Kseniia Harshina

Generating UAV Countermeasures using Machine Learning: A Game-Based Approach

MASTER THESIS

submitted in fulfilment of the requirements for the degree of

Master of Science

Programme: Master's Programme Game Studies and Engineering

UNIVERSITÄT KLAGENFURT
FAKULTÄT FÜR TECHNISCHE WISSENSCHAFTEN
INSTITUT FÜR VERNETZTE UND EINGEBETTETE SYSTEME

Alpen-Adria-Universität Klagenfurt

Evaluator:
Univ.–Prof. Dr. techn. Wilfried Elmenreich
Alpen-Adria-Universität Klagenfurt
Institut für Vernetzte und Eingebettete Systeme

Klagenfurt, September 2023

Affidavit

I hereby declare in lieu of an oath that

- the submitted academic paper is entirely my own work and that no auxiliary materials have been used other than those indicated,
- I have fully disclosed all assistance received from third parties during the process of writing the thesis, including any significant advice from supervisors,
- any contents taken from the works of third parties or my own works that have been included either literally or in spirit have been appropriately marked and the respective source of the information has been clearly identified with precise bibliographical references (e.g. in footnotes),
- I have fully and truthfully declared the use of generative models (Artificial Intelligence, e.g. ChatGPT, Grammarly Go, Midjourney) including the product version,
- to date, I have not submitted this paper to an examining authority either in Austria or abroad and that
- when passing on copies of the academic thesis (e.g. in printed or digital form), I will ensure that each copy is fully consistent with the submitted digital version.

I am aware that a declaration contrary to the facts will have legal consequences.

Kseniia Harshina m. p. Klagenfurt, September 2023

Abstract

Unmanned Aerial Vehicles (UAVs), or drones, have become an integral part of various sectors, revolutionizing industries such as agriculture, logistics, and surveillance. However, their ubiquity has also raised concerns about security, particularly the threat of adversarial UAVs in the form of rogue drones. Traditional countermeasures often rely on confrontational approaches, which can lead to unwanted collateral damage. This thesis explores an innovative solution by developing a game-based framework that leverages Machine Learning, Game Theory, and Swarm Intelligence to train UAV swarms to generate effective non-confrontational countermeasures.

The research employs the Multi-Agent Deep Deterministic Policy Gradient algorithm to facilitate cooperative-competitive learning among UAVs. Two distinct approaches, the Communication-based and Dynamic Shield methods, are designed and evaluated.

Beyond simulation results, this study discusses the practical implications and applications of the proposed framework, highlighting its potential in UAV defense, Multi-Agent Reinforcement Learning, and the ethical considerations surrounding AI-driven simulations.

This research contributes to the development of non-confrontational countermeasures against adversarial UAVs while opening avenues for interdisciplinary collaboration, ethical discussions, and further research in the realm of MARL. Ultimately, it addresses a pressing issue in the era of drone technology, offering a promising approach to enhance security while minimizing harm.

Zusammenfassung

Unbemannte Luftfahrzeuge (im Englischen Unmanned Aerial Vehicles (UAVs)), oder Drohnen, sind heutzutage aus verschiedenen Bereichen nicht mehr wegzudenken und revolutionieren bereits Branchen wie jene der Landwirtschaft, Logistik, oder Überwachungstechnologien. Ihre Allgegenwärtigkeit hat jedoch auch Bedenken hinsichtlich der Sicherheit aufgeworfen, insbesondere die Bedrohung durch feindliche UAVs. Herkömmliche Gegenmaßnahmen beruhen oft auf Konfrontationsansätzen, die zu unerwünschten Kollateralschäden führen können. Dieses Werk präsentiert eine innovative Lösung durch die Entwicklung eines spielbasierten Frameworks, welches Machine Learning, Spieltheorie und Schwarmintelligenz einsetzt, um UAV-Schwärme so zu trainieren, dass sie effektive nicht-konfrontative Gegenmaßnahmen entwickeln.

Es wird der Multi-Agent Deep Deterministic Policy Gradient Algorithmus genutzt, um kooperativ-kompetitives Lernen zwischen UAVs zu ermöglichen. Zwei unterschiedliche Ansätze - die kommunikationsbasierte Methode und die Dynamic Shield Methode, werden entwickelt und ausgewertet.

Neben den Simulationsergebnissen werden in diesem Werk auch die praktischen Implikationen und Anwendungen des vorgeschlagenen Frameworks diskutiert. Das Potenzial für die UAV-Verteidigung durch das Framework und Multi-Agent Reinforcement Learning wird hervorgehoben, und ethische Überlegungen bezüglich KI-gesteuerten Simulationen werden ebenfalls angesprochen.

Diese Forschung trägt zur Entwicklung von nicht-konfrontativen Gegenmaßnahmen gegen gegnerische UAVs bei und eröffnet Wege für interdisziplinäre Zusammenarbeit, ethische Diskussionen und weitere Forschung im Bereich von MARL. Schlussendlich handelt es sich hier um ein drängendes Problem im Zeitalter der Drohnentechnologie, welches hier durch einen vielversprechenden Ansatz für die Bekämpfung von Drohnen angegangen wird, wobei Sicherheit maximiert und potentielle Schäden minimiert werden.

Acknowledgments

I would like to thank my thesis supervisor Wilfried Elmenreich, my employer Lakeside Labs for introducing me to this research, and my project supervisor Melanie Schranz. Additional thanks to Tom Tucek for translating the abstract into German and for the overall help and support.

I would also like to mention that ChatGPT (product version GPT-3.5 [58]), a generative AI model, was used in the context of documenting this research to assist with refining the structure of the outlines and language.

The original project received funding from the Military Counter Unmanned Aerial Systems project supported by FFG Forte under contract number 873514 and the security research program KI RAS of the Austrian Federal Ministry of Agriculture, Regions and Tourism (BMLRT) under grant agreement number 879709 (KI-Secure).

List of Acronyms

ABM = **A**gent-**B**ased **M**odeling

ACO = **A**nt **C**olony **O**ptimization

AI = **A**rtificial **I**ntelligence

API = **A**pplication **P**rogramming **I**nterface

BCI = **B**rain-**C**omputer **I**nterface

BCO = **B**ee **C**olony **O**ptimization

BVLOS = **B**eyond **V**isual **L**ine **o**f **S**ight

CPS = **C**yber-**P**hysical **S**ystem

DAA = **D**etect **a**nd **A**void

DDPG = **D**eep **D**eterministic **P**olicy **G**radient

DQN = **D**eep **Q**-**N**etwork

DRL = **D**eep **R**einforcement **L**earning

EO = **E**lectro **O**ptical

FPV = **F**irst **P**erson **V**iew

GMM = **G**aussian **M**ixture **M**odel

GNN = **G**raph **N**eural **N**etwork

GPS = **G**lobal **P**ositioning **S**ystem

GWO = **G**rey **W**olf **O**ptimizer

FNN = **F**eedforward **N**eural **N**etwork

IR = **I**nfra**r**ed

KPI = **K**ey **P**erformance **I**ndicator

MADDPG = **M**ulti-**A**gent **D**eep **D**eterministic **P**olicy **G**radient

MARL = **M**ulti-**A**gent **R**einforcement **L**earning

MAS = **M**ulti-**A**gent **S**ystem

MDP = **M**arkov **D**ecision **P**rocess

MEC = **M**obile **E**dge **C**omputing

ML = **M**achine **L**earning

MLP = **M**ulti**l**ayer **P**erceptron

Multi-Agent **P**article **E**nvironment

NN = **N**eural **N**etwork

OCSVM = **O**ne-**C**lass **S**upport **V**ector **M**achines

PPO = **P**roximal **P**olicy **O**ptimization

RF = **R**adio **F**requency

RL = **R**einforcement **L**earning

ROS = **R**obot **O**perating **S**ystem

SVM = **S**upport **V**ector **M**achine

UAV = **U**namanned **A**erial **V**ehicle

UAS = **U**namanned **A**erial **S**ystem

UE4 = **U**nreal **E**ngine **4**

Contents

List of Figures

Chapter 1

Introduction

This chapter introduces the research into game-based approaches for tackling challenges in UAV swarm coordination and countermeasures. It provides background and motivation, defines the problem, outlines objectives and research questions, and lays the groundwork for exploring how game-based methodologies can enhance safety and effectiveness in UAV operations.

1.1 Background and Motivation

Unmanned aerial vehicle (UAV) technology has gained a lot of importance in recent years due to its application in multiple diverse fields, such as search and rescue operations, aerial photography and videography, delivery, and logistics, as well as many others. Even though many of those areas pose imminent scientific interest for researchers worldwide, this thesis focuses on UAVs in the fields of military and defense. More specifically, on how to create effective UAV countermeasures against possible adversarial attacks using machine learning (ML) approaches. ML methods will be used for both defending and adversarial agents to ensure robustness and efficiency.

UAVs or drones, as they are commonly known, have been prominently used in military operations equally for defense and offense purposes. In the wrong hands, they pose a serious security threat which makes the development and implementation of defense strategies a task of high priority. An adversarial UAV, also known as an attacking or hostile UAV, is an Unmanned Aerial Vehicle used with malicious intent or aggressive objectives. These UAVs are specifically programmed or operated to execute actions that can include targeting, attacking, surveillance, intrusion, interference, and coordinated attacks. Their primary purpose is to identify, locate, and approach a specific target or objective, after which they may engage in hostile actions, such as deploying weapons or causing disruptions. Adversarial UAVs can also conduct covert surveillance, intrude into restricted airspace, interfere with communication systems, and coordinate attacks when operating in swarms of multiple UAVs.

UAV countermeasures encompass a diverse set of strategies and technologies aimed at detecting, tracking, mitigating, or neutralizing UAV threats, particularly adversarial drones. These countermeasures are pivotal for safeguarding critical infrastructure, ensuring public safety, and protecting sensitive areas from potential UAV-related risks. However, traditional approaches to UAV countermeasures, for the most part, focus on expert knowledge and rule-based systems, which can pose various problems related to a lack of dynamic adaptability, robustness, and flexibility, as well as limited time resources. For these reasons, in recent years, researchers have been exploring different options, primarily those that focus on autonomous AI-related solutions that would solve issues raised by traditional methods. An overview of existing defense strategies can be found in Chapter 2 of this thesis.

The motivation behind this thesis is to explore ML techniques within a game-based setting in order to generate novel and optimal UAV countermeasures that can be potentially used in the future – either for security purposes, for further research of reinforcement learning methods, or as a basis for video game artificial intelligence (AI) implementations.

The advantage of using ML methods stems from the potential autonomy, effectiveness, and efficiency of such methods. The digital simulation environment creates a flexible and safe space for learning and testing agents in a dynamic way in order to achieve the most optimal agent behavior, and training both defensive and adversarial agents at the same time creates opportunities for the defensive UAVs to develop robust countermeasures that are not susceptible to outside influence. Meanwhile, the game-based approach of training used in this project gives essential insights into the role of game elements in creating multi-agent intelligence.

In conclusion, this thesis aims to explore how the use of multi-agent reinforcement learning, swarm intelligence, and game theory can be utilized in order to achieve efficient UAV countermeasures.

1.2 Problem Statement

Considering the prominent relevance of UAV technology in today's world, the research in the field of UAV defense does not seem to follow the same advancements. The current available UAV countermeasures offer limited solutions that often require additional infrastructure. Furthermore, most of the existing methods for UAV countermeasures are being utilized in a military setting and oftentimes lack regard for civilian safety. This fact makes it difficult to translate to other, non-military environments and prompts for the creation of efficient and safe approaches.

Existing literature on UAV countermeasures is also missing defense strategies that focus on non-confrontational approaches. In contrast to conventional methods that involve direct confrontation or neutralization of adversarial drones, non-confrontational approaches focus on more strategic, and less aggressive techniques. These strategies often involve misleading, distracting, delaying, or misguiding attacking UAVs, rather than destroying them. Non-confrontational countermeasures aim to achieve their objectives without causing collateral damage or engaging in overt conflict. However, the lack of representation of such approaches poses many potential dangers as UAV technology continues to develop.

The problem addressed in this thesis is the need for safe non-confrontational methods of employing a defending UAV swarm in order to delay or mislead an attacking UAV swarm. Using a defending swarm provides an opportunity to prevent unnecessary damage which brings an advantage when compared to current approaches. Additionally, ML methodology offers an efficient alternative to less optimal rule-based or expert knowledge approaches. Furthermore, this problem lends itself well to be formulated as an agent-based cooperative-competitive environment, which is a solid foundation for applying game-based Multi-Agent Reinforcement Learning (MARL) and Game Theory approaches. In order to ensure the robustness of generated countermeasures, we will consider both asymmetric scenarios, where the number of adversarial and defending UAVs is different, and symmetric scenarios, where there is an equal number of UAVs on both sides.

Addressing this problem will propel the development of non-confrontational countermeasures for attacking UAV swarms, which will offer safe preventive measures in case of UAV attacks. Additionally, it will contribute to the field of MARL by proposing two novel game-based approaches: The so-called Communication-based approach and the so-called Dynamic Shield approach. Both of these scenarios will employ the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm, which is tailored for cooperative-competitive problems.

The Communication-based approach aims to mislead an attacking UAV swarm by accessing the attacker's communication channel and transmitting information about a fake target, located some distance away from the real target. On the other hand, the Dynamic Shield approach aims to delay the attacking UAV swarm by having the defending agents place themselves around the target, preventing the adversarial agents from getting close. Both of these approaches offer promising UAV countermeasures that function as non-confrontational alternatives to many currently existing countermeasures.

This thesis aims to address the crucial requirement for secure and efficient non-confrontational approaches that can guide defending UAV swarms in delaying or misleading adversarial UAV swarms. The development of these strategies ensures a safer and more versatile UAV ecosystem, allowing UAVs to be employed across various civilian and commercial domains without compromising security. Furthermore, this work aims to propose innovative solutions for airspace protection while unlocking the full potential of UAV technology in diverse sectors and industries.

1.3 Objectives and Research Questions

The primary objective of this research is to develop and evaluate efficient and non-confrontational countermeasures against an intelligent adversarial UAV swarm, so as to prevent it from reaching the designated target. Specifically, the research aims to achieve the following objectives:

- Develop a game-based framework for generating UAV countermeasures using machine learning techniques, specifically the MADDPG algorithm.
- Combine methodologies from Machine Learning, Game Theory and Swarm Intelligence to generate novel and efficient UAV countermeasures.
- Investigate and evaluate different non-confrontational approaches, such as misleading, distracting, or delaying tactics, to effectively counteract UAV swarm attacks.
- Design a cooperative-competitive multi-agent system.
- Investigate the effectiveness and adaptability of the MADDPG algorithm in learning and coordinating UAV countermeasure strategies in dynamic environments.
- Assess the performance and effectiveness of the proposed methods through simulations.

Research Questions:

1. What is an optimal way to generate effective non-confrontational countermeasures, using game-based approaches?
    - Evaluation metrics: defenders' winning probability; adversarial inducement success; robustness to swarm size and initial conditions.
2. What machine learning techniques can be applied to coordinate the behavior of UAV swarms in a cooperative-competitive environment?
    - What benefits are there to training both the defending and the adversarial agents using the MADDPG algorithm?
    - Evaluation metrics: stability of rewards; simulation observations.

By addressing these research questions and achieving the stated objectives, this thesis aims to fill the gaps in the current research on UAV countermeasures and provide novel approaches focusing on preventing confrontation and collateral damage. Furthermore, the research outcomes will contribute to the fields of UAV defense and Game-based AI approaches.

1.4 Definition of Game-based Approaches

In the context of this thesis, the phrase "game-based approaches" is used repeatedly, which encompasses a multifaceted paradigm, which in turn includes both the utilization of game-like environments and the incorporation of fundamental game-like principles in the study and development of intelligent agent strategies. It encompasses the following:

- Game Environments: These are controlled and reproducible settings, often simulations, used to train and test intelligent agents. They provide a foundation for evaluating decision-making, learning, and optimization processes.
- Game-like Principles: Beyond environments, it includes the modeling of strategic interactions, competitive and cooperative behaviors among rational agents, and the application of principles inspired by game theory. These principles underpin the development of intelligent strategies in complex, dynamic, and interactive scenarios.
- Utilizing AI Approaches in Game Contexts: This encompasses the application of methodologies from fields like Swarm Intelligence and other non-game domains within game-like scenarios. It explores how these approaches enhance intelligent agent behavior, coordination, and decision-making.

By adopting this definition, our research explores the strategic nature of UAV defense scenarios, aiming to enhance the coordination, adaptability, and robustness of UAV defense strategies within non-confrontational settings.

1.5 Thesis Organization

This thesis is organized into the following chapters to address the research objectives and answer the research questions outlined in the previous section. Chapter 2 conducts a comprehensive review of the literature related to UAV technology, UAV defense, machine learning techniques, game-based approaches, and UAVs in relation to games. It discusses the existing countermeasures against UAVs, highlighting their limitations and the need for non-confrontational methods. Furthermore, the chapter explores relevant concepts in machine learning, game theory, and swarm intelligence that contribute to the development of UAV countermeasures. Chapter 3 presents the research design and methodology employed in this study. It describes the ML methodology, specifically the MADDPG algorithm, and its applicability to UAV countermeasure generation. Additionally, this section provides details on the environment setup and agent behavior modeling. It also discusses the evaluation metrics utilized to assess the performance of the proposed methods.

Chapter 4 discusses the design and implementation of the game-based framework for generating UAV countermeasures. The chapter provides implementation details and discusses the training and optimization processes. Chapter 5 presents the results obtained from the simulations and experiments conducted. It analyzes and evaluates the performance and effectiveness of the developed non-confrontational methods. Chapter 6 provides a comprehensive discussion of the findings and their implications. It addresses the research questions and objectives of the thesis. Furthermore, this chapter explores the advantages and practical implications of the proposed game-based approaches. Chapter 7 summarizes the main findings of the study, as well as addresses possible future prospects.

Chapter 2

Literature Overview

This chapter examines the various domains where UAVs find utility and the challenges associated with these domains. Additionally, it delves into the spectrum of countermeasures aimed at addressing potential UAV-related issues. The chapter further covers key areas such as the role of machine learning, swarm intelligence, and the interplay between UAVs and video games.

The first section, "Overview of UAV Technology and Applications", provides a foundational understanding of UAVs and their deployment across diverse sectors, from surveillance to research. Subsequently, the second section "Existing Countermeasures against UAVs," focuses on the existing strategies devised to neutralize UAV threats. Furthermore, we explore the domain of "Machine Learning Techniques in UAV Defense." Here, the focus is on the integration of machine learning methodologies to enhance defense mechanisms. The section illustrates which algorithms can be utilized to prevent adversarial UAV behaviors. In "Game-based AI Approaches" we explore game-like principles, strategies, and environments within domains such as Game Theory, Swarm Intelligence, and Multi-Agent Reinforcement Learning. In the "UAVs and Multi-Agent Intelligent Behavior in Games" section, we spotlight the utilization of game design principles to employ and improve simulated environments for evaluating UAV behavior.

2.1 Overview of UAV Technology and its Applications

UAVs have gained significant attention and popularity in recent years due to their diverse range of applications and ongoing advancements in technology and research. This section provides an overview of UAV technology, exploring its various applications in different fields, as well as covering its current main research directions.

2.1.1 Definition and Types of UAVs

According to the US Federal Aviation Administration [21], an Unmanned Aerial Vehicle is defined as an aircraft without a pilot or passengers onboard. Instead, UAVs are either autonomous or remotely controlled. UAVs offer an opportunity to handle tasks that would be otherwise dangerous or impossible for humans or human-manned vehicles. Initially, UAVs were used in the twentieth century primarily for military purposes [19]. Since then, they proved to be an asset in various non-military fields as well. UAVs can be categorized based on their size, operational characteristics, and capabilities. According to the website article "Unmanned Aerial Vehicles (UAVs): Revolutionizing Industries and Beyond" [2], some common types of UAVs include:

- Fixed-wing UAVs: These UAVs have a fixed-wing structure, similar to traditional airplanes. They are known for their long-range capabilities, endurance, and efficient flight patterns.
- Rotary-wing UAVs: Also known as multi-rotor drones or quadcopters, these UAVs use multiple rotors to achieve vertical takeoff and landing. They offer increased maneuverability and stability, making them suitable for close-range operations and hovering tasks. Among their common uses are aerial photography and surveillance tasks.
- Autonomous UAVs: Autonomous UAVs come equipped with flight control systems, sensors, and onboard computing capabilities, enabling them to function without a remote human pilot. They are capable of executing autonomous missions by analyzing real-time data. These types of UAVs find applications in various fields, including aerial surveys and scientific research.

The specifics of the UAV design and hardware are outside the scope of this thesis. However, since the generation of the UAV countermeasures using ML is presumed to be applied to UAVs with autonomous capabilities, autonomous UAVs are especially relevant in the context of this thesis. Thus, for future implementations and deployment, it is vital to keep in mind that the algorithmic behavior of the virtual agents can be then transferred onto real-life technology.

2.1.2 UAV Applications

UAVs have found applications in various industries and domains, revolutionizing traditional processes and enabling new possibilities previously inaccessible to human-manned vehicles. According to the paper by Mohsan et al. [52] some prominent applications of UAV technology include: aerial photography and videography, surveillance and security, agricultural monitoring, disaster management, infrastructure inspection, and many others.

One of the more common and prominent uses of UAVs is aerial photography and videography. UAVs are extensively used for capturing aerial footage in industries such as filmmaking, real estate, and tourism. This particular UAV application is widely accessible to the public and can frequently be observed in everyday usage.

Additionally, UAVs play a crucial role in surveillance and security operations. They provide a cost-effective and efficient means of monitoring large areas, conducting border surveillance, and gathering real-time intelligence. Their ability to access hard-to-reach locations makes them valuable assets in law enforcement and military applications. This specific application area holds utmost relevance to this study as it encompasses UAV countermeasures, which are a subset of the broader usage of UAVs for security purposes.

Furthermore, UAVs equipped with specialized sensors and imaging technologies have proven invaluable in agricultural applications. They aid farmers in monitoring crop health, identifying areas requiring attention, and optimizing fertilizer usage. UAVs contribute to precision agriculture, leading to increased crop yields and reduced environmental impact.

In disaster-stricken areas, UAVs can provide rapid assessment and situational awareness. Equipped with thermal imaging cameras and other sensors, they assist in search and rescue missions, locating survivors, and assessing damage in unsafe or inaccessible environments.

UAVs are also increasingly used for inspecting infrastructure such as bridges, power lines, and pipelines. They can access hard-to-reach areas, capture high-resolution imagery, and detect structural defects or anomalies. This enables proactive maintenance and reduces the need for manual inspections in dangerous environments.

