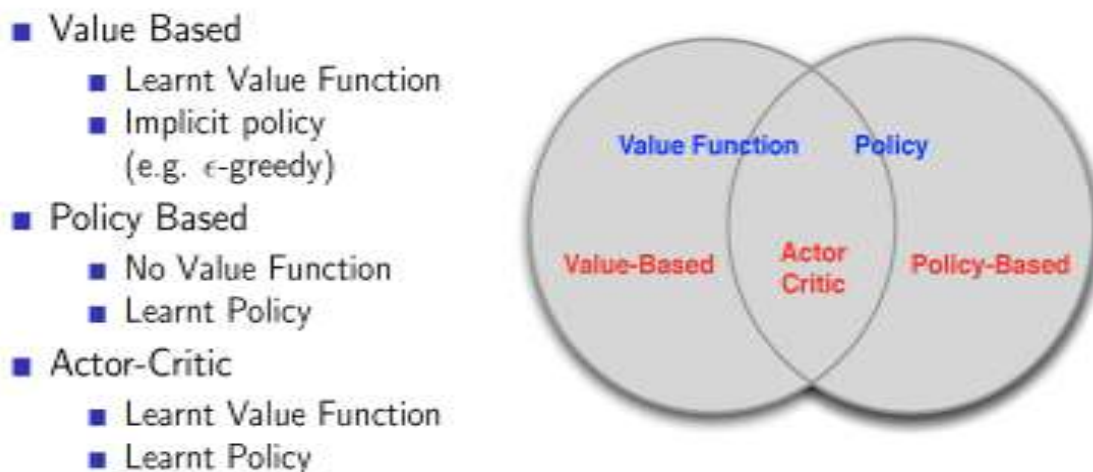


FIGURE 2.11 : Les trois méthodes du Model-Free RL

1. **Value-Based** : Des méthodes impliquent généralement l'optimisation de la fonction de valeur $V(s)$. Les algorithmes basés sur la valeur incluent le [Quality Learning \(Q-Learning\)](#) ,[Deep Q-Network \(DQN\)](#), etc.
 - Les avantages de cette méthode résident dans l'échantillon d'efficacité est élevée.

- Les inconvénients sont qu'il ne peut généralement pas gérer le problème de l'espace d'action continue et la variance de l'estimation de la fonction de valeur est faible et ce n'est pas tombé facilement dans l'optimum local.

2. **Policy-Based** : La méthode optimise directement la politique ,il met à jour la politique de manière itérative jusqu'à ce que le rendement cumules soit maximisé. Certains algorithmes courants basés sur des politiques incluent [Deterministic Policy Gradient \(DPG\)](#) ,[Trust Region Policy Optimization \(TRPO\)](#) et [Proximal Policy Optimization \(PPO\)](#) .

- Par rapport à la méthode basé sur la valeur, une méthode basée sur des politiques présentent les avantages d'un paramétrage plus simple des politiques, d'une meilleure convergence et convient à une action continue ou à grande dimension espace.
- Parmi ses inconvénients : Convergent généralement vers un optimum local plutôt que global et l'évaluation d'une politique est généralement inefficace et à forte variance.

3. **Actor Critic** : Les méthodes du [Actor Critic \(AC\)](#) sont une hybridation entre les deux autres types Value-based et Policy-based tel que :
 $AC = Value - based(Actor) + Policy - based(Critic)$.

Il a deux types de réseaux : Acteur et Critique. L'acteur a décidé l'action à entreprendre et le critique l'informe de la qualité de l'action et de la manière dont elle doit s'adapter.

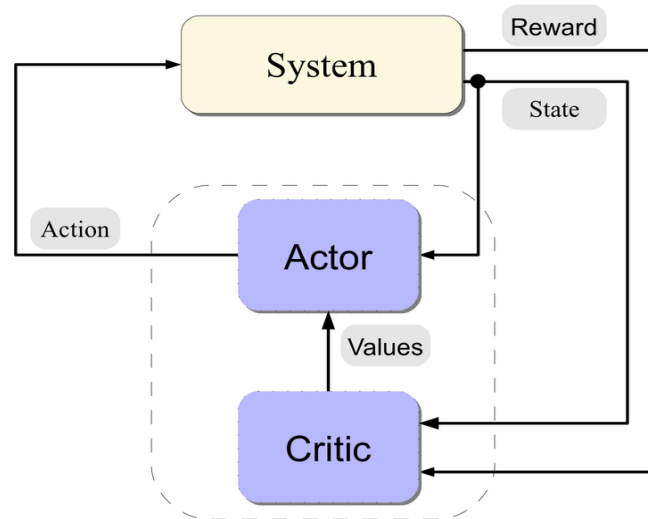


FIGURE 2.12 : Fonctionnement du méthode actor critic

2.6.2 Model-Based RL

Au contrairement au premier modèle **Model-Free**, si l'agent n'a aucune compréhension de l'environnement, il doit explorer toutes les zones de l'espace d'état pour déterminer comment collecter le plus de récompenses. Cela signifie que l'agent devra passer du temps à explorer les zones à faible récompense pendant le processus d'apprentissage. Cependant, vous connaissez peut-être certaines parties de l'espace d'état qui ne valent pas la peine d'être explorées. En fournissant un modèle de l'environnement, ou d'une partie de l'environnement, vous fournissez les agents avec cette connaissance. À l'aide d'un modèle, l'agent peut explorer des parties de l'environnement sans avoir à prendre physiquement cette action. Un modèle accomplir le processus d'apprentissage en évitant les domaines connus pour être mauvais et explorer le reste.

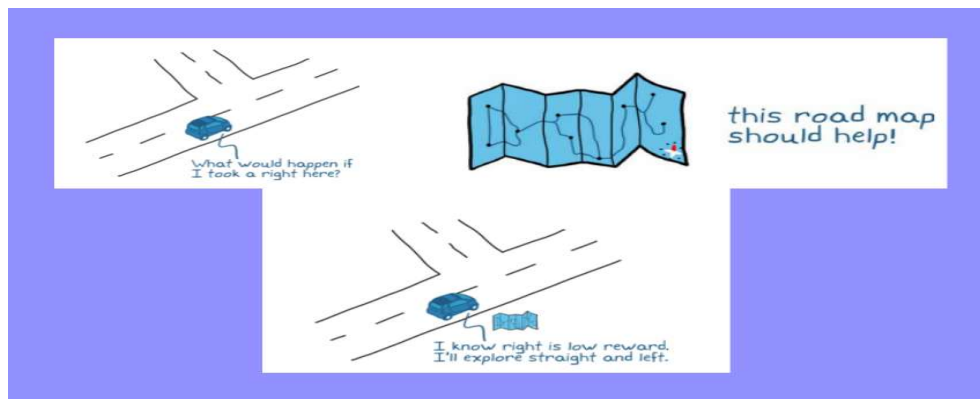


FIGURE 2.13 : Model-Based RL stratégie

2.7 Les principaux défis du Reinforcement Learning

Parmi les défis du **RL** on cite les deux principaux :

1. **Le dilemme entre l'exploration et l'exploitation** : Lorsque l'agent est plongé dans un environnement inconnu, il doit concilier entre :
 - La nécessité d'explorer afin d'acquérir une meilleure connaissance de son environnement.
 - L'exploitation de ses connaissances pour maximiser ses gains.Cet équilibre entre exploitation et exploration représente un enjeu essentiel dans la phase d'apprentissage [2].
2. **Choisir l'algorithme approprié et en régler les paramètres** : Le choix de l'algorithme le plus approprié représente un défi important étant donné la grande variété d'approches existantes [2].

2.8 Conclusion

Ce chapitre a présenté les concepts de base du Machine Learning avec les trois types d'apprentissages , les méthodes de chaque type , leurs domaines d'application et la différence entre eux . Le prochain chapitre contient une présentation des drones (UAVs) , leurs catégories , les types de communications et les domaines d'application de ce composant.

CHAPITRE 3

UNMANNED AERIAL VEHICLES (UAVS)

Sommaire

3.1	Introduction	27
3.2	Unmanned Aerial Vehicles	27
3.3	Composants des UAVs	27
3.4	Classification	29
3.5	Communication	31
3.5.1	Communication UAV à UAV :	32
3.5.2	Communication UAV-sol :	32
3.5.3	Communication par satellite :	33
3.6	Système multi-UAVs (swarm)	33
3.6.1	Architectures de communication multi-UAVs	34
3.7	Domaines d'applications	36
3.7.1	Collecte de données et détection	36
3.7.2	Vidéo et surveillance	37
3.7.3	Opérations de recherche et d'action	38
3.7.4	Couverture de zone (Area coverage)	38
3.7.5	Précision agriculture	39
3.7.6	Livraison (Delivery)	39
3.7.7	Aide à la connexion	40
3.7.8	Edge / fog / cloud computing	41
3.7.9	Missions spéciales	43
3.8	Les défis des UAV's	45
3.9	Conclusion	46

3.1 Introduction

En raison de leurs améliorations impressionnantes et de leur utilité dans les différents domaines d'applications, les véhicules aériens sans pilote **UAV** ou les drones sont devenus un sujet brûlant impliquant la technologie.

Au cours de ce chapitre, nous présentons de manière générale ce qu'est un **UAV**, les types de **UAVs** et leurs classifications puis nous présentons leur architecture de communication et nous terminons par les missions et les applications des drones.

3.2 Unmanned Aerial Vehicles

Les **UAVs** sont partie principale des systèmes aériens sans pilote **Unmanned Aerial System (UAS)**, le terme **UAS** englobe non seulement le drone, mais également la personne ou l'équipe au sol contrôlant le vol, ainsi que le système qui relie les deux (GPS, contrôle au sol, systèmes de transmissions, caméras, logiciels, etc.) [5]

Un drone est essentiellement un robot volant [6]. Il est un aéronef sans pilote qui ne nécessite aucune intervention humaine directe pour voler, il peut naviguer de manière autonome selon un plan de vol pré-programmé ou peut être contrôlé à distance [7]. Pour améliorer les performances des drones sur leurs missions, des recherches sont menées pour rendre les drones coopératifs de manière contrôlée par une station de contrôle ou de manière autonome.

3.3 Composants des UAVs

La plupart des **UAVs** sont constitués de composants modulaires qui s'adaptent sur et à l'intérieur du corps des **UAVs** et sont alimentés par une source d'énergie. Selon le style de vol et les missions à effectuer.

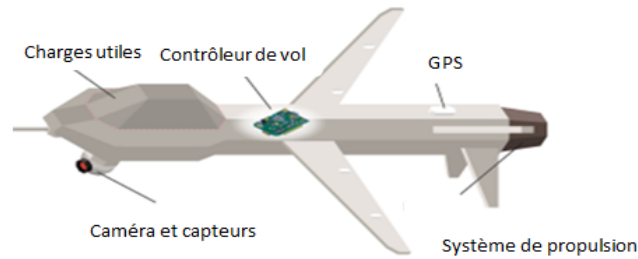


FIGURE 3.1 : Principaux composants d'UAV

1. **Capteurs** : Les drones utilisent une variété de capteurs qui fournissent au contrôleur de vol des données telles que la télémétrie, les mesures d'attitude, le terrain et les obstacles. Les capteurs font également partie intégrante de la connaissance de la situation des UAVs.
2. **Contrôleur de vol** : Le contrôleur de vol est le cerveau du drone. Il est le matériel et le micrologiciel qui traitent toutes les informations collectées du GPS et d'autres capteurs. Le contrôleur de vol contrôle également la charge utile sur l'UAV.
3. **Système de propulsion** : Les UAVs peuvent utiliser une variété des systèmes de propulsion en fonction de la taille , du type et de la mission de l'UAV.

Ce composé de :

Moteurs (rotors) : Ils influent sur la charge maximale supportée et sur le temps de vol.

Hélices : Elles produisent de la poussée.