As the application of UAVs continues to expand across various domains, it is crucial to recognize the potential risks associated with their misuse for adversarial purposes. This research aims to address these risks and develop effective strategies for mitigating them.

2.1.3 Research and Development in UAV Technology

The field of UAV technology is continuously evolving, driven by ongoing research and development efforts. Researchers and industry professionals are exploring advancements in areas such as autonomous navigation, sense-and-avoid systems, payload capabilities, and communication technologies. These advancements aim to enhance the performance, safety, and operational efficiency of UAVs across various applications.

There is a growing focus on developing advanced autonomy and decision-making capabilities for UAVs. This includes research on autonomous navigation, path planning, obstacle detection and avoidance, and collaborative behaviors. One of the earliest publications in this regard, Hadi et al. [30], presented an autonomous UAV system designed for a payload-dropping mission. Their work emphasizes the development of decision-making algorithms to ensure the successful execution of mission tasks. Additionally, Zhang et al. [88] highlight the trends in the development of intelligent unmanned autonomous systems. Their research explores advancements in autonomy, perception, and interaction between UAVs and the environment. Song et al. [72] contribute to the field by addressing the rolling horizon path planning of a UAV system for persistent cooperative service. They present both a mixed-integer linear programming formulation and efficient heuristics for this purpose.

UAV swarms, where multiple drones work together in a coordinated manner, have gained significant attention. Research is being conducted on swarm intelligence algorithms, swarm communication and coordination, and applications of UAV swarms in various domains such as search and rescue, surveillance, and environmental monitoring. Zhou et al. [92] discuss recent advances and future trends in UAV swarm intelligence. Additionally, Tang et al. [75] provide a comprehensive review of swarm intelligence algorithms for multiple unmanned aerial vehicle collaborations.

Beyond Visual Line of Sight (BVLOS) operations refer to the ability of UAVs to fly beyond the operator's visual range. There is ongoing research on developing reliable and safe systems for BVLOS operations, including advanced sensing technologies, communication protocols, and regulatory frameworks. Davies et al. [15] conducted a review of unmanned aircraft system technologies to enable BVLOS operations. This review explores the technologies developed to date that enable BVLOS applications. While BVLOS flight operations hold significant commercial potential, concerns remain about UAV capabilities to detect and avoid airborne hazards. Fang et al. [20] focus on the development of small UAV BVLOS flight operations. They address the need for detect and avoid (DAA) systems on small UAVs for BVLOS flight operations in civil airspace. The authors propose testing small UAV DAA systems in BVLOS flights in restricted airspace to improve operational safety. Politi et al. [63] also present a comprehensive survey of current UAV technologies for BVLOS operations. Their work highlights the main technological challenges and requirements for BVLOS UAV operations and emphasizes the potential of BVLOS UAV features in various industrial sectors.

Researchers are furthermore exploring new ways to integrate different types of payloads and sensors onto UAV platforms, such as high-resolution cameras, thermal sensors, and multi-spectral imaging. This enables diverse applications, some of which were mentioned in the previous section, such as aerial mapping, remote sensing, precision agriculture, and infrastructure inspection. Klemas [38] provides an overview of coastal and environmental remote sensing using UAVs. The author discusses the applications of UAV-based remote sensing in studying coastal environments. Nagai et al. [54] delve into UAV-borne mapping through multi-sensor integration. Their work emphasizes the benefits of combining multiple sensors to enhance mapping capabilities. Feroz and Abu Dabous [24] focus on UAV-based remote sensing applications for bridge condition assessment. The authors highlight the potential of UAVs in assessing and monitoring the condition of bridges using remote sensing techniques.

As UAV technology advances, there is a parallel need for research on UAV security against potential threats posed by malicious drones. This includes developing methods to detect, identify, and mitigate unauthorized or potentially dangerous UAVs. Abro et al. [1] provide a comprehensive review of UAV detection, security, and communication advancements aimed at preventing threats.

The growing body of research and technological advancements in the field of UAVs highlights their increasing relevance and widespread application in recent years. This further underscores the relevance of this research, which primarily focuses on UAV security and countermeasures but also has the potential to contribute to the fields of UAV Swarm Intelligence and Autonomous Systems research.

By providing this overview of UAV technology, its applications, and current research directions, we establish a foundation for understanding the context and significance of UAV defense research. This knowledge sets the stage for exploring existing countermeasures and identifying gaps in current approaches.

2.2 Existing Countermeasures Against UAVs

The current extensive range of UAV application areas has raised concerns regarding their potential misuse or malicious intent. Moreover, there have already been multiple instances of UAVs being used for adversarial purposes [61]. As a result, considerable research has been conducted to develop countermeasures to prevent damage from possible future attacks [11]. This section provides an overview of existing countermeasures and techniques employed to detect, track, and neutralize adversarial UAVs.

2.2.1 UAV Detection and Identification Techniques

Radar-based detection systems utilize radio waves to detect and track UAVs. These systems can estimate the UAV's position, speed, and altitude, thus enabling operators to identify potential threats. Hoffmann et al. [32] present a micro-Doppler-based detection and tracking approach using multistatic radar.

Radiofrequency (RF)-based detection techniques analyze the radio frequency spectrum to detect the presence of UAVs. By examining the unique RF signatures emitted by UAVs, these systems can identify and classify UAVs in the vicinity. Ezuma et al. [18] discuss micro-UAV detection and classification from RF fingerprints using machine learning techniques.

Electro-optical (EO) and infrared (IR) sensors detect UAVs based on their visual and thermal signatures. These systems use cameras and thermal sensors to capture images or heat patterns, allowing for the detection and tracking of UAVs in different lighting and environmental conditions. Jara-Olmedo et al. [36] discuss the interface of optimal EO and IR for unmanned aerial vehicles.

2.2.2 UAV Tracking and Monitoring Techniques

Visual tracking algorithms employ computer vision techniques to track the position and trajectory of UAVs based on visual information captured by cameras or video feeds. These algorithms can estimate the UAV's motion and provide real-time updates on its location. In their research, Li et al. [44] introduce AutoTrack, a high-performance visual tracking approach for UAVs with automatic spatiotemporal regularization.

Another approach can be found in acoustic-based tracking systems, which use audio sensors to detect and localize UAVs by analyzing the unique acoustic signatures generated by their propellers or engines. These systems can provide valuable information about the UAV's direction and proximity. In their work, Dumitrescu et al. [17] describe the development of an acoustic system for UAV detection.

Lastly, combining different sensory data, there are drone detection networks, which consist of multiple sensors strategically placed in an area to create a network that collaboratively detects, tracks, and shares information about UAVs. These networks enhance the coverage and accuracy of UAV detection by combining data from various sensors. For example, Dil et al. [16] present SafeSpace MFNet, a precise and efficient multi-feature drone detection network.

2.2.3 UAV Neutralization Techniques

Many techniques to neutralize adversarial UAVs already exist and are actively being explored. RF jamming techniques disrupt the communication between a UAV and its operator by emitting interfering signals in the same frequency bands used by the UAV's control systems. This can force the UAV to lose control or enter a failsafe mode. Van Voorst [80] presents a counter-drone system using RF jamming.

GPS (Global Positioning System) spoofing involves transmitting fake GPS signals to deceive the UAV's navigation system, causing it to deviate from its intended course or lose its position accuracy. This technique can be used to manipulate a UAV's flight path or force it to land in a controlled area. He et al. [31] propose an effective countermeasure against UAV swarm attacks using GPS spoofing.

There are also snare systems, which are designed to physically capture UAVs. These systems typically employ nets or other mechanisms to immobilize or disable the UAV in flight. In their work Goossen and Martinez [28] detail a catch-and-snare system for unmanned aerial vehicles.

Geo-fencing establishes virtual boundaries that, when breached by a UAV, trigger predefined actions such as warnings or forced landings. This technique helps prevent UAVs from entering restricted or sensitive areas. Chan et al. [12] propose a dynamic geo-fence system for UAVs in their research.

2.2.4 UAV Swarm Defense

A UAV swarm refers to a group of UAVs that operate together in a coordinated manner to achieve a common objective. These swarms can exhibit collaborative behaviors, such as communication, coordination, and distributed decision-making, allowing them to accomplish complex tasks more efficiently and effectively than individual UAVs. Defense against UAV swarms poses particular challenges since they tend to be unpredictable due to the ability to change their local actions and interactions dynamically. Moreover, their quantity and speed could change dynamically as well. In their survey titled "Defending airports from UAS: A survey on cyber-attacks and counter-drone sensing technologies" Lykou et al. [49] delve further into the specifics and difficulties of defending against UAV swarms. Liu et al. [46] conducted an initial study exploring a mechanism for identifying and suppressing the emergent behavior of UAV swarms. The authors found that implementing a comprehensive suppression strategy based on accurate recognition of flocking behavior is an effective approach to combat emergent swarm behavior. Chen et al. [13] introduced a counter approach aimed at disrupting the coordination and clustering capabilities of a drone swarm by dividing it into disconnected components. They proposed the use of a genetic algorithm and a particle swarm optimization algorithm to efficiently identify critical nodes for splitting the swarm. Furthermore, Simonjan et al. [70] investigated the intervention of a UAV swarm by deploying defender UAVs with the objective of misleading the swarm's mission. They hypothesized that both attackers and defenders utilize the same bio-inspired swarm algorithms, and they modified the defenders' objective function and interaction to exert influence over the attackers. However, as the optimization algorithm becomes more robust, it becomes increasingly challenging to mislead the swarm. Additionally, their approach relied on the assumption that defenders possess knowledge of the specific swarm algorithm employed by the attackers, which is a significant difference to the approach described in this thesis.

In summary, the existing countermeasures against adversarial UAVs encompass a spectrum of techniques aimed at detection, tracking, and neutralization. From radar systems and acoustic monitoring to innovative strategies targeting swarm dynamics, the field displays a diverse toolkit of defenses. These efforts contribute to a defense framework, continuously evolving to meet the challenges posed by dynamic and evolving adversarial UAV threats, ultimately enhancing the security and resilience of UAV applications. This thesis ultimately aims to contribute to this ever-growing field of research by combining concepts like machine learning and game-based approaches.

2.3 Machine Learning Techniques in UAV Defense

The field of machine learning has witnessed significant advancements in recent years, and its application in UAV defense has gained considerable attention. ML techniques offer the potential to enhance the effectiveness and adaptability of UAV defense systems by enabling autonomous decision-making and real-time response to evolving threats. This section provides an overview of the key ML techniques employed in the context of UAV defense.

2.3.1 Supervised Learning Approaches

Supervised learning algorithms such as Support Vector Machines (SVM), Random Forests, and Neural Networks (NN) have been applied to UAV defense for tasks such as UAV detection, classification, and identification. These algorithms learn from labeled training data to classify incoming sensor data or detect anomalous behavior indicative of adversarial UAVs.

In their paper, Ihekoronye et al. [35] propose a cyber-edge intelligent intrusion detection framework for UAV networks based on the random forest algorithm. The authors leverage mobile edge computing (MEC) technology and an anomaly-based intrusion detection scheme to address security challenges in UAV networks. They utilize an optimized Random Forest model embedded in dedicated UAV-MEC servers for intrusion detection. The proposed model demonstrates superior performance in detecting and classifying different network attacks. Furthermore, Utubor [77] investigates the application of gradient-boosting-based machine learning techniques to enhance the detection of attacks in cyber-physical systems (CPS). The study evaluates gradient boosting models, including XGBoost and LightGBM, in classifying various cybersecurity attacks using the Edge-IIoTset data set. The research emphasizes the superiority of gradient-boosting models and underscores the impact of appropriate sampling techniques on model efficacy. Zhou et al. [91] introduce SwarmNet, a neural network architecture for predicting and imitating swarm behavior. The authors utilize Graph Neural Networks (GNN) to learn and imitate the behavior of observed swarm agents. The SwarmNet architecture employs 1D convolutions, GNNs, and Multilayer Perceptron (MLP, also referred to as feed-forward neural network (FNN)) layers to predict agents' states based on their positions and velocities. The paper discusses the application of reinforcement learning (RL) methods to enhance swarm behavior in uncertain environments. The authors demonstrate the effectiveness of SwarmNet through experiments involving various swarm behavior models. Furthermore, the authors mention that their framework could potentially be used for UAV defense in the future.

2.3.2 Unsupervised Learning Approaches

Unsupervised learning algorithms, such as Gaussian Mixture Models (GMM) and One-Class Support Vector Machines (OCSVM), are used for anomaly detection in UAV defense. These algorithms learn the normal behavior patterns from the training data and flag any deviations as potential adversarial UAVs.

Zhao et al. [89] propose a UAV signal detection algorithm based on a GMM.

The algorithm utilizes an adaptive threshold calculated using the GMM to identify the start-point of a UAV signal accurately. By leveraging GMM and adaptive thresholding, the algorithm can adaptively detect wireless signals in various noise environments. Furthermore, Panice et al. [60] present an approach for detecting GPS spoofing attacks on UAVs using OCSVMs. The authors employ OCSVMs to analyze state estimation and detect anomalies caused by GPS spoofing attacks, highlighting its potential for identifying adversarial UAVs.

2.3.3 Reinforcement Learning

Reinforcement learning algorithms, like Q-learning, have been utilized for UAV defense tasks, such as path planning and decision-making. These algorithms learn optimal policies through trial and error by interacting with the environment, enabling UAVs to adapt their strategies based on feedback and rewards. Deep reinforcement learning algorithms, such as Deep Q-Networks (DQN) and Proximal Policy Optimization (PPO), combine neural networks with reinforcement learning techniques. These algorithms have shown promising results in UAV defense by learning complex behaviors and enabling UAVs to make intelligent decisions in dynamic environments.

For instance, Lillicrap et al. [45] present continuous control with deep reinforcement learning, showcasing its applicability in solving tasks like cart pole swing-up and car driving through learned policies from raw pixel inputs. Yang et al. [84] investigate UAV air combat autonomous maneuver decisions using the Deep Deterministic Policy Gradient (DDPG) algorithm, demonstrating its efficacy in improving the decision-making process. Among the initial attempts to formulate policies for multi-agent systems, Hüttenrauch et al. [34] pioneered an RL-driven solution utilizing actor-critic learning. This approach laid the foundation for the subsequent deployment of various RL techniques to acquire swarm behaviors that are utilized in UAV countermeasures research.

Currently, several research groups have directed their attention towards investigating the challenge of dynamic swarm versus swarm combat. Notably, Xiang et al. [83] employed RL techniques in their recent study. They used the MADDPG algorithm, which provides the multi-agent extension to the DDPG algorithm, alongside the rule-coupled methodology to intercept adversarial UAV swarms, where defenders engage in shooting down the attacking UAVs. The integration of rules and algorithms produced a winning success rate of 81%.

Another recent contribution in this field comes from Wang et al. [81], who delved into the UAV swarm confrontation problem using hierarchical multi-agent RL methods. Notably, they achieved a performance enhancement of 32% compared to the conventional MADDPG approach.

Within the area of RL, multiple attempts have been made to address the issue of UAV countermeasures. Nonetheless, these strategies predominantly center around countering the attacking swarm with confrontation, potentially leading to collateral damage. Thus, there is a demand for innovative solutions that not only counteract

UAV swarms but also mitigate the risk of UAVs potentially crashing and causing damage. In summary, the application of machine learning techniques, particularly reinforcement learning algorithms, holds immense promise in shaping effective UAV defense strategies.

By exploring the application of various ML techniques in UAV defense, we can identify the strengths, limitations, and potential areas of improvement. This understanding will guide our research in selecting suitable ML approaches and adapting them to address the unique challenges posed by adversarial UAVs in non-confrontational scenarios.

2.4 Game-based Approaches in AI

Within the context of our research, we are applying a game-based approach to the generation of UAV countermeasures alongside machine learning methodology. In this section, we are taking a brief look at how various AI and ML techniques utilize game-based approaches to enhance decision-making, learning, and optimization processes. We will discuss the benefits and advantages of using game-based AI approaches in the context of UAV countermeasures which include their ability to provide a controlled and reproducible environment for training and testing intelligent agents, their capacity to handle complex and dynamic scenarios, as well as their potential for adaptive and strategic behavior generation. Game-based approaches provide a powerful framework for modeling the interaction and strategies of both defending and attacking UAVs.

2.4.1 Game Theory

Game Theory is a mathematical framework used to model and analyze decision-making in situations where the outcomes of one's choices depend on the choices of others. It provides a powerful framework for modeling strategic interactions and decision-making among rational agents. In the context of UAV defense, game theory offers valuable insights into the dynamics between defending and attacking UAVs, helping to develop effective strategies for countering adversarial UAVs. Fudenberg and Tirole [26] provide an authoritative reference on game theory. This book is a foundational resource for understanding the principles of game theory and their application to multi-agent scenarios. Basic concepts and definitions in game theory include the following:

1. Players and Strategies: Game theory considers a set of players, each having a set of strategies to choose from. In the context of UAV defense, the players can represent defending UAVs and attacking UAVs, and their strategies can involve various actions or maneuvers.

2. Payoffs and Utilities: Payoffs or utilities quantify the outcome or reward obtained by each player based on their chosen strategies and the strategies of other

players. In UAV defense, payoffs can represent factors such as successful interception, damage inflicted, or mission completion.

3. Nash Equilibrium: The Nash Equilibrium represents a state where no player has an incentive to deviate from their chosen strategy, given the strategies of other players. In UAV defense, a Nash equilibrium can indicate stable defensive and offensive strategies where neither side can improve their payoff by changing their strategy alone.

4. Optimal Strategies: Game theory enables the identification of optimal strategies that maximize the expected payoffs or utilities for players. In the context of UAV defense, identifying optimal strategies can help in designing effective defensive tactics or countermeasure systems against adversarial UAVs.

5. Non-cooperative Game Theory: Characterized by strategic interactions among self-interested agents, offers insights into the decision-making processes in competitive settings. In the context of UAV defense, non-cooperative games can model the behavior and strategies of both defending UAVs and attacking UAVs, considering their conflicting objectives.

6. Cooperative Game Theory: On the other hand, this focuses on collaboration and cooperation among agents to achieve collective goals. Cooperative game-based approaches in UAV defense can model the coordination and cooperation among multiple defending UAVs in countering an attacking UAV swarm, considering their shared objectives.

Game theory provides a framework for modeling and analyzing the strategies and interactions of defending and attacking UAVs, enabling the assessment of potential threats and vulnerabilities in UAV defense scenarios. Furthermore, game theory-based decision support systems can assist in real-time decision-making for defending UAVs, considering the actions and intentions of attacking UAVs and optimizing defensive strategies to minimize damage and achieve mission objectives.

Researchers have extensively utilized game theory to model and analyze the interactions and behaviors of agents in various domains. Notably, Semsar-Kazerooni and Khorasani [67] proposed a game theory approach for multi-agent team cooperation. Their work explores consensus achievement over a common value using cooperative game theory, which can be applied to the coordination of multiple UAVs in countering adversarial threats. Parsons and Wooldridge [62] discuss the application of decision theory and game theory in multi-agent systems. This provides insights into how game-theoretic concepts can be leveraged to study the behavior of UAVs in complex scenarios involving strategic interactions. In their work, Kolokoltsov and Malafeyev [39] present an in-depth introduction to game theory and its applications in analyzing multi-agent systems with competition and cooperation. This resource offers foundational knowledge for researchers seeking to apply game theory to UAV defense.

By leveraging the principles of game theory, researchers and practitioners in the field of UAV defense can gain valuable insights into the strategic dynamics of UAV interactions and develop robust and effective countermeasure strategies.

2.4.2 Evolutionary Game Theory

Evolutionary game theory explores the evolution of strategies and behaviors in a population of agents over time. It can be applied to study the dynamics of UAV interactions, where defending and attacking strategies can undergo evolution and adaptation based on their success rates.

Evolutionary dynamics have been used to model and analyze the behavior of agents in various domains, including UAV interactions. Notably, Nowak and Sigmund [55] discuss the evolutionary dynamics of biological games, emphasizing frequency-dependent selection using game-theoretic arguments. This provides insights into understanding the evolution of strategies in a population of agents, which can be applied to evolving defending and attacking tactics of UAVs. Moreover, Sandholm [66] provides a comprehensive exploration of population games and evolutionary dynamics, offering insights into the mathematics and concepts behind the application of evolutionary principles to agent behavior.

Evolutionary optimization algorithms, such as genetic algorithms and particle swarm optimization, can be employed to search for optimal strategies or parameters in UAV defense. These algorithms mimic the principles of natural evolution and can lead to the discovery of effective countermeasure strategies against adversarial UAVs. Fogel [25] introduces simulated evolutionary optimization, a population-based optimization process that mimics natural evolution. Genetic algorithms, evolution strategies, and evolutionary programming are discussed, offering potential avenues to optimize UAV defense strategies. Zhao et al. [90] apply evolutionary optimization to server allocation for online cloud games, showcasing the broader applicability of evolutionary techniques to dynamic environments, which is relevant for UAV interactions. Fernández-Ares et al. [23] explore co-evolutionary optimization of autonomous agents in a real-time strategy game. Although not specific to UAV defense, this work highlights the potential of co-evolutionary dynamics to model the interactions and adaptation of agents.

Additionally, in the domain of multi-agent systems and autonomous robotics, Fehérvári and Elmenreich [22] introduced a novel approach for developing controllers for self-organizing robotic teams. Their work leveraged evolutionary algorithms and artificial neural networks to automate the design process, aiming to achieve desired global behaviors through local interactions among robotic agents. While their focus was on robotic teams, the concept of using advanced techniques to coordinate multiple agents in complex environments aligns with the broader field of multi-agent systems, which includes the coordination of UAV swarms. This research serves as a valuable reference for understanding the potential applications of machine learning and evolutionary algorithms in the context of multi-agent systems, which is a key aspect of the present study.

2.4.3 Multi-Agent Reinforcement Learning

Multi-Agent Reinforcement Learning involves the study of how multiple agents can learn to interact with their environment and make decisions to maximize their cumulative rewards. It builds upon the concepts of Reinforcement Learning, extending them to scenarios where multiple agents coexist, collaborate, or compete.

In MARL, the environment is often modeled as a Markov game, an extension of the single-agent Markov Decision Process (MDP). A Markov game captures the interactions between multiple agents, each having their own observations, actions, and rewards. Agents make decisions based not only on their observations but also on the observed actions of other agents. León and Belardinelli [42] introduce the concept of extended Markov games as a novel mathematical model for multi-agent reinforcement learning. The authors provide formal definitions, proofs, and empirical tests of RL algorithms within this framework. This work offers insights into learning diverse tasks through cooperation among agents.

Research has explored how agents can learn to cooperate to achieve common goals. For example, in multi-agent environments with sparse rewards, agents can use techniques like reward shaping or curriculum learning to encourage cooperative behaviors. In competitive scenarios, agents strive to outperform others. Researchers have investigated learning policies that anticipate opponents' strategies and adapt accordingly, leading to sophisticated competitive behaviors.

Notably, Yu et al. [86] delve into the surprising effectiveness of Proximal Policy Optimization (PPO) in cooperative multi-agent games. The study demonstrates that PPO-based multi-agent algorithms exhibit robust performance in diverse multi-agent testbeds, even outperforming competitive off-policy methods in certain cases. Through in-depth analysis and ablation studies, the authors highlight implementation and hyperparameter factors that contribute to PPO's empirical success. Furthermore, Nowé, Vrancx, and De Hauwere [56] present an exploration of the intersection of game theory and multi-agent reinforcement learning. Their work emphasizes the complexities arising when multiple agents apply reinforcement learning in shared environments.

MARL algorithms enable multiple agents to learn and adapt their strategies through interaction and feedback from the environment. In the context of UAV defense, cooperative-competitive MARL frameworks can be employed to train defending UAVs to coordinate their actions and counteract attacking UAV swarms effectively. Zhang et al. [87] propose efficient training techniques tailored for multi-agent reinforcement learning in combat scenarios, stemming from UAV combat situations. The authors present three distinct training methods, including scenario-transfer, self-play, and rule-coupled training, enhancing both convergence speed and performance of multi-agent Deep Q-learning and Deep Deterministic Policy Gradient methods. Their work contributes valuable strategies for improving MARL in adversarial contexts.

Decentralized MARL algorithms distribute decision-making among individual

UAV agents while maintaining coordination and cooperation. Decentralized approaches in UAV defense allow each defending UAV to make autonomous decisions based on local observations and communicate with other UAVs to achieve collective objectives. This approach resembles swarm intelligent agent behavior, which is described further in the subsequent section.

Multi-Agent Reinforcement Learning offers a powerful framework for studying how intelligent agents can interact in complex environments. By extending RL to multiple agents, researchers explore a wide range of behaviors, from cooperation to competition, and even hybrid forms of interaction.

2.4.4 Swarm Intelligence

Swarm intelligence refers to the collective behavior exhibited by decentralized, self-organized systems inspired by the behavior of natural swarms. It has gained significant attention in various domains due to its potential for solving complex problems through the interaction and cooperation of individual agents. In the context of UAV defense, swarm intelligence offers promising possibilities for developing effective countermeasures against adversarial UAV swarms. Kennedy [37] provides an overview of swarm intelligence, discussing its principles, concepts, and applications. This work offers a foundational understanding of swarm intelligence. Key concepts and principles in the field of swarm intelligence include the following points:

1. Emergence: Complex collective behavior emerges from the interactions of simple agents following local rules. The overall behavior of the swarm arises from the self-organization and adaptation of individual agents.

2. Decentralization: Swarm intelligence systems operate without centralized control or global knowledge. Agents make local decisions based on local information, leading to global coordination and problem-solving.

3. Interaction and Communication: Agents in a swarm interact and communicate through various mechanisms such as direct communication, indirect signaling, or the use of environmental cues. Communication enables the exchange of information and coordination among the agents.

4. Cooperation and Collaboration: Swarm intelligence relies on cooperation and collaboration among agents to achieve common objectives. Agents coordinate their actions, share information, and divide tasks based on their individual capabilities and the needs of the swarm.

5. Ant Colony Optimization (ACO): ACO is inspired by the foraging behavior of ants. Agents, representing artificial ants, deposit pheromones on paths, which influences the probability of other ants choosing the same paths. ACO has been used for combinatorial optimization problems.

6. Bee Colony Optimization (BCO): BCO models the behavior of honeybees in finding and exploiting flower patches. Agents, representing bees, use local search and global exploration to optimize solutions.

Swarm intelligence has found applications in various domains. Givigi Jr. and

Schwartz [27] propose a game-theoretic approach to swarm robotics, discussing the use of traits of personality to achieve swarm intelligent robots through reinforcement learning. This work demonstrates the integration of game theory and swarm intelligence. Moreover, Purnomo and Wee [64] present an innovative integration of evolutionary and swarm intelligence algorithms for soccer game optimization. This work showcases the synergy between evolutionary and swarm intelligence techniques.

Zhou et al. [92] provide a comprehensive survey of UAV swarm intelligence, reviewing advances and future trends. This work offers insights into the application of swarm intelligence to UAV systems, including decision-making, path planning, and communication. By leveraging the principles of swarm intelligence, it is possible to achieve efficient and adaptive swarm behaviors, such as formation flying, task allocation, or target tracking. Swarm intelligence techniques can be used for decentralized decision-making within UAV swarms. Agents in the swarm can collectively make decisions based on local observations and communicate to reach a consensus, enabling the swarm to respond effectively to dynamic environments. Swarm intelligence allows for the development of adaptive defense strategies against adversarial UAV swarms. By leveraging the collective intelligence and cooperative behavior of a defending UAV swarm, non-confrontational approaches, such as misleading, distracting, or delaying tactics, can be employed to neutralize attacks.

Swarm intelligence provides a promising avenue for developing efficient and adaptive countermeasures against adversarial UAV swarms. By leveraging the principles of emergence, decentralization, and cooperation, swarm-based approaches offer novel strategies for defending against UAV threats. Understanding swarm intelligence principles and their applications in UAV defense will contribute to developing effective and non-confrontational countermeasures in dynamic and adversarial environments.

Furthermore, by incorporating game-based approaches in AI, we can capture the strategic nature of UAV defense scenarios and develop intelligent decision-making systems for defending UAVs. These approaches offer the potential to enhance the coordination, adaptability, and robustness of UAV defense strategies in non-confrontational settings.

2.5 UAVs and Multi-Agent Intelligent Behavior in Games

In this section, we explore the intersection of Unmanned Aerial Vehicles (UAVs) and multi-agent intelligent behavior within the realm of games. Exploring topics such as games with a purpose, game AI, game engine development, and the gamification of UAV countermeasure training, we discuss how these elements converge to shape the future of UAV swarm coordination and countermeasures. Additionally, ethical considerations regarding game-based training are discussed.

2.5.1 Games with a Purpose and Simulations

Games with a purpose (in the literature oftentimes referred to as serious games) and simulations have gained prominence in the field of UAV defense. These games serve as valuable tools for training UAV operators, developing situational awareness, and testing different countermeasure strategies. Susi et al. [74] provide an overview of serious games, discussing their applications beyond entertainment. They highlight how serious games enable learners to experience situations that are otherwise infeasible, fostering skill development. This work explores various markets, including military, government, and healthcare games, showcasing the diverse applications of serious games. In the realm of UAVs, López et al. [50] explore the plausibility of using UAVs as a serious game for therapeutic purposes, specifically addressing Attention Deficit-Hyperactivity Disorder (ADHD). Through a systematic review, they investigate the coupling of UAVs with Brain-Computer Interfaces (BCIs) to control UAVs via mental commands. The paper identifies challenges and research opportunities in game design, interaction, and cognitive rehabilitation.

Simulation games are virtual environments designed to mimic real-world scenarios and allow users to interact and make decisions in a controlled setting. They have gained significant popularity and have been widely used in various fields, including UAV defense, to simulate real-world scenarios and test different strategies. By accurately modeling the characteristics of UAVs, their capabilities, and the operational environment, simulation games offer a valuable platform for assessing the effectiveness of various countermeasures. Moreover, simulation games can serve as a valuable tool for validating and benchmarking the performance of UAV defense systems. By comparing the outcomes of simulated scenarios with known ground truth or empirical data, researchers can assess the accuracy and reliability of the simulated environment and the effectiveness of the implemented countermeasures. This validation process helps establish the credibility and robustness of the simulation games as a testing platform. Notably, Atkin et al. [3] explore the convergence of military simulation and computer games, specifically focusing on the "Capture the Flag" game. They introduce a Hierarchical Agent Control architecture that supports action abstraction and multi-level computational architecture. The authors propose that game design and AI could mutually benefit from each other's technology. Stone et al. [73] present an insightful retrospective overview of aerospace simulation evolution, progressing from immersive Virtual Reality to serious games. They discuss advancements in simulation hardware and software technologies, emphasizing the application of immersive technologies for visualization, training, and research in the aerospace industry.

Furthermore, simulation games offer an interactive and immersive environment for training operators and decision-makers involved in UAV defense. By simulating diverse attack scenarios and presenting players with challenging situations, simulation games enable individuals to develop their skills in detecting, tracking, and countering adversarial UAVs. This training aspect helps enhance situational awareness, decision-making abilities, and response times in real-world UAV defense operations. Cardona-Reyes et al. [9] address the training of drone pilots through virtual reality environments using a gamification approach. They emphasize the advantages of using virtual reality for drone training, particularly in urban environments where real drone training involves risks. The collaborative creation of virtual environments, involving software designers, drone experts, and educators, enhances the learning experience for acquiring drone piloting skills.

The influence of serious and simulation games in the field of UAV defense is evident in their ability to replicate realistic scenarios, facilitate training and skill development, enable experimental testing and analysis, foster collaboration, and provide a means for validation and benchmarking. As the field of UAV defense continues to evolve, simulation games are expected to play an increasingly important role in the development and evaluation of effective countermeasures.

2.5.2 Game AI and Autonomous Agents

Game AI and autonomous agents are essential components in developing realistic and intelligent behaviors for UAVs and adversaries in game-based UAV defense systems. Research in this area focuses on developing algorithms and techniques for creating autonomous agents that can make intelligent decisions, adapt to dynamic situations, and exhibit realistic behaviors. Key works in this field include the book by Millington [51], "AI for Games," which provides a comprehensive resource on integrating artificial intelligence techniques into video games. This work explores AI concepts and strategies used in game design and development, providing insights into creating intelligent and engaging game experiences through the integration of AI technologies. Furthermore, the paper by Shao et al. [68], provides a comprehensive survey of deep reinforcement learning (DRL) in video games, highlighting its significant achievements and impact on game artificial intelligence. They explore various DRL methods, including value-based, policy gradient, and model-based algorithms, while discussing their applications in diverse video game genres, from Arcade to real-time strategy games. The authors also address challenges when applying DRL to gaming contexts and outline potential research directions. Earlier works in the field include the presentation of Orkin [59] of AI in the video game *F.E.A.R. First Encounter Assault Recon* [53], showcasing how AI technologies can be applied to enhance player experiences. This presentation at the Game Developers Conference delves into the AI design of *F.E.A.R.*, discussing the implementation of AI-driven behaviors and interactions to create immersive and challenging gameplay.

2.5.3 Game Engine Development

Game engines are fundamental tools for creating interactive and realistic environments. In the context of UAV defense, game engine development plays a crucial role in building simulation environments that accurately represent the dynamics of

UAV interactions and countermeasures. Noteworthy resources in this area include the work by Gregory [29] which contributes significantly to the field of game development with the book *Game Engine Architecture*, providing a comprehensive guide to the principles and concepts underlying the design and implementation of game engines. This work explores key architectural components, rendering systems, resource management, and more, offering insights into the complex process of building game engines.

Cowan and Kapralos [14] present a survey that addresses the challenges faced by serious game developers in choosing appropriate game engines and frameworks. Their research highlights the shortage of standard tools for serious game development and discusses the usage of entertainment-based game development tools for serious applications. The authors identify commonly used frameworks and engines, offering valuable insights for developers seeking tools for serious game creation, which is of relevance for incorporating UAVs in serious games.

In the realm of UAVs, there have been multiple works that utilize game engines for UAV simulations. Hu and Meng [33] introduce ROSUnitySim, a novel real-time simulator that enables local planning for miniature UAVs in complex environments. This simulation system combines the Robot Operating System (ROS) with the Unity3D game engine to create a versatile platform for testing flight control and navigation algorithms. The authors emphasize the advantages of using Unity3D and ROS for UAV simulations and the potential for saving effort in flight tests. Yong-kang, Yong, and Daquan [85] also present a UAV simulation training and assessment system designed using the Unity3D game engine. Their system combines virtual simulation technology with UAV maintenance training, providing an interactive learning environment for UAV equipment training and teaching. The authors describe the system's modular design, interactions, and its ability to meet the performance requirements of UAV training and assessment. Wang et al. [82] contribute to the construction of a virtual reality platform for UAV deep learning using the video game engine Unreal Engine 4 (UE4) and AirSim. This platform enables the acquisition of images and data for UAV deep learning research by simulating realistic environments and scenarios. The authors highlight the platform's capability to provide real-time, accurate visual and sensory feedback for advancing UAV autonomous flight and obstacle avoidance research. Furthermore, Bondi et al. [6] present AirSim-W, a simulation environment designed specifically for wildlife conservation applications using UAVs. This environment integrates computer vision techniques and thermal infrared models to assist in monitoring wildlife and poaching activities. The authors detail the creation of an African savanna environment in Unreal Engine, its integration with thermal infrared models, and the potential for using simulated data to improve detection algorithms and autonomous flight strategies.

2.5.4 Game Design and Mechanics

Game design principles play a crucial role in developing effective game-based approaches for UAV defense. Understanding the mechanics of game design, including player engagement, motivation, and interaction, can inform the design of engaging and immersive experiences for training UAV agents and simulating adversarial scenarios.

Salen Tekinbas and Zimmerman [76] present an influential work titled *Rules of Play: Game Design Fundamentals*, offering a comprehensive exploration of the foundational principles and mechanics that underpin the design of games. The authors delve into essential aspects of game design, including game mechanics, dynamics, aesthetics, and interaction, providing a deep understanding of how games create engaging experiences for players. The book serves as a valuable guide for understanding the intricate interplay between design choices and player engagement within the realm of game development. The principles of game design extend beyond entertainment and can be applied to serious scenarios like UAV defense. By leveraging the mechanics that create engaging and immersive experiences in games, developers can create training environments that enhance the skills of UAV operators and foster effective multi-agent cooperation. The article “Gamification science, its history and future: Definitions and a research agenda” [40] delves into the concept of gamification, which involves applying game design elements to non-game contexts. It discusses how gamification principles, such as points, badges, and leaderboards, can be utilized to engage users and motivate desired behaviors. In the context of UAV defense, gamification elements could be integrated to potentially enhance training effectiveness.

The fusion of game design principles with UAV defense not only enhances the engagement of training environments but also cultivates effective multi-agent behaviors among UAVs. By utilizing game mechanics, UAV defense training scenarios can mimic complex real-world scenarios, enabling operators to develop the skills and strategies necessary for successful defense operations.

2.5.5 Gamification of UAV Countermeasure Training

Gamification techniques have been increasingly employed in UAV countermeasure training to enhance user engagement, motivation, and learning outcomes. By incorporating game elements, such as challenges, rewards, leaderboards, and progress tracking, the training experience becomes more immersive and enjoyable. Works like the paper by Landers and Landers [41], "An Empirical Test of the Theory of Gamified Learning: The Effect of Leaderboards on Time-on-Task and Academic Performance," provide insights into the impact of gamification on learning outcomes and engagement. Applying gamification principles to UAV countermeasure training can potentially improve operator performance, decision-making abilities, and overall effectiveness. Smith [71] delves into the historical intertwining of gaming within both military training and entertainment realms in the article "The long history of gaming in military training". The article explores the utilization of various gaming formats, including sand tables, miniatures, board games, and computer games, to achieve training objectives. The author highlights the longstanding relationship between dynamic representations of the physical world and their diverse applications in fields such as military training, education, city planning, and entertainment.

2.5.6 Depictions of UAVs in Video Games

This subsection delves into a selection of games that showcase the diverse ways UAVs are portrayed in video games. Such as LiftOff [48], which is a drone simulation game that focuses on recreating the experience of flying FPV (First Person View) racing drones. It offers a wide range of customizable drones and realistic physics that mimic real-world drone flight. Players can practice drone racing, and freestyle maneuvers, and explore various environments. Similar to LiftOff, VelociDrone [4] is another FPV drone simulator that allows players to practice and refine their piloting skills in a virtual environment. It features a variety of tracks and environments for drone racing and freestyle flying. VelociDrone is an example of how realistic drone simulations contribute to training pilots and understanding drone behavior, which can have implications for UAV defense.

Different than games designed to feature realistic drone simulations, drones are also represented in mainstream video games, usually within futuristic contexts. They can be used in the context of providing environmental storytelling by enhancing aesthetics, or they can be included in gameplay and combat. One such example is *Cyberpunk 2077* [10] developed by CD Projekt Red, which presents a dystopian future where advanced technology is seamlessly integrated into society. Within this world, various types of UAVs are a common sight, reflecting the game's cyberpunk aesthetic. These UAVs serve a multitude of functions, including surveillance, transportation, and even combat. Players encounter surveillance drones that monitor city streets, delivery drones navigating through neon-lit skyscrapers, and combat-oriented UAVs deployed in intense firefights. The game's portrayal of UAVs offers a glimpse into a potential future where these devices are an integral part of urban life, raising questions about privacy, control, and the implications of their widespread use.

Another popular game, *Valorant* [65], developed by Riot Games, is a tactical first-person shooter that incorporates a diverse array of special agents, each equipped with unique abilities. Among these abilities, UAVs play a strategic role, as player characters, referred to in-game as agents, such as one called Sova, possess the ability to deploy UAV drones, allowing players to scout enemy positions and gather crucial intelligence. The UAVs in *Valorant* emphasize the importance of information warfare in gameplay, where mastering drone usage can lead to a significant advantage. By incorporating UAVs into the game's mechanics, *Valorant* highlights the growing influence of drone technology in modern combat scenarios and underscores the value of reconnaissance in tactical decision-making.

The portrayal of UAVs in video games contributes to broader conversations about their roles in various sectors, ranging from urban surveillance to combat tactics. These games provide a unique opportunity to engage players in discussions about the implications, both positive and negative, of UAV integration into our real and virtual worlds.

2.5.7 Ethics of Game-based Training

The use of game-based approaches in UAV defense training raises important ethical considerations. As these systems involve simulating adversarial scenarios and training UAV agents to engage in combat-like activities, it is crucial to address ethical concerns related to the realistic portrayal of violence, psychological impact on users, and potential desensitization.

Olson and Rashid [57] present an ethical analysis titled "Modern Drone Warfare: An Ethical Analysis," discussing the intricate ethical dilemmas arising from the utilization of drone technology in military operations. The authors examine the potential consequences and moral implications of drone warfare, acknowledging its benefits in safeguarding soldiers' lives but also addressing concerns about the dehumanization of warfare and the disconnect between soldiers and the battlefield. They underline the importance of evolving policies and strategies in alignment with the changing landscape of defense technology.

Leonard [43] explores the intersection of video games, the military, and pedagogy of peace in the article "Unsettling the military entertainment complex: Video games and a pedagogy of peace." In the context of post-9/11 American policy shifts, the author delves into the role of video war games in shaping public perceptions and attitudes toward warfare. Highlighting the potential influence of virtual war games on supporting military endeavors and imperialism, Leonard emphasizes the significance of critically analyzing the impact of video games as a pedagogical tool in shaping societal perspectives on war.

The integration of UAVs and multi-agent intelligent behavior in games provides a strong foundation for designing and implementing game-based approaches in UAV defense. By drawing upon the principles, theories, and techniques from these fields, researchers can develop innovative solutions that enhance training, decision-making, and the overall effectiveness of UAV countermeasures.

Chapter 3

Methodology

In this chapter, we will discuss the overall research design, as well as specific procedures employed within the thesis in order to achieve the stated earlier research objectives, primarily the generation of efficient and non-confrontational UAV countermeasures using ML techniques. We will delve into the specifics of ML approaches used for the study, discuss the developed game-based approaches and the baseline approach, and describe the agent behavior and the environmental details, as well as the evaluation metrics used for the testing of the generated countermeasures.

3.1 Research Design

The research design is based on the analysis of the related literature presented in the previous chapter, as well as the research objectives of the thesis. The availability of resources was also taken into consideration during the research design stage.

The analysis of the related literature covered in the previous chapter has provided valuable insights with regard to how to approach the research design, which methodologies to incorporate, and which gaps exist in the current UAV countermeasures research.

Firstly, the overview of existing countermeasures displays the need for countermeasures that prevent confrontation, as well as avoid potential collateral damage. This puts the research design into focus, setting an overarching goal, which is reflected in the problem statement of the thesis as well.

Additionally, the review of ML methodologies has narrowed down the area of machine learning to reinforcement learning approaches. The cooperative-competitive nature of the problem yields itself well for the design of MARL scenarios. Furthermore, the previous applications of MARL algorithms to UAV countermeasures problems have pointed out the benefits of using the MADDPG algorithm. The MADDPG algorithm and the reasoning behind the choice are examined in greater detail in the following section of the Methodology chapter.

Further analysis of the related research and areas of interest have indicated the correlation between MARL methods and Game-based approaches, which further solidifies the choice of using the MADDPG algorithm for our research and enables the design of game-based scenarios. This further aligns with the problem statement and research objectives.

3.2 Machine Learning Methodology

In this section, we look at machine learning methodology, which is a vital aspect of our research. We explore various facets of machine learning, including the reinforcement learning approach, deep learning techniques, actor-critic methodologies, policy gradient methods, and the Multi-Agent Deep Deterministic Policy Gradient algorithm. These techniques form the foundation upon which our non-confrontational countermeasures against UAV swarms are built, demonstrating the fusion of AI and game-based principles in addressing real-world challenges.

3.2.1 Reinforcement Learning Approach

Reinforcement Learning refers to a type of Machine Learning that focuses on intelligent agents that learn by interacting with and getting feedback from their environment. These agents perform actions and receive a reward from the environment based on the impact of that action. RL has gained popularity in recent years due to its potential for versatility and efficiency. Multi-Agent Reinforcement Learning is a type of RL that deals with multiple agents that learn to achieve objectives in a cooperative and/or competitive manner. MARL has been applied multiple times within the context of UAV research, including UAV countermeasures research (see Chapter 2). MARL is a fitting choice for generating UAV countermeasures since the UAVs could be easily mapped to RL agents interacting within and receiving feedback from an environment.

In our case, MARL is an appropriate solution as well, since we could model both the adversarial and the defending UAVs as RL agents with different objectives. The objective of the defending agents is to mislead, distract, or delay the adversarial UAVs from a certain target, while the objective of the adversarial UAVs is to reach the said target. Thus two distinct teams form that exhibit cooperative-competitive behavior. Each team works together to accomplish its unique objective while competing against the other team. The overarching objectives of reaching and protecting the target can then be further divided into multiple smaller objectives that can be expressed in terms of rewards. The environment can be portrayed as an abstract virtual environment containing the defending team, the adversarial team, and the target.