Contrôleurs de vitesses électroniques : ils gèrent la vitesse, la direction et la rotation des moteurs.

Une batterie : Elle fournit l'énergie nécessaire aux autres composants électriques .

4. **Charge d'utile** : Ce drones varient de quelques dizaines de grammes à des centaines de kilogrammes [8]. Plus la charge utile est importante, plus

d'équipements et d'accessoires peuvent être transportés au détriment d'une taille de drone plus grande, d'une capacité de batterie plus élevée et d'une durée de vol plus courte. Les charges utiles typiques comprennent des caméras vidéo et toutes sortes de capteurs, qui pourraient être utilisés à des fins de reconnaissance, de surveillance et à des fins commerciales [9].

5. **Système de communication** : Un opérateur d'UAV peut communiquer avec l'UAV en utilisant l'un des éléments suivants : émetteur de contrôleur radio, station de contrôle au sol **Ground Control Station (GCS)** ou liaisons de données duplex par satellite [10].

3.4 Classification

Il existe de nombreuses classifications pour les drones basées sur une variété de paramètres ou de caractéristiques, telles que **la taille, le poids, la vitesse, l'altitude et l'endurance**. Il n'y a pas de classification standard spécifique pour toutes les applications et aucune classification n'est meilleure qu'une autre.

On peut classer UAVs selon leur :

Taille : La taille est un critère crucial lors la désignation de l'UAV le plus adapté à une mission spécial. Pour cela il est utilisé pour faire la distinction entre les UAVs. De plus, leur taille peut varier des micro et mini drones ($> 2kg$ et $< 20kg$) aux gros drones (pesant plus de 600 kg) [11].

Altitude et endurance : Il existe également trois types d'UAV, chacun avec son propre ensemble de fonctionnalités :

Les **High Altitude UAVs (HAUs)** longue endurance à haute altitude. sont pratiquement stationnaires à une altitude supérieures à 20 kilomètres comme les satellites et plus de 24 h d'endurance. Ils effectuent des reconnaissances et des surveillances à très longue portée. Ils sont généralement exploités par les forces aériennes à partir de bases fixes .

Les [Medium Altitude UAVs \(MAUs\)](#), comme les avions, voyagent à des altitudes moyennes jusqu'à 11 km et peuvent se déplacer plus rapidement du point de vue des nœuds terrestres.

Les [Low Altitude UAVs \(LAUs\)](#), comme les drones, peuvent voler pendant plusieurs heures à des hauteurs allant de quelques dizaines de mètres à quelques kilomètres, se déplacer rapidement et sont adaptables.

Mécanisme de vol : Les drones peuvent être classés en trois types :

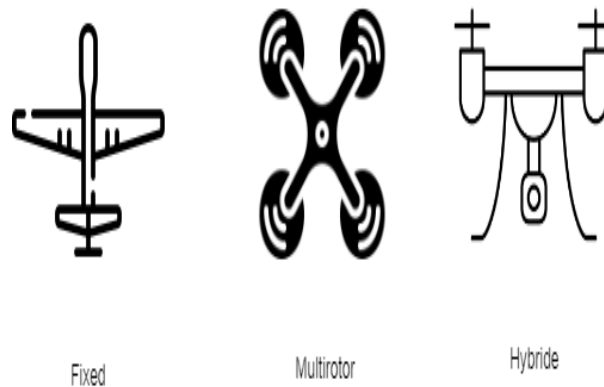


FIGURE 3.2 : Exemples de styles de corps d'UAV

a. Les drones multirotors (également appelés drones à voilure tournante rotary-wings drones) permettent un décollage et un atterrissage verticaux et peuvent survoler un emplacement fixe pour fournir une couverture cellulaire continue dans certaines zones. Cette grande maniabilité les rend aptes à assister les communications cellulaires, car ils peuvent déployer des [Base Station \(BS\)](#) aux emplacements souhaités avec une grande précision, ou voler dans une trajectoire désignée tout en transportant des [BS](#).

b. Les drones à voilure fixe (Fixed-wing) peuvent planer dans les airs, ce qui les rend beaucoup plus économes en énergie et capables

de transporter une lourde charge utile. Le vol à voile aide également les drones à voilure fixe à se déplacer à une vitesse plus rapide. Les inconvénients des drones à voilure fixe sont que (i) ils nécessitent une piste pour décoller et atterrir car le décollage et l'atterrissage verticaux ne sont pas possibles [12], et (ii) ils ne peuvent pas survoler un emplacement fixe. (iii) Les drones à voilure fixe sont également plus chers que les drones multi-rotors.

c. Les drones hybrides (fixed/rotary wing) à voilure fixe/tournante sont récemment arrivés sur le marché pour offrir un compromis entre les deux types de drones, qui peut décoller verticalement, atteindre rapidement sa destination en glissant dans les airs, puis passer en vol stationnaire à l'aide de quatre rotors.

3.5 Communication

Chaque nœud du réseau UAV, Ground Base Station (GBS) et satellites peut fonctionner comme un système final. Cependant, la communication entre deux nœuds distants est soumise à diverses limitations, notamment les pertes de paquets et la fragmentation du réseau. En conséquence, trois formes de communication émergent selon le niveau :

- UAV-to-UAV (U2U) : UAV à UAV
- UAV-to-Ground (U2G) : UAV à sol
- Satellite Communication (SATCOM) : UAV à satellite

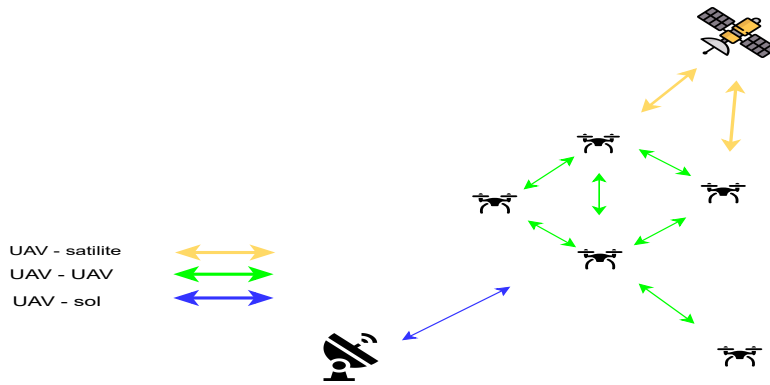


FIGURE 3.3 : Différentes méthodes pour les communications UAVs

3.5.1 Communication UAV à UAV :

Afin de répondre aux besoins des différentes missions, les UAVs communiquent directement en échangeant fréquemment des paquets de données entre eux.

Cependant, en raison des restrictions sur les portées de transmission, la communication multi-sauts est effectuée sur d'autres UAV. Ceci est crucial pour étendre la couverture d'un domaine d'intérêt spécifique. Dans la majorité des cas, la ligne de visée **Line-of-Sight (LoS)** est prédominante dans les communications U2U puisqu'il n'existe aucune obstruction entre les UAVs dans le ciel [13].

3.5.2 Communication UAV-sol :

Pour un meilleur contrôle des drones volants, des infrastructures sous forme de GBS sont fixées au sol afin d'échanger des messages critiques de

contrôle et de commande et même les données.

- Fixe : emplacements statiques et fixes en raison d'une mobilité limitée comme les capteurs, les MeNodeB, serveurs MEC.
- Mobile : emplacements dynamiques et mobiles comme les véhicules, les appareils mobile et les GBS.

3.5.3 Communication par satellite :

les drones sont souvent déployés dans des environnements complexes, où il est difficile d'installer des GBS ou lorsqu'un groupe de drones nécessite une connectivité continue et que le réseau est sévèrement partitionné. À cette fin, il est nécessaire de disposer d'une entité centralisée assurant une connectivité permanente, comme l'utilisation de satellites comme une option adéquate pour servir de relais contrôlant les UAVs de manière centralisée et fournissant également une couverture LoS importante, établissant ainsi Satellite Communication SATCOM [14].

3.6 Système multi-UAVs (swarm)

Un système multi-UAVs est un groupe de drones intelligents, conçus pour atteindre le même objectif dans un environnement complexe et de fournir aux utilisateurs au sol un taux de couverture élevé et d'assurer une couverture équitable entre les utilisateurs d'une manière économe en énergie. Et de les faire fonctionner en groupes de manière coordonnée et collaborative (swarm) ou concurrent pour former un système multi-UAV.

Cela qui donne une capacité d'essaimer a un grand nombre d'UAV pour accomplir des tâches compliquées devient attrayante car elle répond aux limites des systèmes d'UAV uniques, telles que la charge utile et la durée du vol. Il présente également d'autres caractéristiques et avantages, tels que des gains de temps, une réduction de la main-d'œuvre et des coûts d'exploitation réduits. En plus peuvent échanger du travail entre eux.

D'autre part si tombe quelque drone ou un matériel en panne dans un système de drone, les autres continuent leurs missions.

3.6.1 Architectures de communication multi-UAVs

L'architecture de communication joue un rôle important dans le contrôle intelligent et la collaboration autonome des essaims d'UAV. Il existe deux types d'architectures de communication dans un swarm :

Architecture de communication centralisée

Ce type d'architecture de communication centralisée est relativement stable, adopte des algorithmes de routage plus simples et a une échelle plus petite. Lorsque la taille de l'essaim de drones et la zone de couverture de la mission sont petits et la mission relativement simple, ce type d'architecture est plus adapté [15]. Comme montre la partie a de la figure 3.4.

Architecture de communication décentralisée

Ce type d'architecture de communication décentralisée est fonctionnent à grande vitesse et que les missions peuvent couvrir de vastes zones, les réseaux ad hoc en essaim d'UAV sont considérés comme le meilleur choix dans cette architecture [16].

1. Réseau ad hoc UAV : Dans cette méthode la communication entre l'essaim et l'infrastructure est une liaison à point unique s'appuyant sur un drone spécifique appelé passerelle UAV. En ce qui concerne le drone passerelle, deux types d'émetteurs-récepteurs supplémentaires sont nécessaires : l'un pour communiquer avec d'autres drones à faible puissance et à courte distance et l'autre pour communiquer avec l'infrastructure à haute puissance et longue portée. Comme montre la partie b de la figure 3.4

2. Réseau Ad-hoc UAV Multi-Groupe : Cette architecture est intégrée à la fois une « architecture centralisée » et une architecture « réseau ad hoc UAV ». Les communications intergroupes, c'est-à-dire les communications de groupe à groupe, sont effectuées via l'infrastructure, de sorte que les UAV passerelles sont toujours responsables de la communication avec l'infrastructure. Une architecture de «réseau ad hoc en essaim multi-groupes» peut être utilisée lorsque les scénarios de mission nécessitent une variété d'UAV. Il faut cependant faire attention à la robustesse de cette architecture, car l'architecture « multi-group swarm Ad hoc network » est semi centralisée. Comme montre la partie c de la figure 3.4
3. Réseau Ad-hoc UAV Multi-Couche : Ce type d'architecture de réseau ad hoc en essaim multi-couche est approprié pour les scénarios où les missions d'essaim d'UAV sont compliquées et significativement avancé par rapport à l'architecture "multi-group swarm Ad hoc network". Un groupe de drones adjacents du même type constitue un réseau Ad hoc, qui est la couche inférieure de l'architecture de communication. Le drone passerelle le plus proche communique avec l'infrastructure (sol), qui est la couche supérieure de l'architecture. Comme montre la partie d de la figure 3.4