MARL introduces unique challenges compared to single-agent RL. Agents' actions can affect the environment and the observations of other agents, leading to non-stationarity. The presence of other learning agents adds complexity to the environment, making it essential to design algorithms that enable agents to collaborate and compete effectively. In further subsections, we will look into the specific MARL methods and an algorithm that we incorporated into our work.

3.2.2 Deep Learning

Deep Learning is a subfield of machine learning that has revolutionized various domains, including computer vision, natural language processing, and, more recently, reinforcement learning. At the core of deep learning are neural networks, which are computational models inspired by the structure and function of the human brain. These networks consist of multiple layers of interconnected nodes, also known as neurons. Each layer extracts increasingly abstract and hierarchical representations of the input data, allowing the network to learn intricate and meaningful patterns. This makes them a useful tool in many areas, including UAV behavior research.

Deep Reinforcement Learning (DRL) combines both the RL and DL methodologies, where the agents are able to learn complex decision-making strategies and behaviors in order to achieve the highest cumulative reward. The policy function is responsible for mapping the states of the environment to the actions of agents. In other words, it defines the behavior or strategy of agents in the environment. Neural Networks are used to approximate the policy function since oftentimes times policy functions can be extremely complex. The neural network takes the current state as input and outputs a distribution over possible actions. The agent then selects actions based on this distribution. Neural networks can adjust their internal parameters based on the data they receive during training. As the agent interacts with the environment and receives feedback, the neural network gradually refines its policy to improve decision-making. Neural networks can generalize from the data they have seen during training to make decisions in unseen situations. This is particularly valuable in RL, where the agent needs to make good decisions in a variety of scenarios. Furthermore, Neural networks can represent policies that have intricate and non-linear relationships between states and actions. This is crucial for handling potentially complex environments. Instead of explicitly programming each rule, an RL agent equipped with a neural network policy can learn from interactions with the environment. The neural network can generalize from the data it collects to make informed decisions, even in complex and dynamic situations. This makes RL agents with learned policies more adaptable and effective in a wide range of scenarios, including UAV countermeasures generation, where the agents need to process vast amounts of sensory information and make real-time decisions.

While deep learning offers significant advantages, it also presents challenges, such as vanishing or exploding gradients, overfitting, and high computational demands. To address these issues, regularization techniques, batch normalization, and advanced optimization methods like Adam have been employed to stabilize training and improve convergence.

3.2.3 Actor-Critic Approaches

Actor-Critic methods are a popular class of RL algorithms that have gained widespread use in various applications due to their effectiveness in handling both continuous and discrete action spaces. The distinguishing feature of Actor-Critic methods is the presence of two distinct neural networks—the actor and the critic.

The actor-network is responsible for learning the policy function, which maps states to actions. In the context of our UAV countermeasure generation, the actor-network enables the UAV agents to make informed decisions on which actions to take in response to specific environmental states. The policy function is often represented as a deep neural network, allowing the agents to learn complex and non-linear strategies.

The critic network, on the other hand, is responsible for estimating the value function, which quantifies the expected cumulative reward the agent can achieve from a given state following the learned policy. The value function acts as a feedback signal for the actor-network, guiding policy updates to improve the agent's decision-making capabilities over time. By learning the value function, the critic network provides a measure of the desirability of different states, allowing the agent to differentiate between good and bad actions in a given context.

Actor-Critic methods combine the advantages of both value-based and policy-based RL approaches. They offer improved stability and faster convergence compared to policy-based methods while being more sample-efficient than pure value-based algorithms like Q-learning. Moreover, actor-critic architectures enable the agents to learn decentralized policies, making them suitable for multi-agent settings like UAV swarm defense, where individual agents must coordinate without global knowledge.

3.2.4 Policy Gradient Methods

Policy Gradient Methods are a popular category of Deep Reinforcement Learning. In DRL, each agent is typically equipped with at least one neural network, which serves as an approximation for the agent's policy function. The objective of Policy Gradient Methods is to directly enhance the policy function by employing Gradient Ascent in the policy space. Gradient Ascent is a technique in Machine Learning that optimizes a function to maximize its output. In Policy Gradient Methods, agents learn a policy by optimizing it to achieve the highest possible reward. This approach is well-suited for complex problems where agents have a wide range of possible actions to perform, such as problems with a continuous or high-dimensional action space.

This methodology is particularly fitting to our problem since it ensures the possibility of creating intricate flexible and adaptive strategies. Optimizing the policy directly allows for the generation of UAV countermeasures that are suitable for environments of various complexity.

3.2.5 Multi-Agent Deep Deterministic Policy Gradient

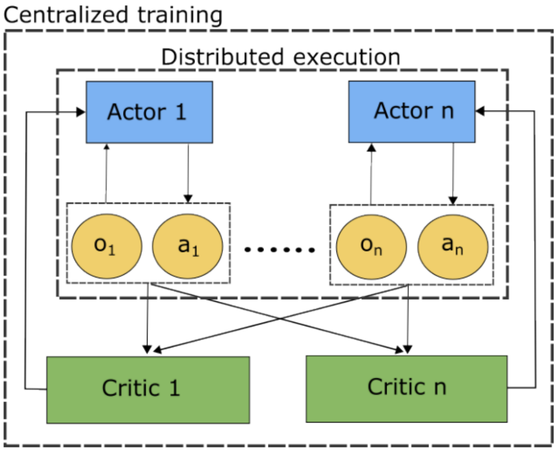


Figure 3.1: Illustration of the MADDPG algorithm. Initially introduced in [69].

The Multi-Agent Deep Deterministic Policy Gradient algorithm is a powerful approach for multi-agent reinforcement learning. It builds upon the ideas of Deep Deterministic Policy Gradient (DDPG) [45] and extends them to environments with multiple interacting agents. In our research on UAV countermeasure generation, we adopt MADDPG to train cooperative-competitive UAV agents and develop effective countermeasure strategies. The following features make the MADDPG algorithm particularly suitable for our problem:

1) Learning policies that only utilize local information during execution, 2) No assumption of a differentiable environment dynamics model or specific constraints on inter-agent communication methods, 3) Applicability not only within cooperative scenarios but also in competitive or mixed interactions, including both physical and communicative aspects

In MADDPG, the algorithm employs a centralized critic and decentralized actor architecture. Each agent maintains its own policy, known as the decentralized actor, which selects actions based on its observations of the environment. A centralized critic is used to estimate the state-value function, which considers the joint actions of all agents. This approach allows the agents to collaborate and learn from each other while maintaining individual policies. During training, every agent's critic network requires access to teammates' observations and actions, establishing a form of implicit information sharing that promotes effective coordination among agents. Consequently, MADDPG undergoes training using a global state (centralized training), while actual execution requires only local observations (decentralized execution). Figure 3.1 presents a schematic overview of the multi-agent actor-critic framework. In our approach, both attackers and defenders harness MADDPG to fulfill their objectives (attackers aim to reach a static target, while defenders try to prevent them from doing so).

Similar to DDPG, MADDPG utilizes experience replay and target networks to stabilize training and improve sample efficiency. Experience replay stores agent experiences in a replay buffer and samples batches of experiences randomly during training. Target networks are used to provide stable target value estimates by periodically updating the target networks with the current actor and critic networks.

MADDPG is known for its robustness in dealing with non-stationarity, a common challenge in multi-agent environments where the actions of one agent directly affect the environment observed by others. The use of target networks and experience replay helps MADDPG handle non-stationary environments and achieve stable learning.

MADDPG's ability to handle multi-agent environments and its cooperative-competitive training setup make it a promising algorithm for UAV countermeasure generation. By learning joint strategies through interaction, the UAV agents can effectively coordinate their actions to defend against adversarial UAV swarms and adapt to changing scenarios.

3.3 Game-based Approach Methodology

Incorporating game elements into the reinforcement learning and multi-agent reinforcement learning framework introduces an inherent gamification aspect to the training process. This section explores how RL and MARL inherently lend themselves to a gamified environment and how the scenarios created for the agents embrace gaming concepts.

Reinforcement learning, by its nature, mirrors the structure of a game. Agents learn optimal strategies by iteratively interacting with an environment, aiming to maximize cumulative rewards. This iterative process aligns with the progression of levels or stages commonly found in games, where players seek to achieve higher scores or unlock new challenges. In RL, the agent's "score" is the cumulative reward obtained over time, and the "challenges" correspond to learning increasingly complex policies to navigate the environment effectively.

Multi-agent reinforcement learning introduces competitive and cooperative dynamics reminiscent of multiplayer games. In the proposed framework, the defending and attacking UAV swarms form distinct teams that engage in cooperative-competitive interactions. The defending team cooperatively works to mislead, distract, or delay the adversarial UAV swarm from reaching its target, while the adversarial team strives to accomplish its mission. This setup creates a gaming-like scenario where two teams compete against each other while working towards their respective objectives, akin to the dynamics seen in multiplayer online games, especially arena shooters, such as *Valorant* [65], *Overwatch* [5], *Team Fortress 2* [78], *Counter-Strike: Global Offensive* [79], etc.

The scenarios crafted for the agents' training can be viewed as different levels within a game. Each scenario presents unique challenges and objectives, serving as analogs to the levels players encounter in video games. As the agents progress through these scenarios, they refine their strategies and decision-making skills, akin to players honing their skills to beat increasingly difficult game levels. This design allows for a structured progression that gradually introduces complexities and encourages the development of versatile countermeasure strategies.

The interaction between agents and their environment, as well as the feedback loop of rewards, align with the core mechanics of gaming. In the proposed framework, agents take actions in response to the environment, similar to how players make decisions based on in-game situations. The rewards obtained by agents serve as a form of feedback, guiding them to learn and improve their strategies, analogous to players receiving feedback on their performance through in-game scores or achievements.

The gamified nature of the framework enhances engagement and adaptability. Agents are driven by the pursuit of higher rewards, mirroring the motivation players have to achieve higher scores or complete game objectives. This engagement leads to agents developing innovative and adaptable strategies, essential traits in both gaming and countering adversarial UAV swarms.

By incorporating this game-based perspective, the RL and MARL framework not only provides a structured training environment but also leverages the intrinsic motivation and engagement seen in gaming scenarios. This approach aligns with the overarching goal of developing efficient and dynamic countermeasure strategies for UAV swarms while harnessing the power of gamification to enhance the learning process.

3.4 Multi-Agent Particle Environment

The simulation environment plays a pivotal role in studying and evaluating the interactions between defending and attacking UAV agents. For our research, we leverage the Multi-Agent Particle Environment (MPE), a versatile platform that facilitates the development, testing, and assessment of multi-agent reinforcement learning algorithms. The MPE is designed to create various scenarios involving multiple agents, enabling us to simulate and analyze the behaviors of both defending and attacking UAVs within controlled settings.

The MPE framework is built upon the OpenAI Gym library, a widely used toolkit for developing and comparing reinforcement learning algorithms. With the MPE, we can define intricate multi-agent scenarios that capture the essence of UAV swarm dynamics, enabling us to investigate the effectiveness of different defense strategies and approaches.

In the MPE environment, agents are the autonomous entities that make decisions and interact within the simulated world. We model both defending and attacking UAVs as individual agents, each equipped with distinct characteristics, objectives, and strategies. These agents possess localized observations and are capable of executing actions that influence the environment based on the information they perceive.

Defending UAV agents are tasked with devising effective strategies to counteract the adversarial UAV swarm. Their primary objective is to mislead, distract, or impede the progress of the attacking UAVs in their pursuit of a designated target. To accomplish this, defending agents rely on their observations of the environment, including their positions, distances to the target, and relative distances to other agents. This localized information informs their decision-making process as they navigate the virtual space to execute actions that align with the overarching defense mechanism.

On the other side, attacking UAV agents are driven by the goal of reaching the predefined target. These agents exhibit behaviors aimed at circumventing the defense strategies and successfully reaching their objective. Similar to defending agents, attacking UAVs rely on localized observations to inform their actions. Their decision-making process is guided by the pursuit of optimal trajectories that enable them to navigate the environment and overcome the defensive measures deployed by the defending UAV agents.

In the MPE, each UAV agent possesses distinct observation and action spaces, shaping how they perceive and interact with the environment. Defending and attacking UAVs receive local observations that include information about their positions, distances to the target, and relative distances to other agents. These observations allow the agents to make informed decisions based on their immediate surroundings.

The MPE environment allows us to model the complex interactions between agents. Our defending UAV agents aim to mislead, distract, or delay the attacking UAV swarm, while the attacking UAVs strive to reach a predefined target. These opposing objectives create a competitive-cooperative dynamic, where each team collaborates internally while competing against the other. To guide the learning process, we define reward functions that provide feedback to the agents based on their actions and outcomes. These rewards encourage UAVs to adopt strategies that align with the overarching goals of the defense mechanism.

Utilizing the MPE environment, we simulate various scenarios involving different swarm sizes and initial conditions. The MPE facilitates the execution of multiple episodes, allowing us to collect data on agent behaviors, reward trajectories, and convergence patterns. We analyze the performance of our proposed defense mechanisms by evaluating key metrics such as win rates, convergence speed, and efficiency in countering the attacking UAV swarm.

By exploiting the capabilities of the MPE environment, we are able to comprehensively explore the intricate dynamics of UAV swarm interactions and validate the effectiveness of our novel defense strategies against adversarial UAVs.

3.5 Game-based Defense Strategies Overview

In the area of UAV swarm countermeasures, the concept of applying game-based strategies is a novel and promising concept. Drawing inspiration from game theory and the principles of strategic decision-making, these defense mechanisms leverage competitive and cooperative interactions between defending and attacking UAV agents. By framing the confrontation between defenders and adversaries as a strategic game, we introduce an innovative paradigm that enhances the robustness and adaptability of UAV defense.

The communication-based defense strategy introduces an element of deception and misinformation into the adversarial UAV swarm scenario. Defenders capitalize on the capability to broadcast messages to infiltrate the attacking UAV swarm, introducing false information and manipulating its trajectory. This approach involves the creation of a decoy or "fake" target, alongside the real objective. Defending agents mimic attacking behavior and transmit messages about the fake target to mislead the adversaries.

In the communication-based approach, defenders assume the role of misinformation agents. They strategically choose to broadcast information to the attacking swarm, aiming to lure UAV adversaries away from the true target. This concept mirrors real-world strategies employed in deception warfare, where the dissemination of false intelligence is utilized to divert and confuse enemy forces.

The Dynamic Shield Approach focuses on the physical protection of the target by strategically positioning defenders around it. The defenders establish a dynamic shield that repels attackers, preventing them from reaching the target. This approach offers an alternative to communication-based methods, relying on direct spatial interactions between agents.

In the dynamic shield approach, defending agents transform into protectors of the target. Their actions focus on strategic positioning and coordinated movement to intercept and repel adversarial UAVs. By forming a barrier around the target, defenders exploit their physical presence to deter and obstruct the attacking swarm's progress. The dynamic shield mechanism introduces a tangible and visual defense tactic, reminiscent of military formations employed to safeguard critical assets.

Both the communication-based and dynamic shield strategies exhibit distinct strengths and trade-offs. The former relies on information warfare and strategic deception, leveraging misinformation to divert adversaries. In contrast, the latter emphasizes physical protection and coordinated movement to physically intercept and deter attackers. The choice between these strategies necessitates a careful evaluation of the scenario's characteristics, including swarm size, environment complexity, and the availability of communication channels.

In the subsequent sections, we delve into the technical details of implementing these game-based defense strategies within the MPE. By employing these approaches, we attempt to disrupt and neutralize adversarial UAV swarms while exploring the dynamics of strategic interactions in UAV defense scenarios.

3.6 Evaluation Metrics

A robust evaluation framework is essential to assess the effectiveness and performance of the devised UAV defense strategies within the MPE. To quantitatively measure the impact of the defense mechanisms and evaluate their success in countering adversarial UAV swarms, we employ a set of evaluation metrics. These metrics offer insights into different facets of defense strategy performance, enabling a comprehensive assessment of their capabilities.

- Winning Probability: The winning probability metric serves as a primary indicator of strategy success. It measures the likelihood of the defending agents achieving their objectives, such as diverting the attacking UAV swarm or preventing them from reaching the target. By analyzing the frequency of successful outcomes over multiple simulation runs, we gain a clear understanding of the defense mechanisms' efficacy.

- Average Reward: The average reward metric provides insight into the overall performance of the agents throughout the learning process. It reflects the cumulative impact of agents' actions and interactions with the environment over the course of training. Monitoring the trajectory of average rewards across episodes enables us to identify convergence and stability points, indicating the point at which the agents have acquired optimal strategies.

- Sensitivity to Swarm Size: The sensitivity of defense strategies to varying swarm sizes is a critical aspect of evaluation. To ascertain the scalability and adaptability of the proposed mechanisms, we examine their performance across a range of swarm sizes. This analysis provides insights into whether the defense strategies maintain their effectiveness when confronted with larger or smaller adversarial UAV swarms.

- Comparative Analysis: In the comparative analysis, we benchmark the performance of the devised defense strategies against a baseline approach. The baseline approach adopts the dynamic shield concept of exploiting physical repulsion, where defenders move in a random manner by choosing the next action randomly from the action space. By contrasting the outcomes and efficiency of our communication-based and dynamic shield strategies with the baseline approach, we gain a comprehensive perspective on their relative advantages and limitations.

In the subsequent sections, we present the results of our evaluation using these evaluation metrics. By extensively analyzing the performance of the communication-based and dynamic shield defense strategies, we aim to provide a thorough assessment of their capabilities and contributions to the field of UAV swarm countermeasures.

3.7 Method Limitations

While our methodology provides valuable insights into UAV swarm defense strategies, it is important to acknowledge certain limitations inherent to our approach. These limitations may influence the generalizability and scope of our findings.

- Abstracted 2D Environment: Our research employs the Multi-Agent Particle Environment, an abstracted 2D environment in which agents and targets are represented as particles. While this abstraction allows us to focus on agent interactions and strategic behaviors, it may not fully capture the complexity and dynamics of real-world UAV swarm scenarios, which could involve three-dimensional movement and environmental features.

- Limited Scalability and Computational Cost: Training and simulating a large number of agents in complex environments can incur significant computational costs. Our approach faces challenges when scaling up to a large number of agents, limiting the scope of our experiments and potentially affecting the real-time feasibility of certain defense strategies.

- Assumption of Defender Knowledge: Our methodology assumes that defending agents have complete knowledge of the positions of attacking agents. This assumption simplifies the decision-making process for defenders and may not hold true in scenarios where real-time communication or accurate enemy tracking is challenging.

- Discrete Actions and Observations: The discrete nature of agent actions and observations in our methodology can impact the precision and realism of agent behaviors. Complex real-world actions, which involve continuous and nuanced movements, may not be fully captured within our discrete action space.

- Rewards and Collision Avoidance: Designing reward functions to encourage desired agent behaviors, such as avoiding collisions between UAVs, can be challenging. Our methodology incorporates reward shaping to mitigate UAV collisions, but the effectiveness of these reward models may vary depending on the specific defense scenarios and environments.

- Applicability to Real-World Dynamics: While our research contributes insights into UAV swarm defense strategies, the transferability of our findings to real-world applications may be influenced by the simplifications and assumptions made in our model. Real-world dynamics, such as wind conditions, sensor limitations, and communication constraints, could introduce additional complexities that are not fully captured in our methodology.

- Algorithm Limitations: Our choice of the MADDPG algorithm, while suitable for our research goals, has inherent limitations. MADDPG can face challenges in training large swarms of agents efficiently and may require careful hyperparameter tuning to achieve optimal results.

These limitations should be considered when interpreting the results and implications of our study. While our methodology offers valuable insights into UAV swarm defense strategies, addressing these limitations could be the focus of future research endeavors to enhance the applicability and robustness of our approach.

Chapter 4

Design and Implementation

This chapter delves into the process of designing and implementing our UAV countermeasure system. It begins with framing the problem, designing the solution, and creating a game-based framework. We then explore the simulation environment, define action and observation spaces, and detail agent behavior. Next, we examine the game elements and discuss the practical implementation of defense strategies, including training, the Communication-based approach, and the Dynamic Shield approach. Challenges related to problem complexity, environment setup, training length, and initial design issues are addressed. This chapter also sets the stage for evaluation as we implement evaluation metrics to assess our game-based UAV countermeasure system.

4.1 Problem Framing and Solution Design

In this section we discuss the conceptualization of the problem, preliminary research in UAV countermeasures and Machine Learning, and the decision to frame the challenge as a game, laying the groundwork for techniques like Multi-Agent Reinforcement Learning in defense strategy development.

4.1.1 Conceptualization and Problem Statement

As my first assignment as a researcher in the field of Swarm Intelligence, I took part in a project aimed at merging the areas of UAV countermeasures and Swarm Intelligence. This project later became closely related to the topic of my master's thesis. At the time of the conceptualization of the project, two potential approaches were considered: "Inducing UAVs" and "UAV Swarm vs. UAV Swarm."

The "Inducing UAVs" potential approach involved deploying a small group of defending drones, most probably two or three, their objective being to disrupt an incoming swarm, either by diverting their attention or delaying them. Initial considerations included swarm algorithms like Grey Wolf Optimizer (GWO) and Slime Mold Algorithm. Key Performance Indicators (KPIs) focused on aspects such as the number of defending UAVs, possibility to mislead, duration of disruption, and response speed.

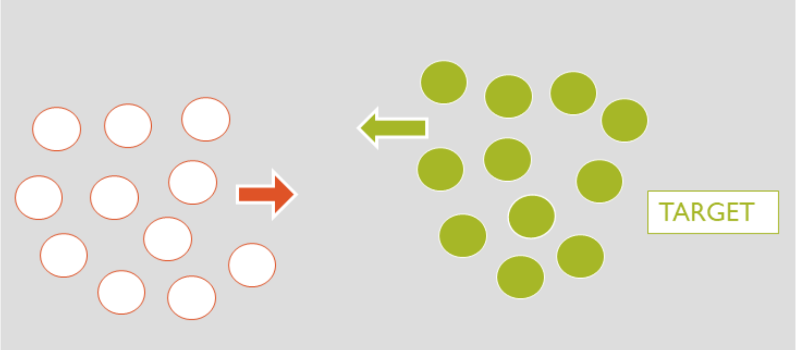


Figure 4.1: Initial design of the "Inducing UAVs" approach

Alternatively, the "UAV Swarm vs. UAV Swarm" approach proposed a counter-attack by a defending UAV swarm against an incoming adversary swarm. However, due to its confrontational nature, this approach was later discarded in favor of non-confrontational methods. KPIs primarily revolved around the required number of defending UAVs. Figure 1 and Figure 2 illustrate the initial approaches. The project later shifted toward a Machine Learning-driven approach. Since I was allowed to choose the ML methodologies freely, I re-framed the problem to resemble a game-based approach, connecting it to my studies and area of interest.

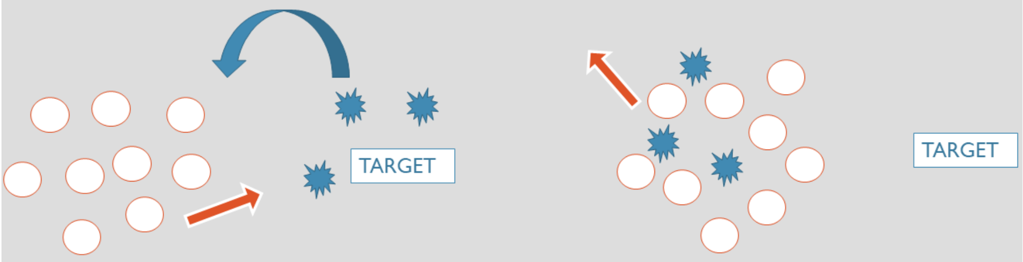


Figure 4.2: Initial design of the "UAV swarm vs. UAV swarm" approach

4.1.2 Preliminary Research and Methodologies

The research process was initiated with an exploration of UAV technologies and current research directions. This process laid the groundwork for approaching the problem with the necessary background knowledge. Subsequently, I focused closely on state-of-the-art UAV countermeasures research, with a specific emphasis on Machine Learning methodologies. The results of this work are presented in the literature review of this thesis (Chapter 2).

In the area of ML-driven UAV countermeasures, two distinct approaches showed promise of fitting with our research: Graph Neural Networks and Multi-Agent Reinforcement Learning, particularly the Multi-Agent Deep Deterministic Policy Gradients framework. Both GNNs and MADDPG have showcased promising potential in the domain of UAV countermeasures. However, neither of these algorithms had yet been applied to the exact problem at hand, which highlights the potential novelty of our approach.

While GNNs would make for a promising area to explore, further investigation of MARL and MADDPG led me to settle on this methodology. Central to this decision was the availability of an accessible simulation environment, namely the Multi-Agent Particle Environment, as well as the inherent gamification of the MARL methods. Furthermore, an attempt was made to utilize Agent-Based Modeling (ABM)-related simulation platforms such as Mesa and NetLogo. However, due to limitations in their ML application programming interfaces (APIs) at the time, they were less accessible for implementing the defense strategies compared to the more adaptable MPE environment.

4.1.3 Formulating the Problem as a Game

In designing the countermeasure, the dynamics of UAV interaction mimicked strategic interactions often seen in gaming scenarios. This transition from a conventional problem to a game-like framework was motivated by the need to model the competitive and cooperative nature of UAV swarm behavior.

The game representation involves two main roles: attackers and defenders. Attackers aim to reach a designated target, while defenders work to prevent this objective. These roles, much like players in a game, follow specific strategies, effectively making decisions based on their observations and accumulated knowledge. This game setup allows us to utilize the principles of MARL to develop intelligent countermeasures.

The core challenge here is to establish the rules, strategies, and decision-making processes that simulate real-world UAV interactions within the confines of the game. Defenders must dynamically adapt their actions to outmaneuver the attackers, who are attempting to reach the target. This interplay captures the essence of the countermeasure problem, transforming it into a strategic game of decisions and outcomes. The game framework also opens the door to defining reward structures that encourage desirable behaviors. For defenders, rewards are associated with successful interception or distraction of attackers, while attackers are rewarded for progressing toward the target. This reward system guides the learning process of our agents, enabling them to develop effective strategies through reinforcement learning.

By framing the UAV countermeasure challenge as a game, we establish a structured environment that captures the complexities of real-world interactions. This approach lays the foundation for applying advanced techniques like MARL, setting the stage for developing effective defense strategies.

4.2 Simulation Environment and Agent Behavior

In this section, we describe the details behind the simulation environment, define action and observation spaces, and elaborate on the scenario setup. We also describe the swarming and the attacker-specific agent behaviors in detail. The equations and formulas presented in this chapter were initially introduced in our paper [69].

4.2.1 Simulation Environment

The simulation environment serves as a virtual space where the UAV swarm confrontation unfolds, providing a controlled setting for testing and evaluating the effectiveness of the proposed defense strategies. In this section, we detail the design and components of the simulation environment, which enables the dynamic interactions between defending and attacking agents.

Our simulation environment leverages the Multi-Agent Particle Environment, a customizable platform that aligns with our research objectives. Our environment models N agents and L landmarks inhabiting a 2D area with continuous space and discrete time. Agents can perform physical actions (e.g., moving in a certain direction) in the environment and communication actions that get broadcast to other agents. This two-dimensional environment allows for the representation of agents and targets as particles, ensuring the scalability required to accommodate various swarm sizes and scenario complexities.

Within the MPE framework, both defending and attacking agents are characterized as intelligent entities equipped with perceptive capabilities and strategic decision-making. These agents possess attributes such as position, velocity, and orientation, enabling them to navigate the environment and interact with their surroundings. Their discrete action space facilitates the execution of a range of actions, while their individual observation spaces encompass critical information about team members, adversaries, and the target.

Dynamic interactions between agents unfold within the simulation environment, closely resembling the cooperative-competitive behaviors observed in real-world UAV swarm confrontations. The foundation of these dynamics lies in multi-agent reinforcement learning principles, where agents perform actions within the environment and receive rewards based on the outcomes. This dynamic interplay facilitates the emergence of strategic behaviors and the exploration of defense strategies.

To facilitate informed decision-making, agents in the simulation environment observe relevant information from their surroundings. Observations include details about the positions and movements of teammates, adversaries, and targets. Communication between agents is limited to mimic realistic scenarios, where defenders may possess additional information and can attempt to mislead or deceive attacking agents through strategic messaging.

The simulation environment accommodates various scenarios that challenge the agents with different contexts and complexities. Scenarios encompass distinct configurations of defender and attacker groups and varying target locations. This variability ensures a comprehensive assessment of defense strategies across a range of challenging conditions.

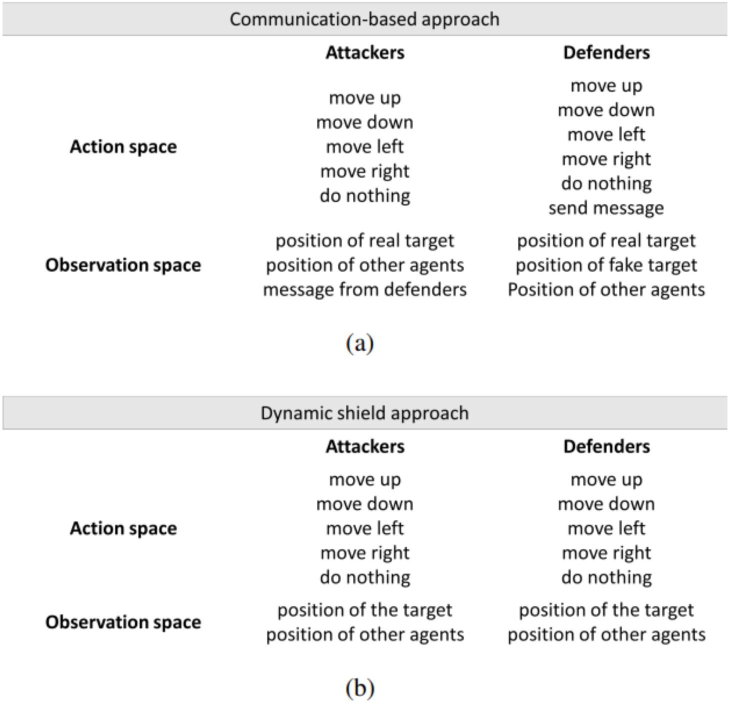
To enhance the analysis and understanding of agent behaviors, a real-time visualization component provides a graphical representation of the ongoing UAV swarm confrontation. This visualization allows researchers to observe and interpret the strategic movements, interactions, and decision-making processes of the agents as they unfold within the virtual environment.

In the following sections, we delve into the behaviors of defending and attacking agents, the utilization of the MADDPG algorithm for policy learning, and the integration of these components to forge effective and adaptive UAV swarm defense strategies.

4.2.2 Action and Observation Space

In the construction of the actor-critic environment, agents need well-defined observation and action spaces. These spaces dictate how agents choose actions based on what they perceive from their surroundings. In the case of the Dynamic Shield defense, both attackers and defenders share the same action and observation spaces. However, in the Communication-based defense, these spaces differ due to an added capability for defenders to send messages to attackers. In the Dynamic Shield approach, which primarily relies on movement and physical interactions, the action space offers five movement choices in different directions. In the Communication-based defense, defenders can go beyond basic movement actions by also sending messages.

The observation space involves key aspects such as the target's position, the locations of other agents, and the distance between an agent and the target. In the Communication-based approach, attackers and defenders each have additional observations: Attackers can receive messages from defenders, while defenders can also observe the position of a fake target. Figure 4.3 presents the (a) action and (b) observation spaces for attackers and defenders in both the Communication-based and Dynamic Shield approaches.



Communication-based approach

	Attackers	Defenders
Action space	move up move down move left move right do nothing	move up move down move left move right do nothing send message
Observation space	position of real target position of other agents message from defenders	position of real target position of fake target Position of other agents

(a)

Dynamic shield approach

	Attackers	Defenders
Action space	move up move down move left move right do nothing	move up move down move left move right do nothing
Observation space	position of the target position of other agents	position of the target position of other agents

(b)

Figure 4.3: Action and observation space of the (a) Communication-based approach and the (b) Dynamic Shield approach. Initially introduced in [69].

4.2.3 Scenario Setup and Assumptions

As stated earlier, the primary objective of the defenders is to maintain distance between the attackers and the target. To accomplish this goal, we have implemented two distinct RL-based defense strategies. The first strategy assumes that defenders have access to the communication channel utilized by the attackers, while the second strategy focuses on physically safeguarding the target. Notably, the attackers also exhibit intelligent behavior in their pursuit of the target, leveraging RL techniques. To formulate these two scenarios, we take into account the following components:

1. Spatial Configuration: The environment is modeled within a continuous 2D space that encompasses a stationary target M, positioned at $\hat{m} = [x_m, y_m]$, and an auxiliary target F at $\hat{f} = [x_f, y_f]$. Notably, the auxiliary target F is exclusively utilized in the communication-based defense approach.

2. UAV Agent Dynamics: The environment involves a collective of attacking UAVs, denoted as $a_k \in A$ with $k = 1, \ldots, K$, and a group of defending UAVs, designated as $d_l \in D$ where $l = 1, \ldots, L$. The unified set of all UAVs is denoted as $U = A \cup D$.

3. Discrete Framework: The environment operates in a discrete manner, with both time and spatial coordinates discretized.

During each discrete time step t_s, every agent selects a singular action from the available action space $\mathcal{S}$, which is explained in the subsequent section. We assume that each UAV is outfitted with a GPS module, facilitating knowledge of its own spatial coordinates.

In terms of notation, spatial coordinates are consistently annotated with a circumflex (e.g., $\hat{a}_k$ signifies the spatial coordinates $[x_k, y_k]$ of attacker a_k). One additional assumption made during our initial endeavors to counter UAV swarm attacks is that defenders possess awareness of the attackers' positions. This assumption could potentially be realized through the incorporation of supplementary hardware or sensors to estimate attackers' positions from the defenders' perspective. Nevertheless, we intend to relax this assumption in our forthcoming work and reconfigure the reward function that relies on this information.

4.2.4 The Swarm Behavior

In this section, we introduce a set of general rewards denoted as r^u_n, designed to establish fundamental behaviors for all UAV agents. The primary objective is to ensure smooth and collision-free interactions among UAVs. To achieve this, the first reward, denoted as r^u_1, encapsulates a typical swarm behavior approach, encouraging repulsion between agents. Mathematically, this is formulated as follows:

$$r^u_1(\hat{u}_i) = \begin{cases} -\varrho, & \text{if } \|(\hat{u}_i - \hat{u}_j)\|_2 < \epsilon \mid u \in U \ \wedge \ i \neq j \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

When the distance between two UAV positions, $\hat{u}_i$ and $\hat{u}_j$, falls below a predefined threshold ϵ, a negative reward is assigned to the agent. Importantly, this reward function applies to every agent $u \in U$. Since UAVs commonly utilize various sensors (e.g., cameras, ultrasound) to measure object distances, repulsion can be effectively achieved between UAVs without necessitating communication.

The second overarching reward, r^u_2, is crafted to motivate agents to remain within a bounded region surrounding the target. This reward takes the form:

$$r^u_2(\hat{u}_i, \hat{m}) = \begin{cases} +\varrho, & \text{if } \|(\hat{u}_i - \hat{m})\|_2 < 2\gamma \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

When an agent u_i is positioned within a radius of 2γ from the target, it receives a positive reward; otherwise, no reward is granted. These two rewards form the fundamental behaviors guiding all agents. They foster cooperative behaviors among UAVs, enabling them to operate as a cohesive swarm while avoiding collisions.

4.2.5 Attacker-Specific Behavior

The reward functions tailored to attackers' behavior are aimed at maintaining a coherent swarm and optimizing their proximity to the target. The initial attacker-specific reward r_1^a is designed to encourage a swarm-like behavior among agents, which is expressed as:

$$r_1^a(\hat{a}_i) = \begin{cases} +\varrho, & \text{if } \exists a_i : \epsilon \leq \|(\hat{a}_i - \hat{a}_j)\|_2 < 2\epsilon \mid a \in A \, \wedge \, i \neq j \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

When the distance between an attacker's position $\hat{a}_i$ and any other attacker's position $\hat{a}_j$ falls within the range of ϵ to 2ϵ, the agent a_i receives a positive reward. The subsequent reward, denoted as r_2^a, hinges on the attackers' proximity to the target. This reward introduces a penalty if attackers move away from the target, and it is mathematically defined as:

$$r_2^a(\hat{a}_i, \hat{m}) = -\min\|(\hat{a}_i - \hat{m})\|_2, \tag{4.4}$$

Here, $\hat{a}_i$ represents the positions of the attackers, and $\hat{m}$ signifies the position of the target. The reward function assigns a negative value based on the minimum distance between the attackers and the target. Consequently, the further the attackers are from the target, the more substantial the negative rewards become.

Finally, a reward to acknowledge victory for the agents is defined as r_3^a. This reward is triggered when any of the attackers successfully reaches the target:

$$r_3^a(\hat{a}_i, \hat{m}) = \begin{cases} +10\varrho, & \text{if } \exists a_i \in A : \|(\hat{a}_i - \hat{m})\|_2 \leq \gamma \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

Here, $\hat{a}_i$ denotes the positions of the attackers, $\hat{m}$ signifies the target's position, and γ is the threshold distance defining when the target is considered reached. These attacker-specific reward functions are designed to shape the agents' behavior in a way that aligns with the desired strategies.

4.3 Game-like Elements

In the context of RL and MARL, the problem of countering adversarial UAV swarms takes on a distinct gamified nature. This section outlines the gamification aspects integrated into our research, emphasizing how RL techniques have been tailored to create engaging and competitive scenarios, mirroring real-world challenges in a game-like environment.

Gamification, the application of game design elements in non-game contexts, offers a powerful approach to enhance learning and problem-solving within complex environments. In the domain of UAV swarm defense, leveraging gamification principles can provide an intuitive and engaging way to develop effective countermeasures. To realize the gamified aspects of the research, specific scenarios were crafted to simulate adversarial encounters between UAV swarms. These scenarios encompass both cooperative and competitive interactions, enabling agents to engage in intricate strategic decision-making while vying to achieve their respective objectives. The gameplay dynamics involve defenders aiming to mislead, distract, or delay attacking UAV swarms from reaching vital targets, while the attackers strive to overcome these obstacles and achieve their mission objectives. Essentially, two teams are formed each with their own winning and losing criteria. Within the implemented scenarios, various game elements and incentives are introduced to drive agent behavior and decision-making. These elements include rewards that encourage effective defensive tactics and penalties for undesirable actions. Reinforcement learning rewards resemble player scores in this context, where the highest agent score correlates to the highest-achieving player. By formulating the UAV swarm confrontation as a game, the reinforcement learning agents adopt strategies that optimize their chances of success while factoring in potential risks and uncertainties.

The gamified nature of the scenarios facilitates efficient learning for the agents. Through repeated interactions and gameplay, agents acquire a deep understanding of the environmental dynamics and the consequences of their actions. Over time, this learning process results in the development of sophisticated defensive and offensive strategies, enhancing the overall efficacy of the countermeasures. By infusing game design elements into the research, the study leverages the intrinsic motivation and engagement inherent to gamified environments. This approach not only advances our understanding of UAV swarm defense strategies but also offers insights into the applicability of gamification principles in addressing complex real-world challenges.

In the subsequent sections, we delve into the technical aspects of the implemented scenarios, exploring the details of the simulation environment, agent behaviors, and the algorithmic design that underpins the game-like UAV swarm confrontation.

4.4 Game-based Defense Strategies Implementation

In this section, we look into the practical implementation details of the defense strategies designed to counter adversarial UAV swarms. We discuss two distinct approaches, namely the Communication-based Approach and the Dynamic Shield Approach. These approaches embody the core mechanisms through which defenders aim to protect a designated target against attacking UAV swarms.

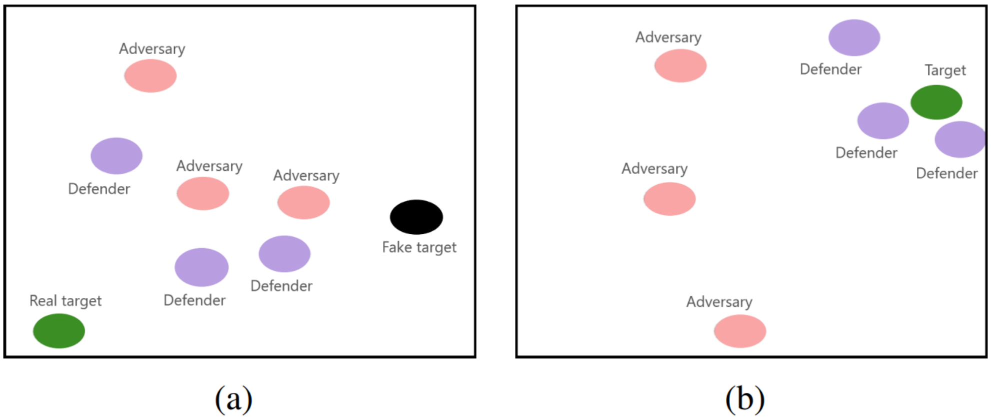


Figure 4.4: Illustrations of the simulation environment for the (a) Communication-based defense and the (b) Dynamic Shield defense. Initially introduced in [69].

4.4.1 Training and Optimization Processes

In the context of training Multi-Agent Deep Deterministic Policy Gradient agents, two key components play a crucial role: Experience Replay and Target Networks. Experience Replay is a technique used in reinforcement learning algorithms, including MADDPG, to improve the stability and efficiency of training. It involves storing past experiences (observations, actions, rewards, and resulting states) in a replay buffer. During training, the agent samples batches of experiences from this buffer to update its neural network parameters. In our case, the replay buffer holds 10^6 past experiences. This approach ensures a balance between exploring new possibilities and exploiting learned knowledge, resulting in improved overall performance.

In reinforcement learning, the target value used in the Bellman equation for updating the Q-network (critic network) is often estimated with another neural network, known as the target network. This helps stabilize the learning process by decoupling the target value calculation from the network being updated. In MADDPG, each agent maintains two sets of neural networks for both the actor and critic: local networks and target networks. Target networks are used to estimate the target values for updating the local networks. The target networks are updated more slowly by periodically copying the weights from the local networks. This soft update process helps maintain a stable and consistent target value estimation. By using target networks, the learning process becomes more robust and less susceptible to oscillations. In our case, hyperparameter tau represents how quickly the target networks get updated. We have set tau to 0.01 to ensure a gradual and stable updating process.

In general, the hyperparameters of our algorithm are aligned with the ones used in the original MADDPG approach, which can be found in more detail in [47]. We have employed the Adam optimizer, which is a widely used optimization algorithm in machine learning and deep learning. It stands for "Adaptive Moment Estimation" and is designed to efficiently adjust the learning rate for each parameter during the training process. In our case, the learning rate for this optimizer is set to 0.01. Another critical parameter is gamma. Gamma, often referred to as the discount factor, determines how much an agent values future rewards when making decisions. In our case, it is set to 0.95.

4.4.2 Communication-based Approach

The Communication-based approach focuses on the injection of false information into the attacking swarm to mislead it away from the real target. Defenders disguise themselves as attackers and leverage a communication channel to broadcast messages. The following components contribute to the implementation of this approach:

- Fake Target Introduction: We create a virtual scenario in which both a real target M and a fake target F coexist. Defenders are aware of the true identity of these targets, while attackers are not. This approach operates under the assumption that defenders have the capability to access and transmit messages via a communication channel.

- Observation Spaces: We design distinct observation spaces for both attackers and defenders. Attackers' observations encompass relative distances to both targets and other agents, along with communication messages. Defenders' observation space includes the same elements as attackers' observations, supplemented by knowledge of the real and fake targets.

- Reward Functions: Defenders' actions are guided by reward functions based on proximity to the fake target. Negative rewards are introduced to encourage defenders to remain closer to the fake target than the attackers. Additionally, winning rewards are defined for scenarios where defenders successfully mislead the attackers toward the fake target or maintain a strategic presence around the real target.

The first defender-specific reward function, r_1^d, is predicated on the distance between defenders and the fake target and can be mathematically defined as:

$$r_1^d(\hat{d}_i, \hat{f}) = -\min\|(\hat{d}_i - \hat{f})\|_2, \tag{4.6}$$

Here, $\hat{d}_i$ symbolizes the defender's position, and $\hat{f}$ denotes the position of the fake target. This equation denotes a negative reward that becomes smaller as defenders get closer to the fake target, stimulating behaviors that mislead attackers.