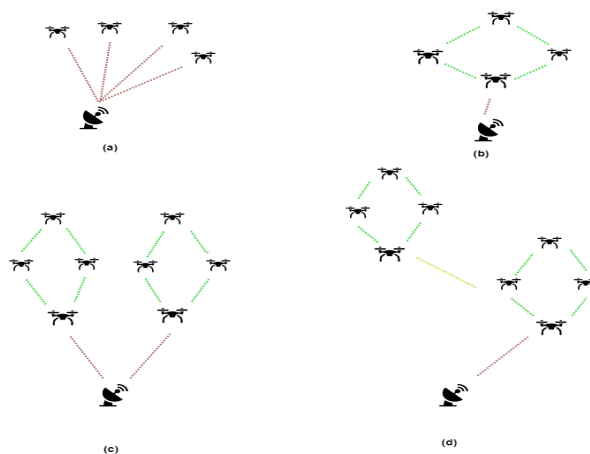


FIGURE 3.4 : Les architectures de communication

3.7 Domaines d'applications

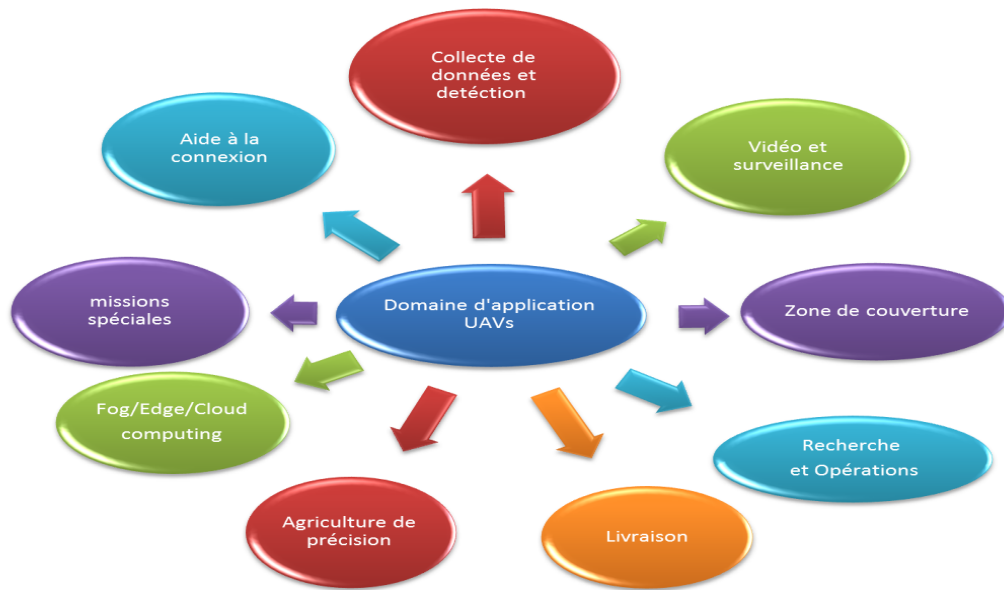


FIGURE 3.5 : les domaines d'application des UAVs

3.7.1 Collecte de données et détection

Un UAV est joué le rôle de capteur mobile qui peut collecter des informations et acquérir des données et des calculs importantes à partir de ressources et d'entités externes à l'aide de ses nombreux capteurs. Les nœuds de capteurs au sol (GSN), les voitures, les appareils Internet des objets (IoT) ou tout autre élément connecté au cloud sont des exemples de ressources externes.

1. **Détection** : (remote sensing) Les drones peuvent être utilisés pour collecter des données à partir de capteurs au sol et transmettre les données collectées aux stations de base au sol [17]. Les drones équipés de capteurs peuvent également être utilisés comme réseau de capteurs aériens pour la surveillance de l'environnement et la gestion des catastrophes [18].
2. **Inspection des infrastructures et exploitation minière** (Infrastructure inspection and mining) Ces UAVs sont utilisés pour inspecter des objets qui sont hors de portée ou des zones qui pourraient être

risquées et dangereuses, telles que des lignes électriques, des conduites de pétrole, des éoliennes et des ponts. Les drones d'inspection et de surveillance peuvent également être utilisés pour inspecter des espaces confinés difficiles à inspecter par des humains, tels que des cheminées.

3.7.2 Vidéo et surveillance

1. **Surveillance en temps réel :** Les drones ont été considérés comme une nouvelle technologie de surveillance du trafic et des endroits, pour collecter des informations. Par rapport aux dispositifs de surveillance traditionnels tels que les détecteurs de boucle, les caméras vidéo de surveillance et les capteurs à micro-ondes.
2. **Média presse et streaming vidéo :** Pour diverses applications de diffusion multimédia en temps réel, les drones connectés au cellulaire doivent générer des vidéos à partir de fichiers de données collectés à l'aide de capteurs et de caméras.
3. **suivi des cibles :** (Tracking Target) Les informations de localisation peuvent être complétées par des images aériennes d'objets cibles [19]. Les drones peuvent être intégrés à des technologies de détection de cibles, notamment des caméras thermiques et de vision. Les caméras thermiques (par exemple, les caméras infrarouges) peuvent détecter le profil thermique pour localiser les personnes disparues.
4. **Application de la loi et surveillance des contrôles aux frontières :**

L'autonomes des UAVs prévoit de surveiller les frontière, et patrouillant ensemble. L'intelligence artificielle de ces drones leur permettrait de reconnaître les humains et de distinguer si ces derniers commettent des infractions comme celui de passer une frontière et leur dangerosité pour ensuite prévenir la police aux frontières.

Ces drones peuvent se mouvoir dans les airs, sous l'eau, sur l'eau et dans

des engins au sol. Dotés de multiples capteurs, en plus de la détection d'activités criminelles.

3.7.3 Opérations de recherche et d'action

1. **Chercher et sauver** : Les drones de recherche et de sauvetage donnent aux premiers intervenants l'avantage de survoler les zones sinistrées ou de se rendre dans des endroits trop difficiles à atteindre pour les humains et à la recherche de victimes. Certaines des applications des UAVs de recherche et de sauvetage sont l'évaluation post-catastrophe, la recherche de personnes disparues et la prestation des premiers soins aux victimes.
2. **Rechercher et attaquer** : Un UAV est considéré comme un type d'attaque s'il est piloté ou fabriqué à des fins militaires, telles que la collecte de renseignements, la surveillance, la reconnaissance, l'acquisition d'objectifs et l'évaluation des dommages au combat, ou si l'UAV peut transporter une variété d'armes et effectuer des frappes aériennes.

3.7.4 Couverture de zone (Area coverage)

1. **Exploration pétrolière, gazière et minière** : Les UAVs à beaucoup pour offrir dans les domaines de l'exploration pétrolière, gazière et minérale et de la surveillance des pipelines et des installations. Ceci est réalisé par le personnel des sociétés d'exploration.
2. **Cartographie** : (mapping) Les UAVs sont utilisés dans les applications de cartographie et d'arpentage en raison des images haute résolution qui peuvent être prises par un UAV volant à basse altitude et qui peuvent être cartographiées à une distance plus proche par rapport aux images obtenues à partir d'un avion ou d'un satellite. Le 3D Robotics Aero est un exemple d'UAV commercial d'arpentage et de cartographie. Il s'agit d'un UAV électrique lancé à la main qui peut avoir plusieurs caméras et des charges utiles personnalisables.

3. Suivi de la croissance des plantes : (Follow up plant growth)

Les lectures de température par des capteurs aériens peuvent être affectées par les stades de croissance des cultures. Au début de la saison de croissance, lorsque les plantes sont petites et clairsemées, les mesures de température peuvent être influencées par la réflectance de la surface du sol.

3.7.5 Précision agriculture

Les drones agricoles sont généralement équipés de caméras, de capteurs multi-spectraux, de capteurs d'évitement d'obstacles, d'un réservoir et de pulvérisateurs. Ces drones sont utilisés pour surveiller la santé des cultures et la qualité des sols.

1. **Management de l'irrigation** : En utilisant des drones pour mesurer l'humidité du sol, la température des plantes et l'évapotranspiration. Par exemple, la distribution spatiale de l'humidité du sol de surface peut être estimée à l'aide d'images multi-spectrales à haute résolution capturées par un [UAV](#).
2. **Pulvérisation de pesticides** : Les drones agricoles jouent un rôle déterminant dans l'agriculture de précision. Les réservoirs équipés sur les drones agricoles sont utilisés pour fournir de l'eau et des engrais aux zones de cultures soumises à un stress hydrique et nutritif et des pesticides aux parties de cultures infestées. Les drones sont également utilisés pour fournir une alimentation en direct pour la surveillance du bétail.

3.7.6 Livraison (Delivery)

Les drones peuvent être utilisés pour transporter de la nourriture, des colis et d'autres marchandises [20]. Dans le domaine des soins de santé, les drones ambulances peuvent livrer des médicaments, des vaccins et des échantillons de sang, dans et hors des endroits inaccessibles. Ils peuvent

transporter rapidement des instruments médicaux dans les quelques minutes cruciales qui suivent un arrêt cardiaque. Ils peuvent également inclure des services de streaming vidéo en direct permettant aux ambulanciers paramédicaux d'observer à distance et d'instruire les personnes sur place sur la façon d'utiliser les instruments médicaux [21].



FIGURE 3.6 : Utilisation de systèmes UAV dans les opérations de livraison .

3.7.7 Aide à la connexion

(a) **Fournir une zone de couverture sans fil :** (wireless coverage)

Les véhicules aériens sans pilote (UAV) devraient jouer un rôle vital dans les systèmes de communication sans fil. Les drones sont capables de fournir une connectivité sans fil même sans infrastructure réseau. Cela facilite les opérations [Search and rescue \(SAR\)](#), comme illustré dans les figures 3.7 et 3.8 pour les environnements extérieurs et intérieurs, respectivement.



FIGURE 3.7 : Utilisez UAV pour fournir une couverture sans fil aux utilisateurs extérieurs .



FIGURE 3.8 : Utilisez UAV pour fournir une couverture sans fil aux utilisateurs intérieurs .

- (b) **Allocation lien de connexion** : Les UAVs peuvent être utilisés pour alloc et lien de connexion en cas d'urgence dont chaque UAV sert de station de base aérienne sans fil où lorsque le réseau cellulaire tombe en panne

3.7.8 Edge / fog / cloud computing

Avec le développement de l'intelligence artificielle IA et la commercialisation progressive des technologies 5G, l'ampleur des réseaux Internet Of

Things (IOT), il y a une présence croissante d'appareils et de capteurs connectés qui prennent en charge une grande variété d'applications .

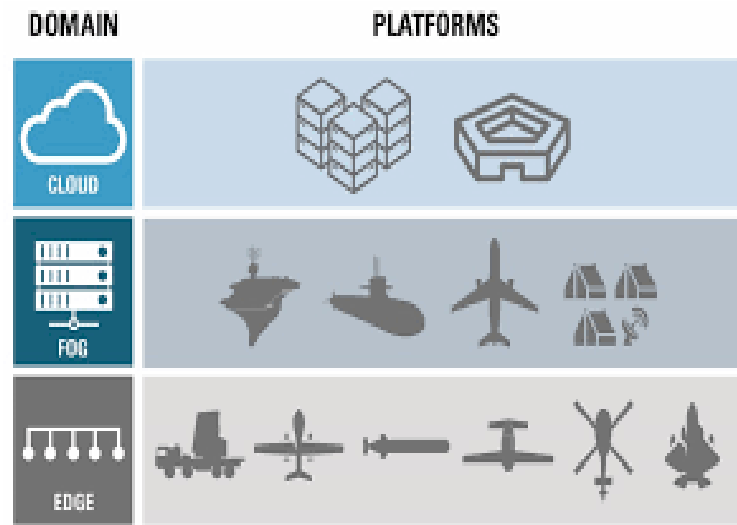


FIGURE 3.9 : edge / fog /cloud présentation .

Le domaine **Mobile Edge Computing (MEC)** assisté par un véhicule aérien sans pilote (**UAV**) est proposé, dans lequel plusieurs **UAVs** ayant des trajectoires différentes survolent la zone cible et prennent en charge les équipements utilisateurs (UE) au sol.