Likewise, the second defender-specific reward, r_2^d, refers to the distance between attackers and the fake target, facilitating the defenders' aim to draw attackers towards the fake goal:

$$r_2^d(\hat{a}_i, \hat{f}) = -\min\|(\hat{a}_i - \hat{f})\|_2, \tag{4.7}$$

In this equation, $\hat{a}_i$ signifies the attacker's position, and $\hat{f}$ represents the fake target's position.

Two winning rewards are designed for defenders. The first, denoted as r_3^d, is obtained when defenders successfully lead attackers to the fake target:

$$r_3^d(\hat{a}_i, \hat{f}) = \begin{cases} +10\varrho, & \text{if } \exists a_i \in A : \|(\hat{a}_i - \hat{f})\|_2 \leq \gamma \\ 0, & \text{otherwise.} \end{cases} \tag{4.8}$$

Here, $\hat{a}_i$ signifies the attacker's position, $\hat{f}$ is the fake target's position, and γ defines the proximity threshold to qualify as reaching the target.

The second winning reward, r_4^d, is applicable if the attackers fail to reach the real target within the defined conditions. In this case, defenders are awarded when they are collectively within the boundaries of the $2D$ area:

$$r_4^d(\hat{d}_i, t_s) = \begin{cases} +10\varrho, & \text{if } t_s > T \land \forall \hat{d}_i : \text{inside boundary} \\ 0, & \text{otherwise.} \end{cases} \tag{4.9}$$

Here, $\hat{d}_i$ represents the defender's positions, t_s denotes the simulation step, and T signifies the total number of simulation steps. This reward acknowledges defenders' success if they managed to remain within the bounded area while attackers were unable to reach the real target. The threshold distance is set as 3γ, ensuring defenders remain in close proximity around the target without reaching it.

4.4.3 Dynamic Shield Approach

The Dynamic Shield approach aims to physically protect the designated target by establishing a dynamic shield of defenders around it. The following components are integral to the implementation of this approach:

- Dynamic Shield Creation: Defenders establish a protective barrier around the target to deter attackers from reaching it. The shield operates on the principle of repulsion, with attackers avoiding both defenders and the target.

- Shared Observation Space: Both attackers and defenders share an observation space that includes relative distances to the target and other agents. This shared space guides the behaviors of both types of agents.

- Reward Functions: Defenders' actions are motivated through reward functions that penalize them for being distant from the target. Negative rewards are based on the minimum distance between defenders and the target, encouraging them to remain close to the protected area. Winning rewards are established for scenarios in which defenders successfully prevent attackers from reaching the target.

To safeguard the target, the first defender reward r_1^d depends on the defender's proximity to the target. Accordingly, a reward function is designed to penalize defenders for positioning themselves too distantly from the target:

$$r_1^d(\hat{d}_i, \hat{m}) = -\min\|(\hat{d}_i - \hat{m})\|_2, \tag{4.10}$$

Here, $\hat{d}_i$ represents the positions of defenders, and $\hat{m}$ signifies the position of the target. This formulation produces a negative reward that increases as the defenders move farther from the target.

Similarly, the second reward function r_2^d evaluates the distance between attackers and the target:

$$r_2^d(\hat{a}_i, \hat{m}) = \sum_{i=1}^{I}\|(\hat{a}_i - \hat{m})\|_2, \tag{4.11}$$

Here, $\hat{a}_i$ denotes the attacker positions, and $\hat{m}$ represents the target position. This reward computes the sum of the Euclidean distances between all attackers and the target. Consequently, the reward magnitude increases with greater distances between attackers and the target.

A winning reward, r_4^d, is defined to signify success in a game. This reward is achieved if, by the end of the game, attackers have not managed to reach the target and defenders have effectively stayed in proximity to the target, preventing attackers from passing through and reaching it:

$$r_4^d(\hat{d}_i, \hat{m}, t_s) = \begin{cases} +10\varrho, & \text{if } t_s > T \land \|(\hat{d}_i - \hat{m})\|_2 \leq 3\gamma \\ 0, & \text{otherwise.} \end{cases} \tag{4.12}$$

Here, $\hat{d}_i$ signifies the defender positions, $\hat{m}$ represents the target position, t_s indicates the simulation step and T is the total number of simulation steps. The threshold distance is set as 3γ, ensuring defenders remain in close proximity to the target without physically converging on it. The winning reward is awarded when defenders effectively maintain close proximity around the target.

4.5 Problems and Solutions

In this section, we confront the various challenges encountered throughout the research process. These include the complexity of the problem domain and the agent behavior design, the struggle to strike a balance between abstraction and real-world relevance, environment setup issues, the duration of training processes, and the initial difficulties faced during the design of a Communication-based approach. We outline the strategies and solutions devised to address these issues and provide insights into their impact on the research trajectory.

4.5.1 Problem Complexity

Addressing the challenges of UAV defense involves navigating complexities at different stages. This subsection explores the depth of research, complexities in designing intelligent agent behaviors, and the decision-making process for modeling real-world dynamics.

4.5.1.1 Extensive Research and Problem Context

Developing a robust UAV defense solution necessitated thorough research in Swarm Intelligence, UAV technology, Machine Learning, and Multi-Agent Reinforcement Learning. This research provided a foundational understanding of UAV swarming, adversarial dynamics, and decision-making, informing the solution's contextual framework.

4.5.1.2 Intricacies of Agent Behavior Design

Designing intelligent agent behaviors within a Multi-Agent System (MAS) presented challenges due to the non-transparent nature of ML methods. Iterative design cycles were essential, guided by empirical exploration and reward-shaping refinement. Balancing rewards to guide desired behavior while minimizing unintended consequences was a crucial challenge in generating efficient countermeasures.

4.5.1.3 Abstraction and Real-World Complexity

Choosing an abstraction level for modeling complexities required a strategic trade-off between realism and computational efficiency. The high-level abstraction of the Multi-Agent Particle Environment was chosen to represent UAV interactions. This allowed for the effective exploration of countermeasure strategies while streamlining computational demands. In summary, the challenges of UAV defense span research depth, agent behavior complexities, and abstraction-level decisions.

4.5.2 Environment Setup Challenges

This subsection delves into the challenges encountered during the initial phases of the implementation process. It highlights the complexities in implementing the tools and libraries for the MADDPG framework.

The implementation of the MADDPG framework necessitated the integration of a diverse array of tools and libraries. Challenges are presented due to version conflicts between reinforcement learning and deep learning components. Resolving these conflicts required systematic testing and experimentation within virtual environments, ensuring compatibility among different tool versions. Employing these tools effectively within the Windows operating system presented further challenges. It became evident that certain tools were not optimally configured for the Windows Operating System, which could be attributed to a range of encountered issues.

Addressing version conflicts and platform compatibility issues was central to establishing a functional and efficient environment for subsequent development.

4.5.3 Training Length

This subsection addresses the challenges stemming from training duration. The study's aim to accommodate diverse swarm sizes for comprehensive analysis, as advised by peer reviews, required training multiple models with escalating swarm sizes. This led to longer training times due to increased complexity, particularly notable in the Communication-based approach. The extended training duration posed time management challenges, potentially jeopardizing submission deadlines for a venue. Balancing complexity, training time, and deadlines were critical factors in navigating the research and submission process.

4.5.4 Initial Communication-based Approach Design Difficulties

The design of the first implemented approach was based on providing adversarial UAVs with false information through the defender's communication channel. This design was partially based on prior work in this area, where the swarm algorithm Grey Wolf Optimizer was used to mislead the attacking UAVs from the target. Similarly, in the first version of the Communication-based approach, we attempted to have the defending UAVs mislead the adversaries by infiltrating them and leading them towards a fake target. The complications arose promptly after the first training attempts. During training, we took note of how the agents' intelligent behavior was evolving by observing the number of wins for defending agents vs. attacking agents. The initial observations pointed out a pattern of high-to-middle winning rate of the defenders in the beginning, followed by a rapid decline and almost 100% attackers' win rate. The cause for this sudden change appeared to be the attackers adjusting and learning to disregard the signals from the defending agents after some iterations of training episodes. Following this realization, we had to adjust the rewards for defending agents as well as reimagine the overarching goal of the defenders to delay the attackers instead of completely misleading them onto a fake target. The reframing of the rewards and the defender goals has shifted the approach into a more effective strategy, that after some more design iterations produced valuable results which will be further covered in the following chapter.

4.6 Evaluation Metrics Implementation

To evaluate the performance of the developed approaches, an analysis of scenario training and resulting data was conducted. Python libraries were leveraged to extract valuable insights and visualize the outcomes of the training process. Matplotlib, Scipy, and Numpy were among the key libraries utilized for data manipulation, analysis, and visualization. By parsing the training data, we were able to monitor critical metrics such as the number of defender and adversary wins per scenario and per episode. This information allowed us to gain a deep understanding of the effectiveness of the implemented strategies in different scenarios and to identify any trends or patterns emerging during the training process.

Another element of our evaluation process was the integration of the OpenAI Gym platform and the Multi-Agent Particle Environment. OpenAI Gym provided a standardized framework for benchmarking and evaluating reinforcement learning algorithms, enabling us to assess the performance of our models against well-defined metrics. The MPE environment served as the testing ground for our trained models. It allowed us to not only observe the performance of the agents as they executed learned strategies but also provided valuable insights into the learning process itself. The ability to visualize the agents' behaviors and interactions in a controlled environment was crucial in validating the effectiveness of our countermeasure approaches.

As we proceed, we will delve into the specific analyses conducted on the training data and present visualizations that illustrate the performance trends and improvements achieved through the training iterations.

Chapter 5

Evaluation of Results

This chapter delves into the evaluation of the proposed game-based defense strategies. We start by outlining the experimental setup. Subsequently, we evaluate the performance of the defense strategies, including an examination of reward stability and winning rates across various swarm sizes. We further look at a comparative analysis of winning rates. To gain a deeper understanding of agent behavior, we present observations, specifically focusing on the Communication-based and Dynamic Shield approaches. This chapter provides valuable insights into the effectiveness and behavior of our defense strategies.

5.1 Experimental Setup

To evaluate the effectiveness of the designed game-based approaches for countering attacking UAV swarms, an experimental setup was established. This subsection outlines the hardware and software environment, simulation specifics, and training parameters employed during the evaluation process.

The experiments were conducted on a laptop equipped with an Intel Core i7 processor and a NVIDIA GeForce 940MX graphics card. Python 3.8 was used as the programming language, and several libraries were utilized to facilitate experimentation, including NumPy for numerical computations, PyTorch for neural network implementations, and Matplotlib for data visualization.

To replicate the real-world scenario, we utilized the Multi-Agent Particle Environment from the OpenAI Gym toolkit. The MPE offers a platform for simulating multi-agent environments, enabling us to train and evaluate our game-based approaches effectively. The environment was configured to represent the attacking and defending UAV swarms, and interactions were modeled within a two-dimensional space.

To ensure the convergence of the trained models, a considerable number of training episodes were conducted. Each episode represented a single iteration of the agents' interactions within the MPE environment. The training was performed using the MADDPG algorithm, which facilitates multi-agent reinforcement learning.

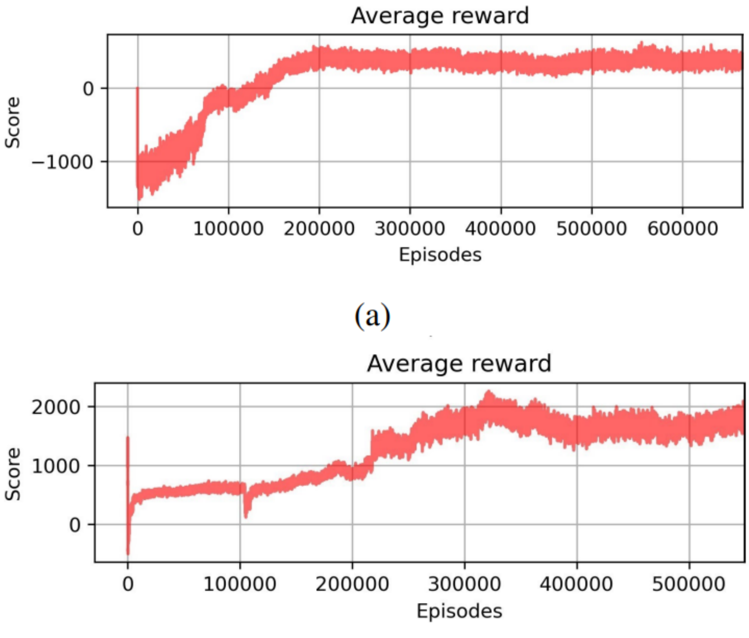
Hyperparameters such as learning rates, discount factors, and batch sizes were utilized to optimize training stability and convergence.

By establishing this experimental setup, we aimed to provide a controlled environment for evaluating the performance of our proposed game-based approaches. The subsequent sections delve into the outcomes of these experiments, shedding light on the stability of rewards, winning rates based on swarm sizes, and comparative analyses of winning rates between the approaches.

5.2 Performance Evaluation

In this section, we evaluate the performance of our game-based defense strategies. To test their effectiveness, we examine the stability of rewards as a measure of learning consistency. Furthermore, we look at the winning rate evaluations, analyzing scenarios with varying swarm symmetries. Through a comparative analysis of winning rates, we illustrate the strategies' comparative advantages.

5.2.1 Stability of Rewards



(a)

(b)

Figure 5.1: Average reward of all agents in the 3 defenders vs. 3 attackers scenario for the (a) Communication-based defense and the (b) Dynamic Shield defense. Initially introduced in [69].

The stability of rewards is an important measure indicating the convergence and stability of the applied defense strategies. Figure 5.1 depicts the average reward trend during the learning phase for the 3 defenders vs. 3 attackers scenario for both the Communication-based defense and the Dynamic Shield defense. The y-axis represents the average accumulated rewards, which are elaborated on in previous sections. The rewards demonstrate consistent stability after approximately 400,000 episodes for both defense strategies. As a result of this observed stability, the training of the models was extended to 600,000 episodes to ensure robust and reliable convergence.

The following figures showcase the stability of rewards for both defense approaches, as well as corresponding plots that display how the agents' winning rate reflects the reward stability. The training of both scenarios took around 500,000 episodes, which can be considered sufficient to obtain a stable model, especially for smaller swarms.

In Figure 5.2, we analyze the stability of rewards during the training of the Dynamic Shield approach with 2 defenders and 3 attackers. This figure provides valuable insights into the learning process of the agents. Here we can observe that the defenders start winning against the attackers after approximately 50,000 games, and how their cumulative wins equalize after around 470,000 games.

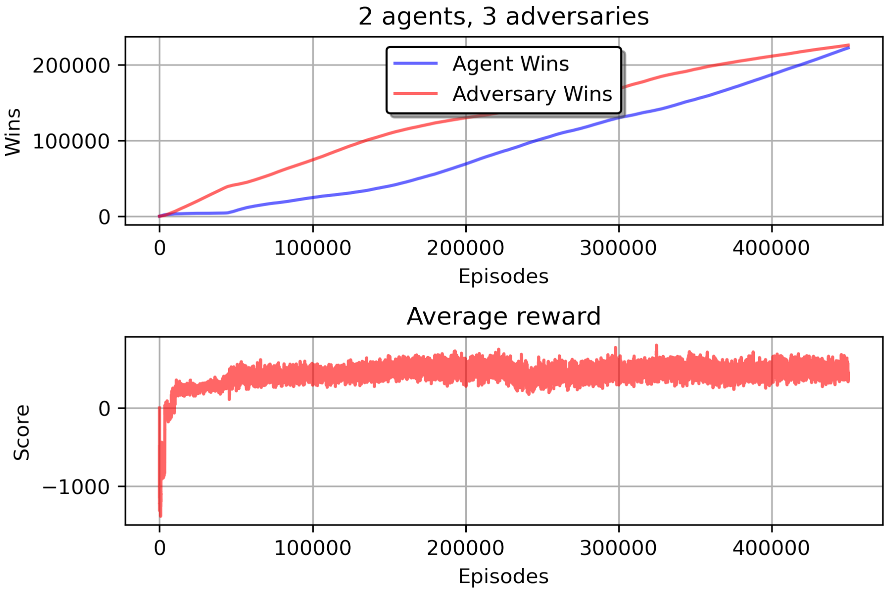


Figure 5.2: Dynamic Shield defense training process

In Figure 5.3, we observe the stability of rewards during the training of the Communication-based approach with 2 defenders and 2 attackers. In contrast to the Dynamic Shield plot, the Communication-based approach demonstrates consistent defender superiority throughout the training. However, a surge in defender performance happens after approximately 300,000 episodes.

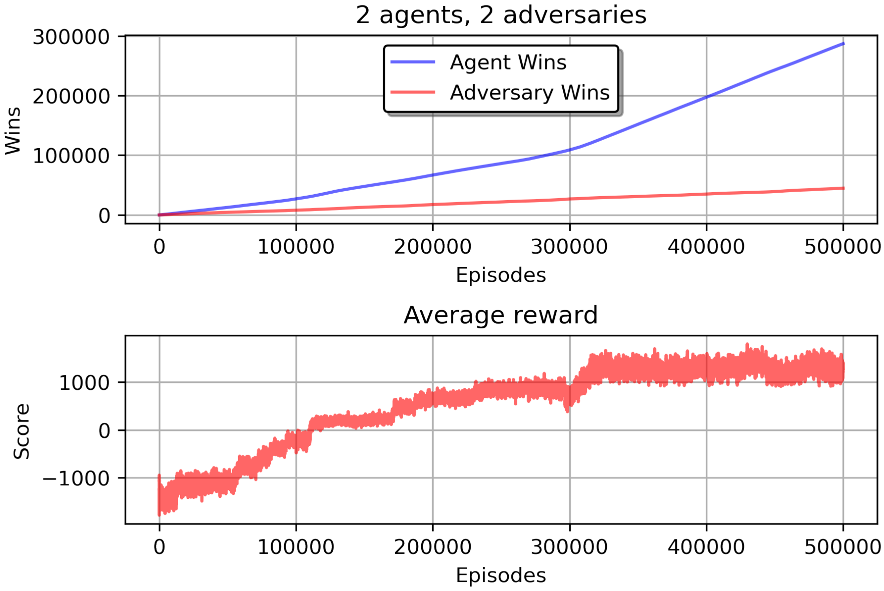


Figure 5.3: Communication-based defense training process

It is worth noting that achieving a stable model is enhanced when the winning rates of both defending and attacking agents exhibit fluctuations during training. This pattern indicates that both teams are continuously evolving and striving to outperform each other, ultimately contributing to the development of more effective strategies.

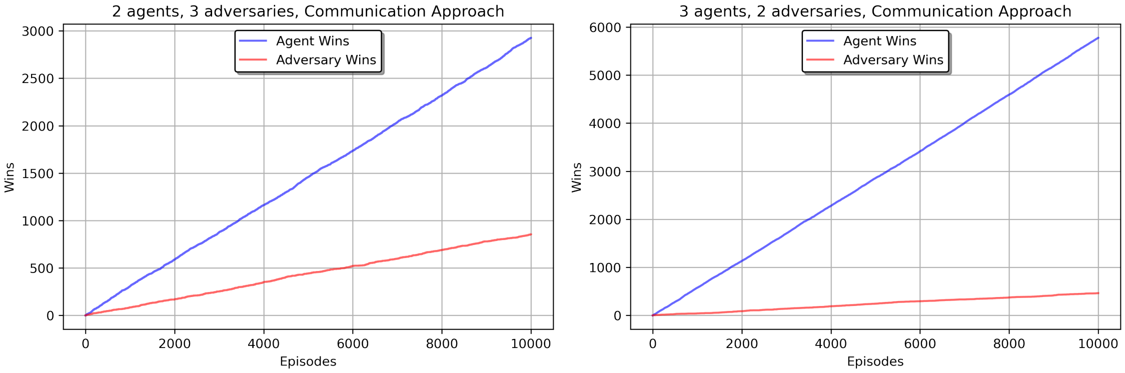
5.2.2 Winning Rate Evaluation based on Swarm Sizes

In this section, we present a thorough evaluation of the performance of the Communication-based defense approach across different scenarios involving varying numbers of defenders and attackers. The primary objective is to assess the ability of the defense approaches to counter adversarial UAV swarms under different swarm compositions and complexities.

For each scenario, a trained model corresponding to the respective defense approach was utilized. The simulation was conducted over 500,000 episodes to ensure an in-depth assessment of the defense mechanisms' performance. The simulation environment is highly abstracted from the real-world conditions, while still incorporating the dynamics of UAV swarm interactions and the unique challenges posed by the specific scenario.

5.2.2.1 Asymmetrical Smaller Swarm Scenarios

From the plots, we can see that the y-axis represents the probability of both defender and attacker winning, while the x-axis denotes the number of episodes/games. In Figure 5.4, we observe the winning probability plot for the Communication-based defense approach in two asymmetrical smaller swarm scenarios. In scenario (a), with 2 defenders against 3 attackers, the defending team achieves a winning rate of approximately 3000, while the adversaries have a winning rate of around 1000. In scenario (b), with 3 defenders against 2 attackers, the defending team exhibits a winning rate of about 6000, while the adversaries' winning rate is approximately 500, representing the lowest number of adversary wins among all scenarios.



(a) 2 defenders vs 3 attackers scenario

(b) 3 defenders vs 2 attackers scenario

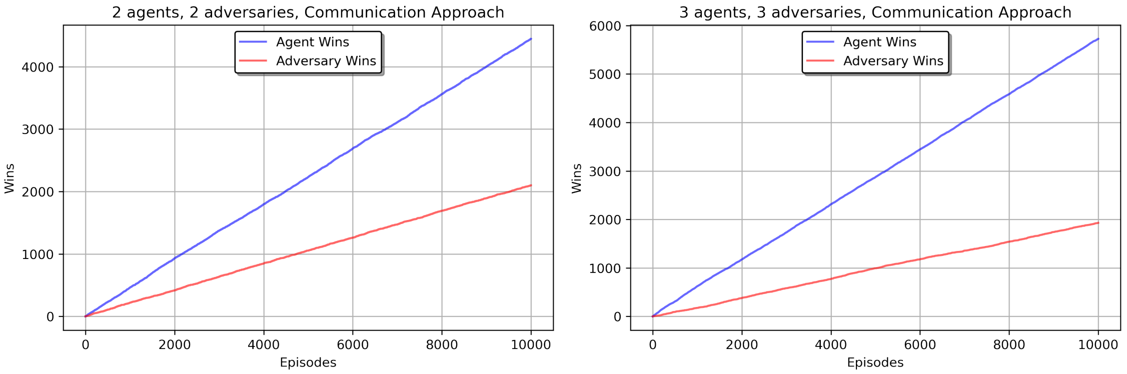
Figure 5.4: Figure presents the winning probability of the Communication-based defense approach for the asymmetrical smaller swarm scenarios

5.2.2.2 Symmetrical Swarm Scenarios

In Figure 5.5, we examine the winning probability plot for the Communication-based defense approach in two symmetrical swarm scenarios. In scenario (a), with 2 defenders against 2 attackers, the defending team attains a winning rate of around 4000, while the adversaries have a winning rate of approximately 2000. In scenario (b), with 3 defenders against 3 attackers, the defending team demonstrates a winning rate of about 6000, while the adversaries maintain a winning rate of around 2000.