- (a) **Déchargement "offloading"** : Le déploiement des **UAVs** dans la ville intelligente peut réduire les charges de nœuds de capteurs grâce au déchargement des données. Les tâches de déchargement des drones sont transférées vers des serveurs cloud ou des stations de base. Il apporte un délai de transmission et une consommation d'énergie élevés.
- (b) **Mise en cache "Caching"** : Dans les réseaux cellulaires assistés par **UAV** permettant la mise en cache, il existe une macro BS au sol (MBS) avec plusieurs BS UAV. La figure 4.2 montre un exemple du scénario pratique. Le MBS au sol à proximité du stade est surchargé, ce qui ne peut pas répondre aux besoins de trafic des utilisateurs aux heures de pointe, par exemple pendant un match de football.

Cependant, chaque drone ne peut collecter qu'un nombre limité de fichiers de données, ce qui peut ne pas être suffisant pour générer toutes les vidéos demandées.

- (c) **Équilibrage de charge "load balancing"** Dans les scénarios d'informatique de périphérie, l'équilibrage de charge des BS de périphérie est important, en particulier pour les petits serveurs informatiques de périphérie tels que les UAVs avec des ressources limitées. L'équilibrage de charge entre les UAVs peut empêcher les UAVs individuels de fonctionner à des charges élevées pendant une longue période, évitant ainsi la perte de matériel et la génération de grandes quantités de chaleur.
- (d) **Allocation des ressources "Resource allocation"** Les ressources de calcul et de stockage sont placées à la périphérie du réseau mobile, est apparue comme une solution potentielle pour résoudre le problème de latence du réseau en repoussant la frontière des données et des services du cloud centralisé à la périphérie du réseau .

3.7.9 Missions spéciales

- (a) **Filming** : La cinématographie aérienne est une application commerciale majeure des drones. Il est plus facile, plus rapide et moins cher de piloter un UAV pour prendre les photos aériennes nécessaires plutôt que de construire un appareil de lavage "grue-caméra" ou de piloter un hélicoptère pour prendre des photos aériennes.
- (b) **Récupération d'énergie "energy harvesting"** : Le développement d'un système comme source d'alimentation qui améliore la fiabilité et l'efficacité énergétique du vol de l'UAV, en mettant en place des capacités pour redémarrer le moteur en vol, est présenté. La structure matérielle du système de gestion se compose d'un générateur électrique, d'une batterie électrochimique et

de stockages de super-condensateurs, de convertisseurs de puissance électroniques et d'un contrôleur spécialisé.

La conception du système présente une fiabilité élevée ainsi que des caractéristiques d'énergie et de puissance de poids spécifique, qui sont essentielles pour le temps de vol global du drone.

- (c) **Applications dans la pandémie de Covid-19** : Récemment, pendant l'épidémie de virus corona, les gouvernements ont plaidé pour l'utilisation de véhicules aériens sans pilote UAV plus que par le passé. L'utilisation de véhicules aériens sans pilote UAV pour transporter des produits et services importants peut avoir un impact bénéfique significatif.

Étonnamment, des véhicules aériens sans pilote UAV ont été déployés pour lutter contre la propagation du virus. Des essaims de drones ont été utilisés pour transporter des fournitures médicales, arroser les routes avec de l'eau et de la désinfection, et diffuser des informations importantes pour encourager la distance physique et rester à la maison.

De nouveaux drones pour la détection de la température corporelle ont été créés en utilisant l'intelligence artificielle et les drones. De plus, pendant une pandémie, les UAVs peuvent améliorer la connaissance de la situation en aidant à faire respecter les mesures de confinement et les cordons sanitaires. Ces configurations nécessiteraient certainement des essaims de drones plutôt que des systèmes de drones uniques.



FIGURE 3.10 : l'utilisation des UAVs dans la pandémie de covid-19

3.8 Les défis des UAV's

- **Consommation d'énergie** : l'**UAV** doit se rendre sur le lieu de l'événement, traiter l'événement avec les capteurs nécessaires et communiquer les données au client tout en effectuant la tâche. Toutes ces opérations demandent de l'énergie, et l'enjeu est de choisir un prestataire ayant l'énergie nécessaire pour gérer la situation.
- **Sécurité** : L'une des principales difficultés est d'assurer la confidentialité, la disponibilité et l'intégrité des informations lors de la communication entre les différents nœuds de l'essaim. Par conséquent, il est essentiel de sauvegarder les données provenant de sources non autorisées afin de protéger les données contre une utilisation abusive.
- **Modèle de mobilité** : Le modèle de mobilité décrit comment différents UAV peuvent travailler ensemble à l'intérieur d'une zone définie pour accomplir une mission. Pour fournir des performances et une flexibilité améliorées, il est essentiel de définir un modèle de mobilité qui soit cohérent avec les caractéristiques des minuscules drones de l'essaim.
- **Coordination** : Le bloc de coordination est l'entité de raisonnement et de décision, qui est responsable de l'utilisation des observations (propres et des autres UAV), des exigences de la mission et

des contraintes du système pour organiser les UAV. Il doit calculer les trajectoires des drones et prendre des décisions sur la manière d'allouer les tâches pour obtenir un comportement d'équipe. La coordination peut signifier la réalisation et le maintien de formations rigides ou peut être la répartition des tâches entre les véhicules d'une manière auto-organisée.

- **Communication** : Les systèmes mono-UAV doivent maintenir en permanence la communication avec les pilotes au sol ou les infrastructures. En revanche, un système multi-UAV n'a qu'un seul UAV spécifique qui communique avec le sol et transmet le message aux autres UAV.

3.9 Conclusion

Ce chapitre a présenté un aperçu sur les UAVs, avec ses principales caractéristiques, applications, défis et architectures de communication. Dans le prochain chapitre, nous avons décrit brièvement Multi-agent Reinforcement Learning définition, structure et l'interaction entre les agents.

CHAPITRE 4

MULTI-AGENT REINFORCEMENT LEARNING

Sommaire

4.1	Introduction	48
4.2	Définition	48
4.3	Structure du Multi-agent Reinforcement Learning	48
4.4	L'interaction entre les agents	50
4.4.1	Coopération	50
4.4.2	Compétition	51
4.4.3	Hybride	51
4.5	Défis de Multi-agent Reinforcement Learning	51
4.6	Applications du MARL	52
4.7	Multi-agent Reinforcement Learning pour multi-UAVs	53
4.8	Conclusion	69

4.1 Introduction

L'apprentissage par renforcement multi-agents [MARL](#) a récemment attiré l'attention des communautés de l'apprentissage automatique, de l'intelligence artificielle et des systèmes multi-agents. A cet effet, dans ce chapitre nous allons présenter une petite introduction sur [MARL](#), ses applications, les défis du [MARL](#). Alors que la 2ème partie contient notre revue sur les articles qui utilisent [MARL](#) avec un essaim des [UAVs](#).

4.2 Définition

L'apprentissage par renforcement [RL](#) est puissant pour renforcer un agent afin qu'il se comporte de manière permanente en fonction des règles et des récompenses reflétées par l'environnement. Cependant, pour diverses applications en intelligence artificielle, lorsque l'environnement est à grande échelle et que les tâches sont compliquées, non seulement on s'attend à ce qu'un seul agent fasse des actions intelligentes, mais nous espérons aussi qu'il existe un groupe d'agents qui peuvent communiquer et prendre des décisions les un avec les autres pour cette raison apparu la notion Multi-agent Reinforcement Learning [MARL](#) nécessitent d'impliquer plusieurs agents afin de gérer différents types de tâches simultanément.

4.3 Structure du Multi-agent Reinforcement Learning

Pour une analyse claire et une meilleure compréhension, nous définissons les composants de base de l'apprentissage multi-agents comme agent, politique, environnement, fonction de valeur et les récompenses qui sont élaborés plus en détail comme suit :

- **Agent** : Nous définissons l'agent comme l'individu autonome, qui peut interagir de manière indépendante avec l'environnement et adopter sa propre stratégie basée sur l'observation des comportements des autres, visant à atteindre un revenu 'reward' maximum ou une perte minimale pour lui-même. L'apprentissage par renforcement multi-agents est identique aux scénarios d'apprentissage par renforcement habituels.
- **Politique** : 'Policy' Chaque agent suit sa propre politique dans l'apprentissage par renforcement multi-agent. La politique est normalement conçue pour maximiser les revenus et minimiser le coût de l'agent alors qu'il est affecté par l'environnement et les politiques d'autres agents.
- **Récompense** : 'Reward' L'environnement [RL](#) à agent unique a une seule définition de récompense qui spécifie l'objectif de l'agent (chaque agent a leur propre récompense). Tandis que les environnements [MARL](#) peuvent avoir :

une seule récompense globale la récompense globale ne dépend que de l'état global. Elle est facile à calculer et à spécifier, mais conduit souvent à des formations, car les agents ne sont jamais explicitement récompensés ou pénalisés pour leurs propres actions.

Plusieurs récompenses locales chaque agent a leur propre récompense. Peut refléter la contribution de chaque agent au groupe coopérant .

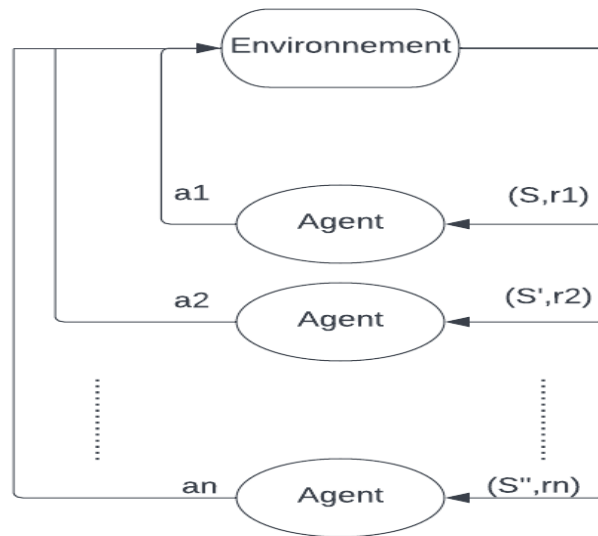


FIGURE 4.1 : La structure du Multi-agent Reinforcement Learning

4.4 L'interaction entre les agents

En conséquence, dans l'apprentissage par renforcement multi-agents [MARL](#), chaque agent effectue un apprentissage de politique de manière autonome, visant à optimiser sa propre récompense, sans tenir compte des autres agents. Donc, il peut exister des compétitions ou des coopérations à travers les interactions avec tous les autres agents.

4.4.1 Coopération

Dans les contextes coopératifs, tous les agents visent à atteindre un objectif commun. Par exemple, tous les agents veulent atteindre leur destination sans dommage. Habituellement, tous les agents partagent une fonction de récompense commune [\[22\]](#).

4.4.2 Compétition

La compétition où la concurrence où chaque agent est indépendant et apprend sa propre politique unique. La récompense est partagée entre tous les agents. Les agents doivent apprendre leurs rôles individuels tout en étant évalués avec les performances d'équipe ou bien la récompense est indépendante pour chaque agent selon , leur performance [23].

4.4.3 Hybride

Dans ce cas les agents peuvent interagir en coopération entre les agents du même groupe et en concurrence contre les agents d'autre groupe.

4.5 Défis de Multi-agent Reinforcement Learning

Bien qu'il s'agisse d'un modèle général qui trouve de nombreuses applications, [MARL](#) souffre de plusieurs défis dans l'analyse théorique, en plus de ceux qui surviennent dans l'agent unique [RL](#). Nous résumons ci-dessous les défis que nous considérons fondamentaux dans le développement du [MARL](#) [24].

- **Objectifs d'apprentissage non uniques :** Contrairement au [RL](#) à agent unique, où le but de l'agent est de maximiser efficacement le rendement à long terme, les objectifs d'apprentissage de [MARL](#) peuvent parfois être vagues.
- **Non-stationnarité :** Un autre défi clé de [MARL](#) est que plusieurs agents apprennent généralement simultanément, ce qui rend l'environnement auquel chaque agent est confronté non stationnaire.
- **Problème d'évolutivité :** Pour gérer la non-stationnarité, chaque agent individuel peut avoir besoin de connaître l'action commune

entre les agents ou la dimension d'espace, augmente exponentiellement avec le nombre d'agents.

- **Diverses structures d'informations** : Par rapport au cas de l'agent unique, la structure d'information de [MARL](#), à savoir qui sait quoi à la formation et à l'exécution, est plus impliqué.

4.6 Applications du MARL

[MARL](#) est utilisé dans différentes applications telles que : appropriation commune des ressources, contrôle des feux de circulation, foot à l'écart, tâches et répartition des ressources ,etc [25].

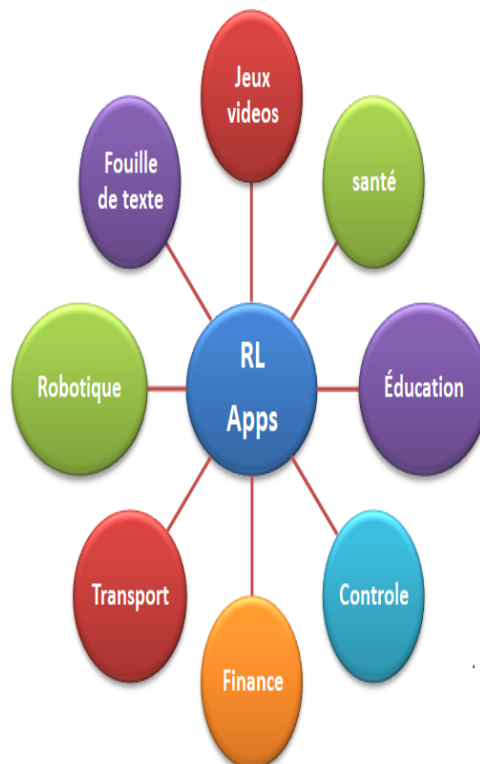


FIGURE 4.2 : Les applications du MARL

4.7 Multi-agent Reinforcement Learning pour multi-UAVs

Dans les tableaux suivants on a étudié plus que 60 articles parmi les 150 articles récents (la plupart entre 2020-2022) qu'on a trouvé sous le thème [RL](#) pour [UAV](#), qui sont utilisés Reinforcement Learning dans les différents domaines d'applications des [UAVs](#).

Les critères traités sont :

- Sous domaine d'application comme déjà défini dans chapitre 2.
- Objectif : on définit le ou les objectifs de chaque article.
- Types de RL : [Value-Based \(VB\)](#) / [Policy-Based \(PB\)](#) / [AC](#).
- Les algorithmes : [Q-Learning](#) / [Deep Deterministic Policy Gradient \(DDPG\)](#) / [DPG](#) / [Advantage Actor Critic \(A2C\)](#) / [AC](#) / [Asynchronous Advantage Actor Critic \(A3C\)](#) / [State-Action-Reward-State-Action \(SARSA\)](#) ...
- Le type des [UAVs](#) ([Rotary](#) / [Fixe](#) / [Hybride](#)) utilisé.
- Types de communication ([U2U](#), [U2G](#) et [SATCOM](#)).
- Ainsi que le nombre des [UAVs](#).
- On a précisé aussi le type de fonctionnement de ces multi-[UAVs](#) ([Coopératif \(P\)](#) / [Concurrent \(C\)](#) / [Hybride \(H\)](#)).

Tous ces critères sont considérés comme des colonnes dans nos tableaux, on a ajouté une colonne "Remarque" pour citer les remarques non communs d'un article spécifié. On a organisé notre travail dans plusieurs sous tableaux selon les domaines d'applications et pour des raisons de présentation.

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P/ C/ H	Remarque
Couverture sans fil	[26] 2019	Maximiser le score de couverture moyen temporel Maximiser l'équité géographique Minimiser les consommations d'énergie	AC	DDPG	[3-8]	Rotary	U2G (BS) U2U	P	modéliser chaque UAV par des réseaux de neurones profonds (Deep Neural Network)
	[27] 2022	Maximiser la connectivité des appareils mobiles au sol tout en améliorant son utilisation de l'énergie Optimiser la trajectoire de vol 3D de chaque UAV sur une série de pas de temps	VB	Q-Learning	4	Fixe	U2G (BS)	P	Offrir des autres méthodes coopératives ayant une convergence plus rapide dans des conditions de réseau dynamique.
	[28] 2021	Maximiser le débit accumulé du réseau IoT Proposer un nouveau réseau IoT assisté par IRS auxiliaire alimenté par UAV maître	AC	DDPG	2	Rotary	U2G(BS)	P	Adapter méthode intelligent reflecting surface (IRS).
	[29] 2022	Maximiser le débit dans des environnements multicellulaires . Maximiser la somme ou le taux individuel dans les réseaux multi-UAV	VB	Q-Learning	3	Rotary	U2G(BS)	P	Adapté les futurs réseaux sans fil de sixième génération (6G).
	[30] 2021	Optimiser dynamiquement la conception de la trajectoire et l'attribution de la bande passante. Maximiser la capacité totale des canaux du réseau.	AC	DDPG	3	Fixe	U2G(BS)	P	Une fonction de récompense appropriée est bien conçue pour guider les agents à exploiter le potentiel du réseau aérien
	[31] 2021	Reconstruire un réseau de communication pour UAV cluster. Résoudre le problème de la convergence difficile causée par la rareté de la récompense.	AC	DDPG	{8,10,12}	Rotary	U2U U2G	P	/
	[32] 2019	Maximiser la durée de vie du réseau. Étudier la planification de trajectoire de plusieurs drones dans un réseau maillé dense de routeurs sans fil (WMN) au sol.	VB	SARSA	[3-9]	Rotary	U2U U2G	P	/

TABLE 4.1 : Classification (1) selon le domaine aide à la connexion

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C/ H	Remarque
Couverture sans fil	[33] 2020	Maximiser la capacité sécurisée en optimisant conjointement la trajectoire des UAVs, la puissance d'émission de l'émetteur UAV et la puissance de brouillage des brouilleurs UAV	AC	DDPG	3	Fixe	U2G (User/ High Altitude Plat forms (HAPs)) U2U	P	Tous les UAVs utilisent le CAA-MADDPG (continuous action Attention) pour trouver une trajectoire optimale pour obtenir une meilleure récompense
	[34] 2022	maximiser le taux de sécurité total de tous les utilisateurs en optimisant conjointement les trajectoires des drones, la puissance de transmission et la puissance de brouillage.	AC	DDPG	3	Hybride	U2U	P	2 brouilleurs UAV et 1 émetteur UAV jammer HAP
	[35] 2019	Étudier le problème du spectre limité dans les réseaux de drones et envisagé une coopérative basée sur des relais.	VB	Q-Learning	[2,6]	/	U2G (BS)	P	/
	[36] 2019	Maximisation du débit.	PB	PPO	3	Rotary	U2G (BS)	P	Utiliser méthode K-Means
	[37] 2019	Maximiser le débit de transmission de la somme instantanée tout en satisfaisant l'exigence de débit des utilisateurs.	VB	Q-Learning	4	Fixe	SATCOM U2G (fixe/mobile)	P	/
	[38] 2021	Réduire le coût énergétique du réseau mmWave compatible UAV .Maximiser le débit d'un réseau cellulaire mmWave assisté par UAV .Résoudre le problème d'allocation des ressources d'énergie renouvelable pour la maximisation du débit.	AC	DDPG	4	Rotary	U2G (Bs /Ms)	P	l'algorithme propose "outperforms the Random Power (RP), Maximal Power (MP) et value-based Deep Q-Learning (DQL)" basés sur la valeur en termes de débit du réseau

TABLE 4.2 : Classification (2) selon le domaine aide à la connexion

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Allocation de lien de connexion	[39] 2020	Explorer la trajectoire optimale. Optimiser la trajectoire du drone pour une stabilité sans fil	AC	Soft Actor Critic (SAC)	3	Rotary	U2G ((High-Speed Trains(HST)))	P	/
	[40] 2020	Optimiser le trajectoire conjointe d'UAV optimiser la formation de faisceaux(Beamforming) passifs au RIS, optimiser l'ordre de décodage, optimiser la détermination de l'allocation de puissance.	VB	DQN	2	Rotary	U2G (BS)	P	Un nouveau cadre est proposé pour integrating reconfigurable intelligent surfaces (RIS) in unmanned aerial vehicle (UAV Nonorthogonal multiple access (NOMA) technique
	[41] 2020	Minimiser l'age de l' information (AOI : Age of information) des UAV's en concevant leurs trajectoires Evaluer la fraîcheur des données sensorielles.	AC	DDPG	{6,8,10,12}	Hybride	U2G (BS/device)	P	/
	[42] 2021	Maximiser le débit du système; Garantir l'équité au niveau de l'utilisateur ; Garantir une consommation énergétique équilibrée des UAV-BS.	AC	DDPG	[2-6]	Fixe	U2G (BS)	P	Proposer une méthode de contrôle des UAV-BS multi-agents nommée MAUC basée sur le cadre de la formation centralisée avec exécution distribuée
	[43] 2020	Formuler le problème du signal de groupe d'UAV tout en satisfaisant couvrant l'environnement de tâche.	VB	Q-Learning	[5-10]	Rotary	SATCOM U2G (BS)	P	Implementer " graph convolutional multi-agent reinforcement learning (DGN)

TABLE 4.3 : Classification (3) selon le domaine aide à la connexion

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P/ C/ H	Remarque
Allocation de lien de connexion	[44] 2021	Maximiser l'efficacité énergétique. minimiser le nombre d'utilisateurs en panne	VB	Q-Learning	{2,4}	Rotary	U2G (BS)	P	Réseaux de sixième génération (6G) adaptés Utiliser algorithme de K-Means
	[45] 2022	étudier le problème de conception de trajectoire pour les réseaux multi-UAV en duplex intégral Maximiser le taux de somme des équipements utilisateurs (UE) .	AC	TRPO / PPO	4	Fixe	U2G (BS /User Équipement)	P	Découple uplink (UL)- downlink (DL)
	[46] 2022	Maximiser les débits de données des utilisateurs et trouver la meilleure trajectoire des drones. Maximiser le taux de liaison descendante des utilisateurs ainsi que le regroupement des utilisateurs et le placement initial des UAVs	AC	AC proposé	4	Rotary	U2G (BS/GU)	P	Propose AC algorithme combiné avec K-Means
	[47] 2021	Optimiser les trajectoires UAV-BS ou le contrôle d'accès GU. Maximiser le débit équitable défini Améliorer le débit total et de maintenir l'équité GU.	AC	DDPG modifié (Air Ground- Probabilistic DDPG)	[2-5]	Rotary	U2G (BS/GU)	H	/

TABLE 4.4 : Classification (4) selon le domaine aide à la connexion

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Resource Management	[48] 2020	maximiser le nombre de tâches déchargées tout en satisfaisant leurs exigences hétérogènes de qualité de service (QoS)	VB	DDPG	3	Fixe	U2G (fixe MeNB) macro eNodeB	P	Etudié la gestion multidimensionnelle des ressources dans les réseaux de véhicules assistés par MEC et UAV.
	[49] 2021	Minimiser la consommation énergétique totale du système UAV-MEC Optimiser les décisions de déchargement Optimiser l'allocation des ressources l'association des utilisateurs et le contrôle de l'alimentation.	AC	DDPG	{2,4,6} {8,10}	Rotary	U2G U2U	P	introduire les différentiels Gaussien pour protéger la vie privée de tous les utilisateurs
Équilibrage de charge	[50] 2021	Résoudre le problème de migration des tâches. Réaliser l'équilibrage de charge Améliorer le temps de vol. Réduire la consommation totale d'énergie.	AC	AC proposé (MDPUR)	[1-35]	Rotary	U2G (SMD (smart mobile device) /Bs)	P	Proposer un algorithme d'itération de valeur basé sur les avantages (ABVI) pour obtenir la stratégie de migration des tâches efficace