5.2.2.3 Asymmetrical Larger Swarm Scenarios

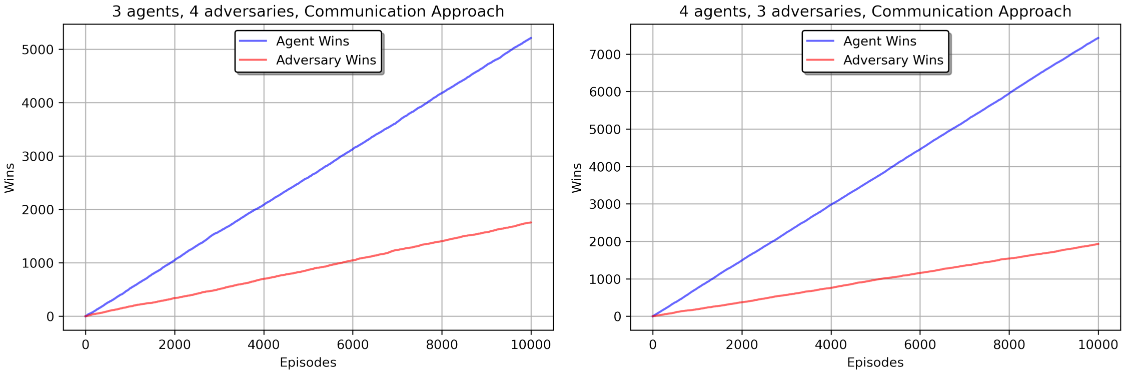
Figure 5.6 illustrates the winning probability plot for the Communication-based defense approach in two asymmetrical larger swarm scenarios. In scenario (a), with 3 defenders against 4 attackers, the defending team achieves a winning rate of approximately 5000, whereas the adversaries have a winning rate of around 2000. In scenario (b), with 4 defenders against 3 attackers, the defending team exhibits the highest winning rate among all scenarios, reaching about 7000, while the adversaries maintain a winning rate of around 2000. These results provide valuable insights into the performance of the Communication-based defense approach across various swarm scenarios.



(a) 2 defenders vs 2 attackers scenario

(b) 3 defenders vs 3 attackers scenario

Figure 5.5: Figure presents the winning probability of the Communication-based defense approach for the symmetrical swarm scenarios



(a) 3 defenders vs 4 attackers scenario

(b) 4 defenders vs 3 attackers scenario

Figure 5.6: Figure presents the winning probability of the Communication-based defense approach for the asymmetrical larger swarm scenarios

The winning probability analysis underscores the robustness and adaptability of the defense approaches across diverse scenarios. The observed convergence of winning probabilities provides compelling evidence of the strategies' ability to effectively counter adversarial UAV swarms, even in scenarios with varying swarm sizes and complexities. These findings demonstrate the practicality and promise of our proposed defense mechanisms in real-world scenarios involving UAV swarm confrontations.

5.2.3 Winning Rate Comparative Analysis

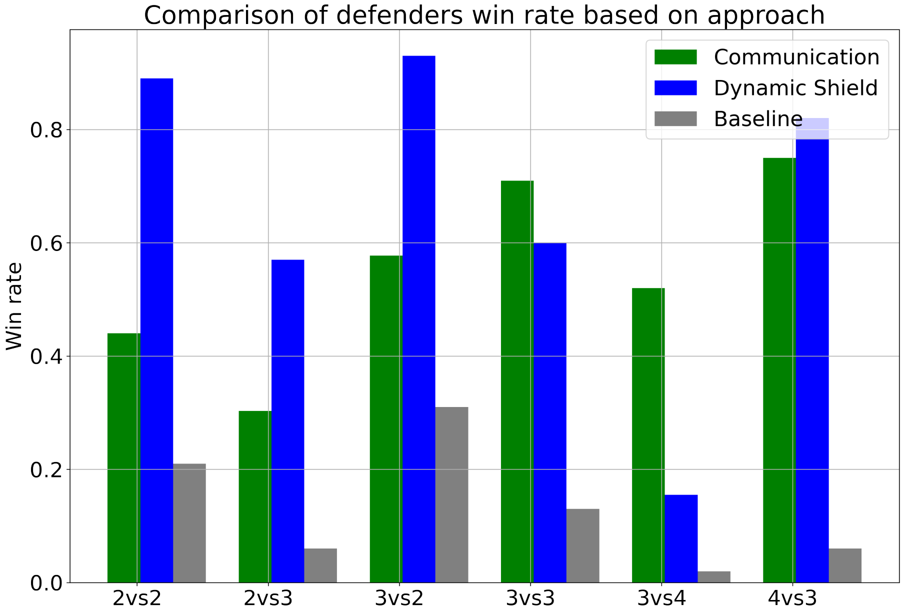


Figure 5.7: Comparison of the defenders' win rates between the Communication-based, Dynamic Shield, and baseline approaches, across various symmetrical and asymmetrical scenarios. Initially introduced in [69].

The winning rate is a crucial metric providing insights into the overall success of the defense mechanisms in diverse scenarios. Figure 5.7 offers a comparative analysis of the winning probabilities of the defending team for different scenarios and approaches. Specifically, the green bars denote the winning probability for the Communication-based approach, the blue bars represent the Dynamic Shield defense, and the gray bars illustrate the performance of the baseline approach.

Notably, in scenarios involving two attackers, the Dynamic Shield defense consistently outperforms the Communication-based approach, achieving impressive win rates ranging from 60% to 90%. This performance distinction remains significant even as the number of attackers increases. While the overall performance of the defense mechanisms experiences a slight decline with an increased number of attackers, the Dynamic Shield defense exhibits its efficacy when the number of defenders equals or surpasses the number of attackers. This distinction arises due to the Dynamic Shield defense's ability to effectively counter attacks from multiple directions, a challenge that is more difficult for fewer defenders to handle. Across various scenarios, our proposed defense mechanisms consistently attain winning rates exceeding 60%. Furthermore, it is important to note that the dynamic shield defense offers a more pragmatic solution by avoiding the need for communication channel access assumptions.

In light of the complexity inherent in the MADDPG algorithm, we opted to limit the number of agents to a total of 7 within the field. Future research directions might focus on refining algorithms and introducing macro-actions to accommodate larger swarm sizes. The outcomes of the performance evaluation showcase promising advancements and mark a significant step towards the establishment of non-confrontational defense strategies against adversarial UAV swarms.

5.3 Agent Behavior Observations

In this section, we discuss the behaviors exhibited by the UAV agents during the simulations of the Communication-based approach and the Dynamic Shield approach. These behaviors provide valuable insights into how the trained agents interact with each other and the dynamics of their strategic decision-making. Both approaches aim to generate effective non-confrontational countermeasures through cooperative-competitive interactions.

5.3.1 Communication-based Approach Observations

The Communication-based approach employs a mechanism of utilizing false information to mislead attacking UAVs. Through our observations of the simulation outcomes, several noteworthy insights emerged, which are described below, accompanied by screenshots taken from the live visualization of agent behavior during the simulated confrontations. The color codes for the visualization remain consistent across approaches – the green circle represents the target that adversarial UAVs (in red) try to reach and the defending UAVs (in blue) try to protect, while the black circle, if present, represents an artificial fake target used by the Communication-based approach.

- The attacking team's behavior appears to be influenced by a combination of the fake messages transmitted by the defending team and the swarming behavior shared by both teams. Defending agents exhibit swarming characteristics like attraction and repulsion, which seem to play a role in luring the attackers towards the fake target, along with the influence of the false signals.

- Interestingly, in certain instances, defending agents rush toward the real target in a race-like manner at the start of the game. This behavior is particularly pronounced when the attacking agents approach the real target from the start of the episode as well.

- Notably, when defending agents are not challenged by the attacking agents they circle around the fake target, potentially indicating an attempt to invite the attacking team to join them in that location.

- Training the Communication-based approach agents proved to be more time-consuming than the Dynamic Shield approach due to the complex reward structure. In instances where the model was not trained enough, issues in the cooperative agent behavior appeared, such as singular agents wandering off on their own and not following the expected objectives.

- Initial agent spawning positions appear to influence their subsequent strategies, with some games resembling races and others focusing on exploration of the environment.

- It is important to mention, that the Communication-based approach underwent a transformation during the course of the conducted research. Initially centered on luring attackers away from the real target, the approach was later redefined to focus on delaying the attackers from reaching the target throughout the game. This evolution led to behaviors sometimes reminiscent of the Dynamic Shield strategy, including shielding the real target from attackers which can be observed in Figure 5.8.

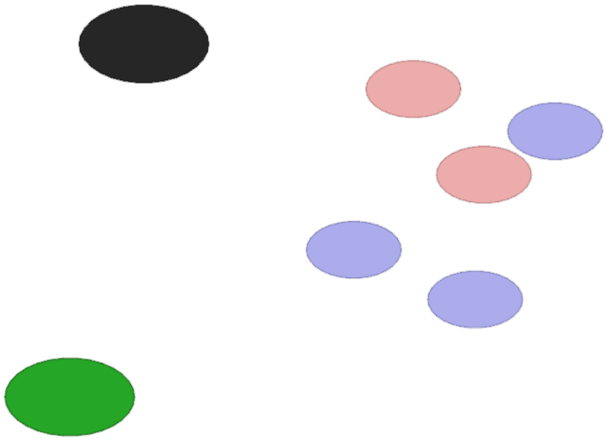


Figure 5.8: Communication-based approach simulation for the 3 defenders vs. 3 attackers scenario

5.3.2 Dynamic Shield Observations

The Dynamic Shield approach aims to shield the target from adversarial UAVs through cooperative swarming behaviors. Insights from the agent behaviors observed during simulations include:

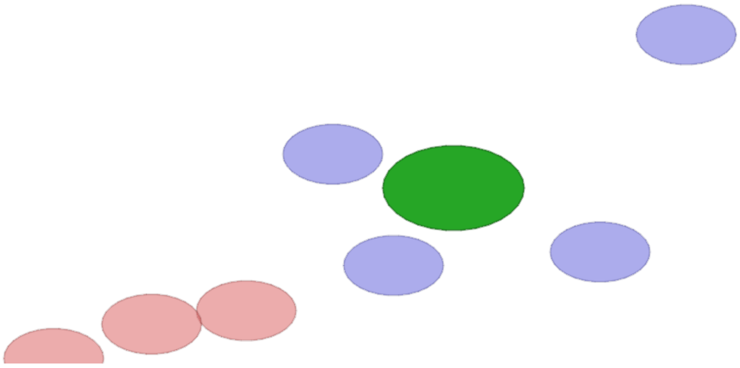


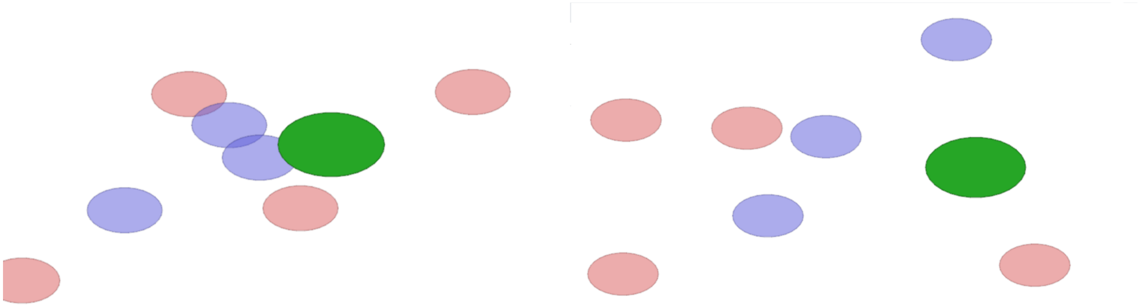
Figure 5.9: Dynamic Shield approach simulation for the 4 defenders vs. 3 attackers scenario. Note that the camera space for the visualization is predefined, whereas the space that the agents can move in is not limited, thus resulting in some UAV depictions being cut off at times, such as one of the adversarial agents here.

- At the start of each game, the defending agents promptly converge on the target, effectively forming a protective barrier around it to obstruct the attacking team. This behavior can be observed in Figure5.9.

- A team consisting of three or more agents often exhibits a strategic advantage, as the cooperative swarming behavior becomes more apparent with a larger number of agents.

- In a scenario involving 3 defenders and 4 attackers, as depicted in Figure5.10, an unexpected emergent behavior could be observed. One attacker would split from the group, approaching the target from behind while the remaining attackers engaged the defending agents from the front. This behavior resulted in a rare overwhelming victory for the attacking team.

In summary, the behaviors exhibited by UAV agents in both the Communication-based and Dynamic Shield approaches provide valuable insights into their decision-making processes, cooperative strategies, and interactions with adversaries. These observations contribute to a deeper understanding of how game-based training can shape agent behaviors in cooperative-competitive scenarios.



(a) First instance of an emergent behavior (b) Second instance of an emergent behavior

Figure 5.10: Dynamic Shield approach simulation for the 3 defenders vs. 4 attackers scenario

Chapter 6

Discussion

In this chapter, we elaborate on our research findings, exploring insights from non-confrontational UAV countermeasures using game-based and Multi-Agent Reinforcement Learning approaches. We address research questions, interpret results, and discuss practical implications.

We begin the chapter with an "Interpretation of Results," examining trends, performance metrics, and agent behaviors. Then, we proceed to "Addressing Research Questions," evaluating optimal methods for countermeasures and MARL techniques.

The chapter continues further by considering the broader implications of our work. We explore real-world deployment possibilities, discuss gamification, and draw comparisons with online multiplayer games. Lastly, we reflect on the abstraction and disconnect present in video games and real-world UAV operations.

This chapter offers an overview of our research's key insights and sets the stage for future research in UAV technology and cooperative-competitive multi-agent systems.

6.1 Interpretation of Results

The results obtained from the experiments conducted in this study provide valuable insights into the behavior and performance of the Communication-based and Dynamic Shield approaches. This section delves into the interpretation of these findings, discussing aspects such as the stability of rewards, sensitivity to swarm size, comparative agent behavior, and the effectiveness of non-confrontational strategies. Through an examination of these results, we aim to derive an understanding of the capabilities and limitations of each approach in simulating cooperative-competitive scenarios for training UAV agents.

- Stability of Rewards

  Achieving reward stability in the training process requires a substantial number of iterations, typically exceeding 500,000 episodes or more. It is notable that the Dynamic Shield approach tends to reach reward stability faster when compared to the Communication-based approach. This discrepancy in training times can be attributed to the complexity of the reward structure of the Communication-based defense, which naturally extends the training duration. Furthermore, the results indicate that larger swarm sizes necessitate a longer training period to attain reward stability, regardless of the chosen approach.

- Sensitivity to Swarm Size

  The study showcased interesting insights concerning swarm size. Cooperative behavior appears to be more pronounced when a team comprises three or more agents. Interestingly, addition of even a single agent can significantly impact team performance, as exemplified by the scenario featuring the Dynamic Shield approach with 3 attackers and 4 adversaries. In this particular case, an unexpected emergent behavior led to the rare occurrence of an overwhelming victory by the attacking team.

- Comparing Performance of the Approaches

  A comparative analysis of the behaviors exhibited by the two approaches reveals distinct characteristics. The Dynamic Shield approach consistently demonstrates more cohesive and stable performance, showcasing its effectiveness across various swarm sizes. Notably, its cooperative behavior remains robust, making it suitable for both smaller and larger teams. While the Communication-based approach provides valuable insights, produces results with high winning probability for defenders, and offers an alternative perspective, the Dynamic Shield approach stands out for its robustness and scalability. This is further illustrated by the bar plot from the previous chapter, which showcases the Dynamic Shield's performance.

- Non-Confrontational Strategies

  Both the Communication-based and Dynamic Shield approaches exhibit non-confrontational behaviors. However, it's important to acknowledge that further research and testing are necessary, especially in more complex environments, before stating that these approaches can ensure the absence of collateral damage.

6.2 Multi-Agent Simulation as a Game

The utilization of trained multi-agent reinforcement learning models within simulated environments, such as the Multi-Agent Particle Environment in this study, could be compared to playing a game. Especially the visualization of the agents (which can be seen in the figures of the previous chapter, such as 5.8) shows them interacting in real-time, and can be seen as a game prototype with minimal interactivity - it could also be compared to two computer-controlled characters (non-player characters) fighting. Furthermore, it would be possible to enhance the visualization application further, for example by including interactive elements like a user interface or selectable agent models, to make it more game-like. Regardless of the visual aspect, the open-source nature of this project implies that anyone with an interest in the field can access and experiment with these models in dynamic and customizable scenarios. One could modify parameters like agent count, target specifications, colors, and agents' objectives, enhancing the project's interactability and also turning it into a game.

Moreover, this multi-agent simulation environment holds significant potential within educational contexts. It offers an accessible means to introduce students to the sometimes complex dynamics of multi-agent systems. Through these simulations, students can gain hands-on experience with concepts such as cooperation, competition, and emergent behavior, fostering a deeper understanding of the complexities underlying multi-agent interactions.

Importantly, cooperative-competitive game-based simulations inherently offer engaging observation experiences. The game-like nature of the implemented approaches produces not only analytical insights but also an element of entertainment. Observing intelligent agents strategically navigate scenarios and adapt their behaviors creates a compelling narrative. This dynamic could even be reminiscent of watching sports matches or streams of multiplayer games, although presented in a much more abstracted manner. The cooperation and competition between agents, along with the emergence of diverse strategies, form an enticing narrative that could potentially serve a purpose beyond just the technical analysis.

Looking ahead, the potential for expanding the 2D MPE environment into more intricate representations holds exciting possibilities. The incorporation of 3D environments, perhaps within platforms like Unity, combined with the utilization of creative assets for agents and targets, could amplify the visual aspects of the simulations. This expansion could further enhance the game-like nature of the simulation, fostering even more captivating observations and explorations.

In essence, viewing multi-agent simulations as games underlines the versatile nature of these environments. They serve as accessible educational tools, offer interactive experimentation opportunities, and could potentially present engaging narratives. The evolution of these simulations toward more intricate representations only amplifies the potential for broader engagement and exploration in the realm of multi-agent systems.

6.3 Practical Implications and Applications

In this section, we discuss the practical implications and applications of our research findings. We explore the potential for real-world deployment of our game-based defense strategies and their applicability in relevant fields, and we also consider the ethical and societal implications stemming from the adoption of these approaches.

6.3.1 Real-World Deployment

Our UAV countermeasure system holds promising potential for real-world applications in various domains. One notable application is in the realm of UAV defense. As the deployment of UAVs becomes more prevalent, the need for effective countermeasures to defend against adversarial attacks increases. By training defending UAV swarms using our system, organizations can enhance their ability to protect critical assets, infrastructure, and events from potential threats posed by rogue UAVs. Additionally, the principles underlying our system's non-confrontational strategies could find applications in scenarios involving security patrols, surveillance, and monitoring tasks. Beyond UAV defense, our approach could also be adapted for use in video game environments. Game developers could integrate our trained agents into games to enhance the realism and complexity of AI-controlled agents, making game experiences more challenging and engaging for players.

6.3.2 Applications in Relevant Fields

The findings of this study hold promising applications across various fields. In the realm of Multi-Agent Reinforcement Learning, the non-confrontational countermeasure strategies developed through cooperative-competitive game-based simulations could greatly enhance the efficiency and effectiveness of UAV swarm operations. MARL is a rapidly evolving field that focuses on agents interacting with other agents to optimize their behaviors. Our system's approach to training both defending and adversarial UAV agents using the MADDPG algorithm aligns well with MARL's objectives. The countermeasure system's ability to generate non-confrontational strategies showcases the potential to advance the study of cooperative-competitive scenarios within MARL. By using this system as a benchmark, researchers in MARL can study and develop novel algorithms that promote non-confrontational behaviors in various multi-agent settings. However, an important challenge lies in generalizing these findings to more complex scenarios with larger numbers of agents and diverse environments. Scalability and adaptability of the countermeasure techniques are areas of concern that need to be addressed to ensure practical application across a range of MARL problems. Nevertheless, our research presents an opportunity to contribute to the ongoing research efforts in MARL by introducing a concrete framework for generating non-confrontational strategies.

6.3.3 Ethical and Societal Implications

The deployment of UAV countermeasure systems and AI-driven strategies raises several ethical and societal implications. The emphasis on non-confrontational methods in our research holds the potential to reshape the landscape of warfare and security operations. If adopted and integrated into real-world defense strategies, these methods could contribute to reducing casualties and collateral damage by avoiding direct confrontations. This shift aligns with broader discussions on the ethics of warfare and the human cost of armed conflicts. Furthermore, the convergence of AI and game-based simulations calls for conscientious representation in video games. As games increasingly draw inspiration from real-world scenarios, developers should be mindful of accurately representing the complexities and ethical considerations associated with military and security operations. By portraying the disconnect between human operators and UAVs, game developers can encourage players to reflect on the potential dehumanization of warfare and the ethical dilemmas posed by AI-driven decision-making.

The practical implications and applications of this research extend to various domains, including the advancement of MARL, educational contexts, and real-world defense scenarios. The ethical and societal implications underscore the potential for promoting non-confrontational approaches, interdisciplinary collaboration, and the responsible representation of complex issues in AI-driven simulations and games.

6.4 Relevance of Non-Confrontational Methods

One observation that emerged during the course of this research relates to the lack of non-confrontational approaches within the field of machine learning-generated UAV countermeasures. A review of existing literature and methodologies showcased a notable lack of studies related to non-confrontational strategies. The field of multi-agent systems and artificial intelligence appears to lean toward more competitive and adversarial paradigms. While these approaches contribute crucial insights into strategic decision-making and emergent behaviors, the absence of proportionate attention to non-confrontational strategies is an issue worthy of investigating.

Furthermore, this observation prompts broader considerations about the role that researchers, especially in the field of AI, play in the design and development of countermeasures. There is oftentimes a noticeable gap between the theoretical formulations of algorithms and their practical deployment. This is particularly relevant to scenarios such as the MPE, where agents are abstracted into particle entities. The level of abstraction in such simulations, while making them computationally manageable, emphasizes the divide between the technical research and the realities of deploying these systems.