TABLE 4.5 : Classification (1) selon le domaine Fog /Edge /Cloud Computing

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Offloading	[51] 2021	Optimiser l'équité géographique entre tous les User Equipements. Optimiser les décisions d'offloading des UE.	VB	DDPG	{3,4}	Fixe	U2G (mobile)	P	introduisons une approche peu complexe pour optimiser les décisions de déchargement des UE, et ne considérons pas la contrainte de la ressource informatique maximale disponible dans chaque UAV.
	[52] 2021	offloaded dynamiquement les tâches des drones vers le cloud périphérique. Améliorer conjointement l'efficacité énergétique (EE)	AC	AC modifié	3	Rotary	U2G(BS)	P	Le système est partagé de manière appropriée entre les nœuds et utilisé
	[53] 2021	Minimiser le coût global de calcul du réseau tout en garantissant les exigences de qualité de service (QoS) des appareils IoT ou des UE dans le réseau IoT.	AC	AC modifié	2	Rotary	U2G (fixe/mobile) U2U SATCOM	P	adapter la technologie 5G afin d'améliorer la qualité du service (QoS)
	[54] 2021	Optimiser le délai de déchargement, et l'efficacité énergétique.	AC	DDPG	[1-9]	Fixe	U2G ("GU" ground users)	P	propose potential game combined
	[55] 2021	Proposer un schéma coopératif de détection de données et de déchargement des calculs pour les systèmes MCS assistés par UAV	AC	AC (modifié AT-MARL)	[1-5]	Rotary	U2G (Bs / Zone de détection)	P	Propose attention -critic network l'exploitation des UAV's dans le crowdsensing mobile (MCS)

TABLE 4.6 : Classification (2) selon le domaine Fog /Edge /Cloud Computing

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Ressource allocation	[56] 2019	Étudie l'allocation dynamique des ressources de plusieurs réseaux de communication activés par les UAV . Maximiser les récompenses à long terme	VB	Q-Learning	4	Rotary	U2G (mobile User)	P	/
	[57] 2021	Planifier conjointe des trajectoires, ordonnancement des données et gestion des ressources. Minimiser des systèmes MEC assistés par UAV	AC	DDPG	6	Rotary	U2G (BS & device)	P	adapter l'approche edge federated multi-agent actor-critic (MAAC). Utiliser 6 mec server
Ressource allocation & Offlo ading	[58] 2021	Minimiser le délai d'exécution des tâches et la consommation d'énergie Minimiser les coûts globaux de l'environnement	AC	DDPG	{3,5}	Rotary	U2G (AerialBS) U2U	P	utiliser réseaux sans fil de cinquième génération (5G)
Ressource allocation & mise en cache	[59] 2022	Planifier la trajectoire des UAV.Allocation des ressources de mise en cache, de calcul et de communication des UAV. Optimiser l'efficacité énergétique .	AC	DDPG	4	Rotary	U2G (GS) U2U	P	Adapté la méthode Tabu Search pour éviter la répétition des services

TABLE 4.7 : Classification (3) selon le domaine Fog /Edge /Cloud Computing

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Collection de donnée	[60] 2019	Déterminer les informations de mobilité des utilisateurs sur la base d'un ensemble de données collecté. Maximiser le débit de transmission. Déterminer les positions optimales des drones en fonction de l'emplacement initial des utilisateurs.	VB	Q-Learning	4	Hybride	U2G	P	echo state network (ESN) based prediction algorithm est proposé pour prédire les positions futures des utilisateurs sur la base de la base de données.
	[61] 2021	formuler le problème de minimisation de l'âge moyen de l'information (AoI) du système MEC assisté par UAV.	AC	DDPG	{-}	Rotary	U2G (BS & device)	P	adapter l'approche edge federated multi-agent actor-critic (MAAC)
	[62] 2021	Minimiser la consommation totale d'énergie d'UAV et des capteurs. Minimiser la consommation énergétique globale des UAV's pour accomplir la collecte de données	AC	DDPG & Q-Learning	{1,2}	/	U2G (Capture)	P	Comparer l'utilisation d'un seul UAV et de multi-UAV
	[63] 2022	Optimiser les trajectoires des deux équipes, minimiser l'âge de l'information (AoI) attendu Maximiser le débit des dispositifs IoT. Minimiser l'utilisation d'énergie des drones et améliorer le transfert d'énergie	AC	AC modifié (TEAM méthode)	[2-20]	Rotary	U2G (device) SATCOM où BS U2U	P	Comparer l'utilisation d'un seul UAV et de multi-UAV

TABLE 4.8 : Classification (1) selon le domaine collection des données et détection

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Détection	[64] 2020	Minimiser l'AoI cumulé pendant une durée donnée.	AC	A2C	2	Fixe	U2G	P	sense-and-send protocol à été proposé . Les UAV's ne doivent pas seulement considérer l'AoI actuel, mais aussi l'AoI que les tâches auront à l'avenir
	[65] 2018	résoudre le problème de conception de trajectoire d'UAV de manière décentralisée.	VB	Q-Learning	3	Fixe	U2G(BS)	P	sense-and-send protocol
	[66] 2021	minimiser AOI pendant la collection des données par les UAVs. Adapter (AoI) age of information	AC	DDPG	{ 6,8 ,10,12 }	Hybride	U2G (device (U2D) / BS)	P	Utilise 2 modes de transmission UAV to device via bs (5G) La détection et la transmission des UAV's sont couplées les unes aux autres

TABLE 4.9 : Classification (2) selon le domaine collection des données et détection

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Suivi de cible	[67] 2020	Localiser une cible mobile radiofréquence (RF). Rechercher la cible avec une précision de suivi spécifiée.	VB	Q-Learning	{2,3,5}	Rotary	U2G	P	Formuler un problème de décision de vol sous la forme d'un processus de décision de Markov contraint (CMDP) avec l'objectif principal d'éviter une trajectoire de vol UAV redondante
	[68] 2022	Trouver des chemins optimaux pour la navigation multi-essaims et l'évitement d'obstacles dans des environnements 3D compliqués invisibles.	AC	PPO / SAC	18	Fixe	U2G	P	Diviser un swarm en plusieurs sous-swarm et combier ces sous-swarm pour former un grand swarm unique en fonction du nombre de cibles et de la distance qui les sépare.
	[69] 2020	Résoudre le problème de suivi des cibles par multi-uav .	VB	Q-Learning	{2,3,5}	Fixe	U2G	P	Utiliser la méthode de régression du processus gaussien (GP)
	[70] 2022	Imiter le scénario selon lequel des multi quadri coptères poursuivent la cible en milieu urbain avec des multi-obstacles	AC	DDPG	[1-5]	Rotary	U2G	H	/
	[71] 2021	Explorer un nouveau schéma de planification de trajectoire de vol pour un système multi-UAV. Planification de trajectoire de plusieurs drones pour suivre des cibles. Former le cluster UAV construit par Le système d'IA.	AC	DDPG DQN	/	Rotary	U2G (GBS)	P	/

TABLE 4.10 : Classification (1) selon le domaine vidéo & surveillance

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Suivi de cible / Surveillance en temps réel	[72] 2019	Optimiser l'affectation de cible. Optimiser la planification de trajectoire séparément. Suivre plusieurs cibles dans des lieux différents par multi-UAV.	AC	DDPG	{3,5}	Rotary	U2G	P	/
	[73] 2020	Explorer l'environnement Evitant les chevauchements entre les observations d'UAV. Trouver toutes les cibles en un minimum de temps. Eviter les collisions	AC	DDPG	3	Rotary	U2U	P	Combiner méthode DDPG avec LSTM (type ANN)

TABLE 4.11 : Classification (2) selon le domaine vidéo & surveillance

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Récupération d'énergie	[53] 2021	fournir des services de déchargement de tâches informatiques et d'allocation de ressources aux appareils IoT.	AC	AC modifié	2	Hybride	SATCOM U2U	P	/
	[74] 2022	Optimiser la tâche de transfert d'énergie entre les sources d'énergie volantes (FES flying energy sources) et les UAVs. Maximiser la somme d'énergie reçue par tous les UAV basés sur les FES en utilisant le wireless power transfer (WPT). Calculer les trajectoires les plus économes en énergie des FES tout en effectuant leurs tâches .	VB	DQN	10	Rotary	U2G(BS)	P	/
	[75] 2021	Réduire le coût énergétique du réseau mmWave compatible UAV. Maximiser le débit d'un réseau cellulaire mmWave assisté par UAV. Résoudre le problème d'allocation des ressources d'énergie renouvelable pour la maximisation du débit.	AC	DDPG	4	Rotary	U2G (BS/MS)	P	l'algorithme proposé outperforms the Random Power (RP), Maximal Power (MP) and value-based Deep Q-Learning (DQL) basés sur la valeur en termes de débit du réseau . Il existe 12 utilisateurs

TABLE 4.12 : Classification (1) selon le domaine missions spéciales

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Énergie Ressource	[76] 2021	Effectuer le traitement du big data de manière distribuée. Minimiser/optimiser l'achat d'énergie auprès des entreprises de services publics externes afin de minimiser les coûts d'exploitation du système.	VB	DQN modifié CommN	30	/	U2G (mobile/fixe)	P	/
Recharge d'énergie	[63] 2022	Optimiser les trajectoires des deux équipes, minimiser l'âge de l'information (AoI) attendu. Maximiser le débit des dispositifs IoT. Minimiser l'utilisation d'énergie des drones et améliorer le transfert d'énergie. proposer comment un algorithme MARL peut être appliqué pour résoudre un problème de couverture optimale.	AC	AC modifié (TEAM méthode)	[2-20]	Rotary	U2G (device) SATCOM où BS U2U	P	Combiner deux tâches collecte de donnée et récupération d'énergie Utiliser deux type des UAVs (collecte de donnée /récupération d'énergie)
Cartogr aphie	[77] 2018	Combiner la méthode DDPG avec une Bio-inspired méthode pour résoudre le problème de la navigation autonome des UAV's.	VB	Q-Learning	3	Rotary	U2G	P	Utiliser Game Theory
Livraison	[78] 2018		AC	DDPG	{3,5,7}	Rotary	U2G (BS)	P	méthode proposée inspirée des bancs de poissons et des troupeaux d'oiseaux (fish schooling and bird flocking)

TABLE 4.13 : Classification (2) selon le domaine missions spéciales

Sous Domaine	N° Année	Objectif	Type RL	Algorithme Proposé	Nombre UAV	Type UAV	Communication UAV	P / C / H	Remarque
Recherche et attaque	[79] 2021	Proposer un réseau de transfert d'agrégation pour le combat multi-UAV contre des agents fixes.	AC	DDPG	{3,5}	Fixe	U2G	P	Combiner le réseau de neurones graphiques (Graph Neural Network) avec le DDPG
	[80] 2019	Utilise l'expérience acquise dans des tâches de combat plus simples pour aider à la formation pour des tâches complexes . Améliorer les performances en formant de manière itérative les agents et leurs homologues.	VB/AC	DQN DDPG	{3 vs 3}	Fixe	U2G	H	Utiliser deux : scenario-Transfer training. Et self-Play training. Comparer deux techniques
	[81] 2021	Appliquer MARL à la prise de décision du combat collaboratif multi-UAV pour un ground attaque.	AC	DDPG	4	Fixe	U2G (BS/MS)	P	2 reconnaissance-attack(RA) UAVs and 2 jamming(JAM) UAVs
	[82] 2021	Proposer DDPG modifié pour un scénario des combats entre les UAVs dans des environnements non stationnaires.	AC	DDPG modifié	{10 vs 10}	/	U2G	H	groupe de 10 uav bleu vs groupe 10 uav rouge
	[83] 2020	Utilise MARL pour swarm des UAVs contre les attaques des swarms des UAVs .	AC	DDPG	{10 vs 10}	Fixe	U2G (fixe)	H	Les résultats de cet article ont une certaine importance de référence pour la mise en place d'un système d'entraînement d'essaim d'UAV et la recherche sur le couplage de règles et d'algorithmes
	[84] 2020	Utiliser RL pour un combat aérien 1vs1 UAV . Prédire la position future de la cible.	AC	DDPG	{1 vs 1}	Fixe	/	C	Utiliser un simulateur de combat aérien
	[85] 2021	étudier les problèmes de contrôle de la formation tels que la génération de la formation, la transformation de la formation et l'évitement des collisions.	AC	DDPG	{1,2,3}	Rotary	U2U	C	/
	[86] 2021	Obtenir la politique de manœuvre des UAVs par RL pour des combats aérienne entre les UAVs .	AC	AC proposé (air combat model)	2 vs 2	Fixe	U2G (Target)	P	un réseau neuronal récurrent bidirectionnel (BRNN) est utilisé pour établir une communication entre des individus UAV et le multi-UAV

TABLE 4.14 : Classification (1) selon le domaine recherche et opération

4.8 Conclusion

Ce chapitre a présenté les concepts de base du [MARL](#), structure générale, défis et ses applications comme une première partie du chapitre. Tandis que la 2ème partie du chapitre a contient les articles qui sont fait des études sur multi-[UAVs](#) avec l'apprentissage [MARL](#).

CHAPITRE 5

CONCLUSION ET PERSPECTIVES FUTURES

Sommaire

5.1	Conclusion	70
5.2	Perspectives futures	71

5.1 Conclusion

Récemment, les [UAVs](#) en swarm jouent un rôle essentiel dans un large éventail d'applications et de missions complexes dans plusieurs domaines. En effet, le Reinforcement Learning joue un rôle crucial afin de résoudre un large éventail de tâches en utilisant une architecture simple avec une bonne fonctionnalité et grande efficacité.

Dans ce travail, nous avons présenté une revue sur les applications ou bien les articles qui faisaient des études sur les swarm des [UAVs](#) avec le Reinforcement Learning pour multi-agent [MARL](#).

Premièrement , nous avons présenté l'intelligence artificielle comme une introduction pour présenter le Machine Learning [ML](#) avec les trois types

du [ML](#) : (i) Supervised Learning , (ii) Unsupervised Learning et (iii) Reinforcement Learning on a basé sur le troisième type [RL](#) qui était bien détaillé.

Deuxièmement, nous avons présenté les [UAVs](#) et discuté de leurs classifications. Nous avons également présenté les principales fonctionnalités, applications, architectures et défis d'un swarm de [UAVs](#).

Finalement, nous avons présenté le dernier chapitre qui était divisé en deux parties .

D'abord, la première partie était sur le [RL](#) avec multi-agent [MARL](#), on a discuter sur la structure, l'interaction entre les agents , défis et applications.

Enfin, dans la deuxième partie on a classifié et organiser les articles dans des tableaux, selon plusieurs critères.

5.2 Perspectives futures

Pour s'appuyer sur les réalisations des recherches et de nouveaux articles, certaines perspectives futures doivent être adaptées sur ce thème en raison que c'est un sujet en tendance et très utilisé dans les grands domaines. on vise future-ment à rendre notre revue survey à publier.

BIBLIOGRAPHIE

- [1] Supervised learning. <https://www.diegocalvo.es/en/supervised-learning/>, 2020. Visité le 21 juin , 2022.
- [2] Reinforcement learning. <https://datascientest.com/reinforcement-learning>, 2021. Visité le 22 juin ,2022.
- [3] Ku Chhaya A Khanzode and Ravindra D Sarode. Advantages and disadvantages of artificial intelligence and machine learning : A literature review. *International Journal of Library & Information Science (IJLIS)*, 9(1) :3, 2020.
- [4] <http://www.alloprof.qc.ca/fr/eleves/bv/mathematiques/les-types-de-variables-m1108>. Visité le 16 juin , 2022.
- [5] Uav. <https://www.missiongo.io/unmanned-aerial-vehicles-uav-unmanned-aerial-systems-uas-and-au> 2021. Visité le 22 juin , 2022.
- [6] AM Bappy, Md Asfak-Ur-Rafi, Md Islam, Ali Sajjad, Khan Nafis Imran, et al. *Design and development of unmanned aerial vehicle (Drone) for civil applications*. PhD thesis, BRAC University, 2015.
- [7] Amartya Mukherjee, Vaibhav Keshary, Karan Pandya, Nilanjan Dey, and Suresh Chandra Satapathy. Flying ad hoc networks :

- A comprehensive survey. *Information and decision sciences*, pages 569–580, 2018.
- [8] Lina Tang and Guofan Shao. Drone remote sensing for forestry research and practices. *Journal of Forestry Research*, 26(4) :791–797, 2015.
- [9] Thomas J Gleason. *Introduction to UAV systems*. Wiley, 2012.
- [10] Joseph Barnard. Small uav command, control and communication issues. *IEEE on communicating with UAV*, page 2007, 2007.
- [11] Yong Zeng, Rui Zhang, and Teng Joon Lim. Wireless communications with unmanned aerial vehicles : Opportunities and challenges. *IEEE Communications Magazine*, 54(5) :36–42, 2016.
- [12] Kenzo Nonami, Farid Kendoul, Satoshi Suzuki, Wei Wang, and Daisuke Nakazawa. *Autonomous flying robots : unmanned aerial vehicles and micro aerial vehicles*. Springer Science & Business Media, 2010.
- [13] Mohammad Mozaffari, Walid Saad, Mehdi Bennis, Young-Han Nam, and Mérouane Debbah. A tutorial on uavs for wireless networks : Applications, challenges, and open problems. *IEEE communications surveys & tutorials*, 21(3) :2334–2360, 2019.
- [14] Naser Hossein Motlagh, Tarik Taleb, and Osama Arouk. Low-altitude unmanned aerial vehicles-based internet of things services : Comprehensive survey and future perspectives. *IEEE Internet of Things Journal*, 3(6) :899–922, 2016.
- [15] Mitch Champion, Prakash Ranganathan, and Saleh Faruque. Uav swarm communication and control architectures : a review. *Journal of Unmanned Vehicle Systems*, 7(2) :93–106, 2018.

- [16] Lav Gupta, Raj Jain, and Gabor Vaszkun. Survey of important issues in uav communication networks. *IEEE Communications Surveys & Tutorials*, 18(2) :1123–1152, 2015.
- [17] Emmanuel Tuyishimire, Antoine Bagula, Slim Rekhis, and Nouredine Boudriga. Cooperative data muling from ground sensors to base stations using uavs. In *2017 IEEE Symposium on Computers and Communications (ISCC)*, pages 35–41. IEEE, 2017.
- [18] Markus Quaritsch, Karin Kruggl, Daniel Wischounig-Strucl, Subhabrata Bhattacharya, Mubarak Shah, and Bernhard Rinner. Networked uavs as aerial sensor network for disaster management applications. *e & i Elektrotechnik und Informationstechnik*, 127(3) :56–63, 2010.
- [19] David Christopher Macke Jr. *Systems and image database resources for UAV search and rescue applications*. Missouri University of Science and Technology, 2013.
- [20] Gerald N Fandetti. Method of drone delivery using aircraft, February 9 2017. US Patent App. 14/817,356.
- [21] Delivery drone. https://en.wikipedia.org/wiki/Delivery_drone, 2022. Visité le 22 juin ,2022.
- [22] Lukas M Schmidt, Johanna Brosig, Axel Plinge, Bjoern M Eskofier, and Christopher Mutschler. An introduction to multi-agent reinforcement learning and review of its application to autonomous mobility. *arXiv preprint arXiv :2203.07676*, 2022.
- [23] Niranjana Balachandar, Justin Dieter, and Govardana Sachithanandam Ramachandran. Collaboration of ai agents via cooperative multi-agent deep reinforcement learning. *arXiv preprint arXiv :1907.00327*, 2019.

- [24] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning : A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control*, pages 321–384, 2021.
- [25] Thanh Thi Nguyen, Ngoc Duy Nguyen, and Saeid Nahavandi. Deep reinforcement learning for multiagent systems : A review of challenges, solutions, and applications. *IEEE transactions on cybernetics*, 50(9) :3826–3839, 2020.
- [26] Chi Harold Liu, Xiaoxin Ma, Xudong Gao, and Jian Tang. Distributed energy-efficient multi-uav navigation for long-term communication coverage by deep reinforcement learning. *IEEE Transactions on Mobile Computing*, 19(6) :1274–1285, 2019.
- [27] Babatunji Omoniwa, Boris Galkin, and Ivana Dusparic. Energy-aware optimization of uav base stations placement via decentralized multi-agent q-learning. In *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, pages 216–222. IEEE, 2022.
- [28] Jingren Xu, Xin Kang, Ronghaixiang Zhang, Ying-Chang Liang, and Sumei Sun. Optimization for master-uav-powered auxiliary-aerial-irs-assisted iot networks : An option-based multi-agent hierarchical deep reinforcement learning approach. *arXiv preprint arXiv :2112.10630*, 2021.
- [29] Suhyeon Lim, Heejung Yu, and Howon Lee. Optimal tethered-uav deployment in a2g communication networks : Multi-agent q-learning approach. *IEEE Internet of Things Journal*, 2022.
- [30] Weijian Wang and Yun Lin. Trajectory design and bandwidth assignment for uavs-enabled communication network with multi-agent

- deep reinforcement learning. In *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, pages 1–6. IEEE, 2021.
- [31] Zixiong Zhu, Nianhao Xie, Kang Zong, and Lei Chen. Building a connected communication network for uav clusters using de-maddpg. *Symmetry*, 13(8) :1537, 2021.
- [32] Wei Zhao, Wen Qiu, Taoyang Zhou, Xun Shao, and Xiujun Wang. Sarsa-based trajectory planning of multi-uavs in dense mesh router networks. In *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–5. IEEE, 2019.
- [33] Yu Zhang, Zhiyu Mou, Feifei Gao, Jing Jiang, Ruijin Ding, and Zhu Han. Uav-enabled secure communications by multi-agent deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 69(10) :11599–11611, 2020.
- [34] Yu Zhang, Zirui Zhuang, Feifei Gao, Jingyu Wang, and Zhu Han. Multi-agent deep reinforcement learning for secure uav communications. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–5. IEEE, 2020.
- [35] Alireza Shamsoshoara, Mehrdad Khaledi, Fatemeh Afghah, Abolfazl Razi, and Jonathan Ashdown. Distributed cooperative spectrum sharing in uav networks using multi-agent reinforcement learning. In *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–6. IEEE, 2019.
- [36] Diego Fernando Cabrera-Castellanos, Alejandro Aragón-Zavala, and Gerardo Castañón-Ávila. Closing connectivity gap : An overview of mobile coverage solutions for not-spots in rural zones. *Sensors*, 21(23) :8037, 2021.

- [37] Xiao Liu, Yuanwei Liu, Yue Chen, and Lajos Hanzo. Trajectory design and power control for multi-uav assisted wireless networks : A machine learning approach. *IEEE Transactions on Vehicular Technology*, 68(8) :7957–7969, 2019.
- [38] Omar Sami Oubbati, Mohammed Atiquzzaman, Abderrahmane Lakas, Abdullah Baz, Hosam Alhakami, and Wajdi Alhakami. Multi-uav-enabled aoi-aware wpcn : A multi-agent reinforcement learning strategy. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE, 2021.
- [39] Yu Min Park, Yan Kyaw Tun, and Choong Seon Hong. Optimized deployment of multi-uav based on machine learning in uav-hst networking. In *2020 21st Asia-Pacific Network Operations and Management Symposium (APNOMS)*, pages 102–107. IEEE, 2020.
- [40] Xiao Liu, Yuanwei Liu, and Yue Chen. Machine learning empowered trajectory and passive beamforming design in uav-ris wireless networks. *IEEE Journal on Selected Areas in Communications*, 39(7) :2042–2055, 2020.
- [41] Fanyi Wu, Hongliang Zhang, Jianjun Wu, Lingyang Song, Zhu Han, and H Vincent Poor. Aoi minimization for uav-to-device underlay communication by multi-agent deep reinforcement learning. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pages 1–6. IEEE, 2020.
- [42] Zhenquan Qin, Zhonghao Liu, Guangjie Han, Chuan Lin, Linlin Guo, and Ling Xie. Distributed uav-bss trajectory optimization for user-level fair communication service with multi-agent deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 70(12) :12290–12301, 2021.

- [43] Anna Dai, Rongpeng Li, Zhifeng Zhao, and Honggang Zhang. Graph convolutional multi-agent reinforcement learning for uav coverage control. In *2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1106–1111. IEEE, 2020.
- [44] Seungmin Lee, Heejung Yu, and Howon Lee. Multiagent q-learning-based multi-uav wireless networks for maximizing energy efficiency : Deployment and power control strategy design. *IEEE Internet of Things Journal*, 9(9) :6434–6442, 2021.
- [45] Chen Dai, Kun Zhu, and Ekram Hossain. Multi-agent deep reinforcement learning for full-duplex multi-uav networks. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2232–2237. IEEE, 2022.
- [46] Maedeh Nasr-Azadani, Jamshid Abouei, and Konstantinos N Plataniotis. Single-and multi-agent actor-critic for initial uav’s deployment and 3d trajectory design. *IEEE Internet of Things Journal*, 2022.
- [47] Ruijin Ding, Yadong Xu, Feifei Gao, and Xuemin Shen. Trajectory design and access control for air-ground coordinated communications system with multiagent deep reinforcement learning. *IEEE Internet of Things Journal*, 9(8) :5785–5798, 2021.
- [48] Haixia Peng and Xuemin Shen. Multi-agent reinforcement learning based resource management in mec-and uav-assisted vehicular networks. *IEEE Journal on Selected Areas in Communications*, 39(1) :131–141, 2020.
- [49] Yiwen Nie, Junhui Zhao, Feifei Gao, and F Richard Yu. Semi-distributed resource management in uav-aided mec systems : A multi-agent federated reinforcement learning approach. *IEEE Transactions on Vehicular Technology*, 70(12) :13162–13173, 2021.

- [50] Wu Ouyang, Zhigang Chen, Jia Wu, Genghua Yu, and Heng Zhang. Dynamic task migration combining energy efficiency and load balancing optimization in three-tier uav-enabled mobile edge computing system. *Electronics*, 10(2) :190, 2021.
- [51] Liang Wang, Kezhi Wang, Cunhua Pan, Wei Xu, Nauman Aslam, and Lajos Hanzo. Multi-agent deep reinforcement learning-based trajectory planning for multi-uav assisted mobile edge computing. *IEEE Transactions on Cognitive Communications and Networking*, 7(1) :73–84, 2020.
- [52] Alessio Sacco, Flavio Esposito, Guido Marchetto, and Paolo Montuschi. Sustainable task offloading in uav networks via multi-agent reinforcement learning. *IEEE Transactions on Vehicular Technology*, 70(5) :5003–5015, 2021.
- [53] Abegaz Mohammed Seid, Gordon Owusu Boateng, Bruce Mareri, Guolin Sun, and Wei Jiang. Multi-agent drl for task offloading and resource allocation in multi-uav enabled iot edge network. *IEEE Transactions on Network and Service Management*, 18(4) :4531–4547, 2021.
- [54] Ang Gao, Qi Wang, Wei Liang, and Zhiguo Ding. Game combined multi-agent reinforcement learning approach for uav assisted offloading. *IEEE Transactions on Vehicular Technology*, 70(12) :12888–12901, 2021.
- [55] Ting Cai, Zhihua Yang, Yufei Chen, Wuhui Chen, Zibin Zheng, Yang Yu, and Hong-Ning Dai. Cooperative data sensing and computation offloading in uav-assisted crowdsensing with multi-agent deep reinforcement learning. *IEEE Transactions on Network Science and Engineering*, 2021.
- [56] Jingjing Cui, Yuanwei Liu, and Arumugam Nallanathan. Multi-

- agent reinforcement learning-based resource allocation for uav networks. *IEEE Transactions on Wireless Communications*, 19(2) :729–743, 2019.
- [57] Zheqi Zhu, Shuo Wan, Pingyi Fan, and Khaled B Letaief. An edge federated marl approach for timeliness maintenance in mec collaboration. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2021.
- [58] Abegaz Mohammed Seid, Gordon Owusu Boateng, Stephen Anokye, Thomas Kwantwi, Guolin Sun, and Guisong Liu. Collaborative computation offloading and resource allocation in multi-uav-assisted iot networks : A deep reinforcement learning approach. *IEEE Internet of Things Journal*, 8(15) :12203–12218, 2021.
- [59] Shu Fu, Meng Zhang, Min Liu, Chen Chen, and Fei Richard Yu. Towards energy-efficient uav-assisted wireless networks using an artificial intelligence approach. *IEEE Wireless Communications*, 2022.
- [60] Xiao Liu, Yuanwei Liu, Yue Chen, Luhan Wang, and Zhaoming Lu. Machine learning aided trajectory design and power control of multi-uav. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2019.
- [61] Zheqi Zhu, Shuo Wan, Pingyi Fan, and Khaled B Letaief. An edge federated marl approach for timeliness maintenance in mec collaboration. In *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2021.
- [62] Seyed Saeed Khodaparast, Xiao Lu, Ping Wang, and Uyen Trang Nguyen. Deep reinforcement learning based energy efficient multi-uav data collection for iot networks. *IEEE Open Journal of Vehicular Technology*, 2 :249–260, 2021.

- [63] Omar Sami Oubbati, Mohammed Atiquzzaman, Hyotaek Lim, Abderrezak Rachedi, and Abderrahmane Lakas. Synchronizing uav teams for timely data collection and energy transfer by deep reinforcement learning. *IEEE Transactions on Vehicular Technology*, 2022.
- [64] Jingzhi Hu, Hongliang Zhang, Lingyang Song, Robert Schober, and H Vincent Poor. Cooperative internet of uavs : Distributed trajectory design by multi-agent deep reinforcement learning. *IEEE Transactions on Communications*, 68(11) :6807–6821, 2020.
- [65] Jingzhi Hu, Hongliang Zhang, and Lingyang Song. Reinforcement learning for decentralized trajectory design in cellular uav networks with sense-and-send protocol. *IEEE Internet of Things Journal*, 6(4) :6177–6189, 2018.
- [66] Fanyi Wu, Hongliang Zhang, Jianjun Wu, Zhu Han, H Vincent Poor, and Lingyang Song. Uav-to-device underlay communications : Age of information minimization by multi-agent deep reinforcement learning. *IEEE Transactions on Communications*, 69(7) :4461–4475, 2021.
- [67] Yu-Jia Chen, Deng-Kai Chang, and Cheng Zhang. Autonomous tracking using a swarm of uavs : A constrained multi-agent reinforcement learning approach. *IEEE Transactions on Vehicular Technology*, 69(11) :13702–13717, 2020.
- [68] Suleman Qamar, Saddam Hussain Khan, Muhammad Arif Arshad, Maryam Qamar, and Asifullah Khan. Autonomous drone swarm navigation and multi-target tracking in 3d environments with dynamic obstacles. *arXiv preprint arXiv :2202.06253*, 2022.
- [69] Yu-Jia Chen, Deng-Kai Chang, and Cheng Zhang. Autonomous tracking using a swarm of uavs : A constrained multi-agent reinfor-

- cement learning approach. *IEEE Transactions on Vehicular Technology*, 69(11) :13702–13717, 2020.
- [70] Ruilong Zhang, Qun Zong, Xiuyun Zhang, Liqian Dou, and Bailing Tian. Game of drones : Multi-uav pursuit-evasion game with online motion planning by deep reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
- [71] Bo Liu, Xuan Li, Tianci Zheng, Wenyu Gao, Xinyu Zhang, Xiaoyan Wang, et al. Research on multi uav attack defense confrontation algorithm based on machine learning. *Academic Journal of Computing & Information Science*, 4(3), 2021.
- [72] Han Qie, Dianxi Shi, Tianlong Shen, Xinhai Xu, Yuan Li, and Liujing Wang. Joint optimization of multi-uav target assignment and path planning based on multi-agent reinforcement learning. *IEEE access*, 7 :146264–146272, 2019.
- [73] Vidyasagar Sadhu, Chuanneng Sun, Arman Karimian, Roberto Tron, and Dario Pompili. Aerial-deepsearch : Distributed multi-agent deep reinforcement learning for search missions. In *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pages 165–173. IEEE, 2020.
- [74] Omar Sami Oubbati, Abderrahmane Lakas, and Mohsen Guizani. Multi-agent deep reinforcement learning for wireless-powered uav networks. *IEEE Internet of Things Journal*, 2022.
- [75] Mari Carmen Domingo. Power allocation and energy cooperation for uav-enabled mmwave networks : A multi-agent deep reinforcement learning approach. *Sensors*, 22(1) :270, 2021.
- [76] Soyi Jung, Won Joon Yun, Joongheon Kim, and Jae-Hyun Kim. Coordinated multi-agent deep reinforcement learning for energy-aware uav-based big-data platforms. *Electronics*, 10(5) :543, 2021.

- [77] Huy Xuan Pham, Hung Manh La, David Feil-Seifer, and Aria Ne-fian. Cooperative and distributed reinforcement learning of drones for field coverage. *arXiv preprint arXiv :1803.07250*, 2018.
- [78] Chao Wang, Jian Wang, and Xudong Zhang. A deep reinforcement learning approach to flocking and navigation of uavs in large-scale complex environments. In *2018 IEEE global conference on signal and information processing (GlobalSIP)*, pages 1228–1232. IEEE, 2018.
- [79] Dongsheng Xu, Peng Qiao, and Yong Dou. Aggregation transfer learning for multi-agent reinforcement learning. In *2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, pages 547–551. IEEE, 2021.
- [80] Guanyu Zhang, Yuan Li, Xinhai Xu, and Huadong Dai. Efficient training techniques for multi-agent reinforcement learning in combat tasks. *IEEE Access*, 7 :109301–109310, 2019.
- [81] Jiayong Fang, Yun Han, Zhongliang Zhou, Shitao Chen, and Sheng Sheng. The collaborative combat of heterogeneous multi-uavs based on marl. In *Journal of Physics : Conference Series*, volume 1995, page 012023. IOP Publishing, 2021.
- [82] Shengang Li, Haoang Chi, and Tao Xie. Multi-agent combat in non-stationary environments. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2021.
- [83] Lei Xiang and Tao Xie. Research on uav swarm confrontation task based on maddpg algorithm. In *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pages 1513–1518. IEEE, 2020.
- [84] Mari Carmen Domingo. Power allocation and energy cooperation for

- uav-enabled mmwave networks : A multi-agent deep reinforcement learning approach. *Sensors*, 22(1) :270, 2021.
- [85] Baowei Li, Shuai Li, Chen Wang, Ruifeng Fan, Jinyan Shao, and Guangming Xie. Distributed circle formation control for quadrotors based on multi-agent deep reinforcement learning. In *2021 China Automation Congress (CAC)*, pages 4750–4755. IEEE, 2021.
- [86] Zhang Jiandong, YANG Qiming, SHI Guoqing, LU Yi, and WU Yong. Uav cooperative air combat maneuver decision based on multi-agent reinforcement learning. *Journal of Systems Engineering and Electronics*, 32(6) :1421–1438, 2021.