Moreover, this abstraction resembles the relationship between UAVs and their operators. The detachment between the design and execution of algorithms within a multi-agent simulation is reminiscent of the disconnect between UAVs and their remote pilots, which poses potential ethical dilemmas. This disconnect could potentially raise concerns mentioned earlier regarding the ethics of game-based training, where the removal of direct human involvement in decision-making could lead to the dehumanization of warfare.

The noticeable absence of non-confrontational strategies in the realm of machine learning-generated UAV countermeasures prompts a thoughtful examination of the research focus in this area. Hopefully, this encourages AI researchers to analyze the underlying ethical considerations and preferences guiding their work. Furthermore, the similarities between the abstraction in simulations and the detachment in human-machine interactions could also facilitate reflection on the wider impacts of our research efforts.

6.5 Gamification and Multi-Agent Cooperative-Competitive Systems

In recent years, the integration of game elements and mechanics, commonly referred to as gamification, has gained significant attention across various domains. The concept of gamification involves applying game-design principles and techniques to non-game contexts to engage users and enhance their experiences. When considering the application of MARL techniques to domains such as UAV defense and coordination, the idea of gamification can provide valuable insights and possibilities.

One of the core aspects of gamification is the engagement factor it introduces. By incorporating game-like elements, such as rewards, challenges, and competition, the process of training and optimizing MARL agents can be made more interactive and dynamic. In the context of UAV defense, gamification can be used to create training environments that simulate real-world scenarios more accurately. For example, the concept of "missions" or "levels" can be introduced, each presenting distinct challenges for the agents to overcome.

Furthermore, gamification can be instrumental in causing the agents to learn and adapt continuously. As agents strive to achieve objectives in a simulated environment, they may learn to collaborate more effectively, respond to dynamic threats, and optimize their strategies to achieve higher scores or rewards. These learned behaviors can potentially translate into more effective real-world deployment of UAVs in defense scenarios.

It is worth noting that while gamification offers a novel approach to enhancing MARL training, its implementation requires careful consideration. The balance between introducing game elements and maintaining the integrity of the learning process should be monitored to ensure that the agents' learned behaviors remain applicable to real-world situations.

As the field of MARL continues to evolve, exploring the integration of gamification concepts can open new avenues for research and innovation. By embracing this interdisciplinary approach, we can harness the power of games to enhance the capabilities of UAVs and other multi-agent systems, ultimately contributing to more efficient and effective defense strategies.

While it is accurate that agents in the context of MARL do not possess human-like emotions, the concept of gamification can be applied to their training and optimization in a more abstract manner. Instead of directly aiming for an experience that humans would perceive as "fun," the focus shifts towards leveraging game-like mechanisms to enhance learning, engagement, and performance.

The core principle of gamification involves introducing elements that encourage active participation, trial-and-error exploration, and the pursuit of goals. In the case of MARL agents, these elements could include reward structures, challenges, objectives, and dynamic scenarios. Agents are motivated not by human-like emotions but by mathematical algorithms designed to seek optimal strategies to achieve predefined objectives. This process aligns with the underlying philosophy of gamification—enhancing engagement and performance through structured mechanisms.

While agents may not feel "fun" in the same sense as humans, the use of gamification techniques can still lead to several benefits. For instance, introducing dynamic challenges and varying levels of difficulty can stimulate agents to explore different strategies and adapt to changing scenarios more effectively. The competitive aspects of gamification can drive agents to continually improve their performance and collaborate more intelligently.

In essence, gamification in the context of agents is about creating an environment that encourages exploration, learning, and optimization, even if these processes are devoid of human-like emotional experiences. By designing training scenarios that utilize game-like mechanisms, we can create an environment in which agents naturally develop more sophisticated strategies and behaviors, aligning with the goals of enhancing their effectiveness in real-world applications.

In summary, while the term "gamification" may indeed evoke human-centered concepts, its application to agents involves creating structured environments that promote learning, adaptation, and performance improvement. The aim is to leverage these mechanisms to achieve optimal results in a way that aligns with the capabilities and objectives of the agents themselves, even if their experiences are fundamentally different from those of humans.

Despite the lack of prior research in this domain, the convergence of game theory, reinforcement learning, and gamification could lay the groundwork for an interesting discussion, which could signify the start of innovative methodologies that bridge disciplinary boundaries and encourage new approaches for enhancing agents' performance and adaptability in dynamic scenarios.

6.6 Comparative Analysis to Online Multi-Player Games

Within the domain of multi-agent reinforcement learning, which involves agents cooperating and competing within complex environments, parallels can be observed between the cooperative-competitive MARL approach in our study and the context of online multiplayer games. These parallels offer insights into strategic dynamics and decision-making processes in both artificial intelligence and human-player interactions.

In our research's cooperative-competitive approach, UAV agents are organized into teams to achieve objectives, either defending or attacking a target. This framework of cooperative collaboration and competitive engagement shares conceptual similarities with online multiplayer games, particularly exemplified by games like *Valorant* [65]. Further examples include *Overwatch* [5] and *Counter-Strike: Global Offensive* [79]. In these games, players navigate intricate virtual environments, working collectively to secure goals while concurrently countering adversaries.

Valorant is a widely played online multiplayer game characterized by tactical team-based engagements. In the game, two teams of players compete to accomplish objectives on distinct maps. One example of such an objective would be for one team to defend a certain spot on the map, while the other team tries to breach their defenses and secure the location – which may involve strategies strikingly similar to the ones exhibited by the trained agents in this work. Each player assumes a unique role, utilizing diverse abilities and strategies to secure victory. Similarly, in our research, UAV agents are organized into teams with shared objectives of defense or attack. The strategic interplay between cooperative efforts and competitive maneuvers, evident in both *Valorant* and our study, underscores the core theme of achieving objectives while responding to dynamic challenges. While the embodiment of agents differs – AI agents in our study versus human players in *Valorant* – the parallel highlights the common thread of cooperative-competitive dynamics central to both contexts.

As mentioned, a fundamental distinction emerges in participant embodiment. Embodiment, in the context of artificial intelligence and cognitive science, refers to the concept of giving or attributing physical form to an entity, often an agent (see also [7]). It involves the idea that cognition and intelligence are not solely products of abstract computation, but are influenced by the entity's interactions with its environment and its physical characteristics. In other words, embodiment recognizes the role of the physical body or representation in shaping an entity's perceptions, actions, and overall cognitive processes.

In our study, AI agents adhere to learned strategies, whereas in games like *Valorant*, human players engage in real-time decision-making, informed by personal intuition and creativity, in addition to a possible base of learned strategies. This contrast underscores distinct attributes and challenges within each domain.

The comparison between these two domains sheds light on decision-making complexities, team dynamics, and adaptability. The mutual pursuit of objectives and navigation of intricate environments by simulation agents and human players in games like *Valorant* illuminate parallels and disparities in underlying mechanisms and motivations. Acknowledging these parallels not only advances our comprehension of AI research but also forges connections between simulations and relatable human activities. This connection fosters deeper explorations into teamwork, strategy, and decision-making dynamics, contributing to a more comprehensive understanding of both artificial and human intelligence in dynamic scenarios.

6.7 Comparing Abstraction and Disconnect: Video Games and Real-world Operators

The notion of abstraction and disconnect within the context of UAVs finds resonance not only in academic debates but also in the realm of popular culture. By examining parallels between the portrayal of UAVs in video games and the experiences of real-world operators, we can gain valuable insights into the complexities of human-machine interaction and decision-making.

In video games, UAVs are often represented as remote-controlled entities, embodying a form of abstraction that mirrors their real-world counterparts. Players assume the role of operators, remotely guiding these vehicles to achieve objectives. This abstraction, however, raises intriguing questions about the portrayal of warfare, as players manipulate agents without direct physical presence, a concept that resonates with the real-world detachment experienced by operators.

Interestingly, this abstraction extends beyond video games into real-world military operations, where UAV operators similarly engage in remote warfare from a position of physical detachment. This comparison highlights the delicate balance between operational effectiveness and ethical considerations. Just as players in video games can make strategic decisions with a sense of detachment, real-world operators face similar challenges in maintaining a sense of connection to the consequences of their actions.

A notable literary work that explores these themes is *Ender's Game* by Orson Scott Card [8]. The novel delves into the training of young children as remote operators of space fleets, where the disconnect between actions and consequences is central to the narrative. This book serves as a reminder that the ethical complexities posed by abstraction and detachment are not confined to the realms of academia and technology but extend into cultural narratives.

By acknowledging these parallels, we are encouraged to reflect on the broader implications of our research endeavors. The convergence of video game representations and real-world operational experiences underscores the significance of ethical considerations and the human impact of remote-controlled warfare. Through interdisciplinary exploration, we can engage in thought-provoking conversations about the intersections of technology, morality, and human agency.

This comparison invites us to consider not only the technical aspects of our research but also the broader societal and ethical ramifications of our contributions. It underscores the relevance of our study in a world where abstraction and detachment are becoming increasingly intertwined with the modern landscape of conflict and technology.

Chapter 7

Conclusion and Outlook

This chapter explores the core objectives, findings, and future directions. We discuss the significance of our work, highlighting the game-based framework developed for non-confrontational UAV countermeasures using the Multi-Agent Deep Deterministic Policy Gradient algorithm. Our research explores the cooperative-competitive dynamics within multi-agent systems, emphasizing the potential for enhancing UAV operations' safety and efficacy. Additionally, we outline promising avenues for future research, addressing limitations, integrating advanced technologies, and advancing the field of UAV swarm coordination and countermeasures.

7.1 Findings

The overarching objectives of this thesis were to develop a game-based framework for generating non-confrontational UAV countermeasures using machine learning techniques, particularly the Multi-Agent Deep Deterministic Policy Gradient algorithm. This framework hinged on the combination of methodologies from Machine Learning, Game Theory, and Swarm Intelligence to devise novel and effective countermeasures against UAV swarm attacks.

Throughout the course of this research, several key objectives were addressed:

- Objective 1: Game-Based Framework Development

  The first objective centered on the establishment of a comprehensive game-based framework to generate non-confrontational countermeasures. The framework was successfully designed and implemented during the course of this study.

- Objective 2: Investigating Non-Confrontational Approaches

  A central facet of this research was the exploration and evaluation of non-confrontational approaches, encompassing misleading, distracting, and delaying tactics. These approaches were effectively designed to counteract UAV swarm attacks.

- Objective 3: Designing Cooperative-Competitive Multi-Agent Systems

  To execute these non-confrontational strategies, a cooperative-competitive multi-agent system was designed. This system enabled agents to collaborate, compete, and adapt dynamically within the confines of UAV swarm scenarios.

- Objective 4: MADDPG Algorithm Evaluation

  An important objective was the assessment of the MADDPG algorithm's effectiveness and adaptability in learning and coordinating UAV countermeasure strategies. This evaluation served as a testing ground for the algorithm's potential applicability in dynamic, real-world environments. While limited in its potential for scalability, the MADDPG algorithm proved to be effective within the scope of this research.

- Objective 5: Performance Assessment via Simulations

  Finally, to evaluate the efficacy of the proposed methods, extensive simulations were conducted. These simulations provided valuable insights into the performance of non-confrontational countermeasures across various UAV swarm scenarios.

This research has provided an understanding of how game-based frameworks and MADDPG-driven machine-learning techniques can be used to tackle the challenge of UAV swarm coordination. The findings of this thesis underscore the potential for innovative, non-confrontational strategies to enhance the safety and efficacy of UAV operations across diverse domains.

7.2 Future Directions

This section explores potential directions for future research, addressing limitations, integrating additional techniques and technologies, and refining the methodologies employed.

7.2.1 Further Research

Future research could focus on the development of new game-based defense strategies or the improvement of existing ones. The Communication-based and Dynamic Shield approaches provide a foundational framework, offering an opportunity to design new methodologies employing multi-agent reinforcement learning. The objective is to advance non-confrontational countermeasures, ensuring the safety and efficacy of UAV operations.

Another possible research direction delves into the possibility of combining the Communication-based and Dynamic Shield approaches. The potential between these two methodologies, as advised by a peer reviewer, is a direction that is worth investigating. Such integration may result in a comprehensive defense strategy with dynamic adaptability to evolving threat scenarios.

7.2.2 Addressing Limitations

Acknowledging and mitigating research limitations is critical for fostering more robust investigations. The constraint imposed by the MADDPG algorithm regarding swarm size is a recognized limitation. Future research could explore hybrid approaches that combine MADDPG with complementary machine-learning techniques. This fusion might enhance scalability and overall performance, enabling the coordination of larger UAV swarms.

The current reliance on 2D simulated environments introduces a level of abstraction. To attain more realistic results, transitioning to 3D simulations using platforms like Unity3D is a promising direction. This upgrade offers the potential for a more realistic emulation of real-world conditions, producing deeper insights into UAV swarm behaviors.

Assumptions about certain agent capabilities, such as access to adversaries' communication channels or knowledge of adversarial agent positions, represent limitations. To alleviate these constraints, future research might focus on iteratively refining strategies, and reducing dependency on these assumptions.

7.2.3 Integration of additional Techniques and Technologies

The transition to 3D simulation environments, facilitated by platforms like Unity3D, is instrumental in achieving more complex, and more realistic experimental settings. This technological shift allows for a more accurate assessment of the strategies' real-world applicability.

Potentially combining the MADDPG algorithm and Graph Neural Networks is another promising direction to consider. GNNs have demonstrated proficiency in modeling complex relationships within graph-structured data. Their application to UAV swarm coordination holds the potential for breakthroughs in optimizing swarm behaviors.

Upon establishing the safety and reliability of realistic simulations, the logical progression is to conduct real-life experiments involving physical UAVs. These experiments bridge the gap between simulation and practical deployment, facilitating the implementation of non-confrontational countermeasures and machine-learning-guided swarm coordination.

In conclusion, this section outlines the prospective research trajectories, emphasizing the need to address limitations and harness advanced technologies. These future directions are pivotal in the continuous evolution of the UAV swarm coordination and countermeasures field, ultimately contributing to the enhanced safety and efficiency of UAV operations across diverse applications.

Bibliography

[1] G. E. M. Abro, S. A. B. Zulkifli, R. J. Masood, V. S. Asirvadam, and A. Laouti. Comprehensive review of UAV detection, security, and communication advancements to prevent threats. *Drones*, 6(10):284, 2022.

[2] Aonic. Unmanned aerial vehicles (UAVs) revolutionizing industries and beyond. https://www.aonic.com/my/blogs-drone-technology/unmanned-aerial-vehicles-uavs-revolutionizing-industries-and-beyond/ . Accessed: 30.06.2023.

[3] M. S. Atkin, D. L. Westbrook, and P. R. Cohen. Capture the flag: Military simulation meets computer games. In *Proceedings of AAAI Spring Symposium Series on AI and Computer Games*, pages 1–5, 1999.

[4] Bat Cave Games. *VelociDrone - FPV Racing Simulator*. Bat Cave Games, 2016. https://www.velocidrone.com/. Accessed: 20.08.2023.

[5] Blizzard Entertainment. *Overwatch*. Blizzard Entertainment, 2016. https://overwatch.blizzard.com/en-us/. Accessed: 01.09.2023.

[6] E. Bondi, D. Dey, A. Kapoor, J. Piavis, S. Shah, F. Fang, B. Dilkina, R. Hannaford, A. Iyer, L. Joppa, et al. Airsim-w: A simulation environment for wildlife conservation with UAVs. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, pages 1–12, 2018.

[7] P. Calvo and T. Gomila. *Handbook of cognitive science: An embodied approach*. Elsevier, 2008.

[8] O. S. Card. *Ender's game*. Tor Books, 1985.

[9] H. Cardona-Reyes, C. Trujillo-Espinoza, C. Arevalo-Mercado, and J. Muñoz-Arteaga. Training of drone pilots through virtual reality environments under the gamification approach in a university context. *Interaction Design and Architecture(s) Journal - IxD&A*, 49:64–83, 2021.

[10] CD Projekt Red. *Cyberpunk 2077*. CD Projekt, 2020. https://www.cyberpunk.net/. Accessed: 22.08.2023.

[11] V. Chamola, P. Kotesh, A. Agarwal, N. Gupta, M. Guizani, et al. A comprehensive review of unmanned aerial vehicle attacks and neutralization techniques. *Ad hoc networks*, 111:102324, 2021.

[12] Y. L. Chan et al. Dynamic geo-fence for drone. *US Appl*, (14/951,533), 2017.

[13] W. Chen, X. Meng, J. Liu, H. Guo, and B. Mao. Countering large-scale drone swarm attack by efficient splitting. *IEEE Transactions on Vehicular Technology*, 71(9):9967–9979, 2022.

[14] B. Cowan and B. Kapralos. A survey of frameworks and game engines for serious game development. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*, pages 662–664. IEEE, 2014.

[15] L. Davies, R. C. Bolam, Y. Vagapov, and A. Anuchin. Review of unmanned aircraft system technologies to enable beyond visual line of sight (BVLOS) operations. In *2018 X International conference on electrical power drive systems (ICEPDS)*, pages 1–6. IEEE, 2018.

[16] M. Dil, M. U. Khan, M. Z. Alam, F. A. Orakazi, Z. Kaleem, and C. Yuen. Safespace mfnet: precise and efficient multifeature drone detection network. *arXiv preprint arXiv:2211.16785*, pages 1–13, 2022.

[17] C. Dumitrescu, M. Minea, I. M. Costea, I. Cosmin Chiva, and A. Semenescu. Development of an acoustic system for UAV detection. *Sensors*, 20(17):4870, 2020.

[18] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc. Micro-UAV detection and classification from rf fingerprints using machine learning techniques. In *2019 IEEE Aerospace Conference*, pages 1–13. IEEE, 2019.

[19] P. G. Fahlstrom, T. J. Gleason, and M. H. Sadraey. *Introduction to UAV systems*. John Wiley & Sons, 2022.

[20] S. X. Fang, S. O'Young, and L. Rolland. Development of small UAS beyond-visual-line-of-sight (BVLOS) flight operations: System requirements and procedures. *Drones*, 2(2):13, 2018.

[21] Federal Aviation Administration. Unmanned aircraft systems (UAS) frequently asked questions. `https://www.faa.gov/uas/faqs/`, 2020. Accessed: 30.06.2023.

[22] I. Fehérvári, W. Elmenreich, et al. Evolving neural network controllers for a team of self-organizing robots. *Journal of Robotics*, 2010, 2010.

[23] A. Fernández-Ares, A. M. Mora, M. García-Arenas, J. J. M. Guervós, P. García-Sánchez, and P. A. Castillo. Co-evolutionary optimization of autonomous agents in a real-time strategy game. In *Applications of Evolutionary Computation: 17th European Conference, EvoApplications 2014, Granada, Spain, April 23-25, 2014, Revised Selected Papers 17*, pages 374–385. Springer, 2014.

[24] S. Feroz and S. Abu Dabous. UAV-based remote sensing applications for bridge condition assessment. *Remote Sensing*, 13(9):1809, 2021.

[25] D. Fogel. An introduction to simulated evolutionary optimization. *IEEE Transactions on Neural Networks*, 5(1):3–14, 1994.

[26] D. Fudenberg and J. Tirole. *Game theory.* MIT press, 1991.

[27] S. Givigi Jr and H. Schwartz. A game theoretic approach to swarm robotics. *Applied Bionics and Biomechanics*, 3(3):131–142, 2006.

[28] E. R. Goossen and S. D. Martinez. Catch and snare system for an unmanned aerial vehicle, 2013. Patent filed via `https://patents.google.com/patent/US8375837B2/en`. Accessed: 30.08.2023.

[29] J. Gregory. *Game engine architecture.* CRC Press, 2018.

[30] G. S. Hadi, R. Varianto, B. Trilaksono, and A. Budiyono. Autonomous UAV system development for payload dropping mission. *The Journal of Instrumentation, Automation and Systems*, 1(2):72–22, 2014.

[31] D. He, G. Yang, H. Li, S. Chan, Y. Cheng, and N. Guizani. An effective countermeasure against UAV swarm attack. *IEEE Network*, 35(1):380–385, 2020.

[32] F. Hoffmann, M. Ritchie, F. Fioranelli, A. Charlish, and H. Griffiths. Micro-doppler based detection and tracking of UAVs with multistatic radar. In *2016 IEEE radar conference (RadarConf)*, pages 1–6. IEEE, 2016.

[33] Y. Hu and W. Meng. Rosunitysim: Development and experimentation of a real-time simulator for multi-unmanned aerial vehicle local planning. *Simulation*, 92(10):931–944, 2016.

[34] M. Hüttenrauch, A. Šošić, and G. Neumann. Guided deep reinforcement learning for swarm systems. *arXiv preprint arXiv:1709.06011*, 2017.

[35] V. U. Ihekoronye, S. O. Ajakwe, D.-S. Kim, and J. M. Lee. Cyber edge intelligent intrusion detection framework for UAV network based on random forest algorithm. In *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1242–1247. IEEE, 2022.

[36] A. Jara-Olmedo, W. Medina-Pazmi no, R. Mes'ıas, B. Araujo-Villaroel, W. G. Aguilar, and J. A. Pardo. Interface of optimal electro-optical/infrared for unmanned aerial vehicles. In *Developments and Advances in Defense and Security: Proceedings of the Multidisciplinary International Conference of Research Applied to Defense and Security (MICRADS 2018)*, pages 372–380. Springer, 2018.

[37] J. Kennedy. Swarm intelligence. In *Handbook of nature-inspired and innovative computing: integrating classical models with emerging technologies*, pages 187–219. Springer, 2006.

[38] V. V. Klemas. Coastal and environmental remote sensing from unmanned aerial vehicles: An overview. *Journal of coastal research*, 31(5):1260–1267, 2015.

[39] V. N. Kolokoltsov and O. A. Malafeyev. *Understanding game theory: introduction to the analysis of many agent systems with competition and cooperation*. World scientific, 2020.

[40] R. N. Landers, E. M. Auer, A. B. Collmus, and M. B. Armstrong. Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming*, 49(3):315–337, 2018.

[41] R. N. Landers and A. K. Landers. An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, 45(6):769–785, 2014.

[42] B. G. León and F. Belardinelli. Extended markov games to learn multiple tasks in multi-agent reinforcement learning. *arXiv preprint arXiv:2002.06000*, 2020.

[43] D. Leonard. Unsettling the military entertainment complex: Video games and a pedagogy of peace. *Studies in Media & Information Literacy Education*, 4(4):1–8, 2004.

[44] Y. Li, C. Fu, F. Ding, Z. Huang, and G. Lu. Autotrack: Towards high-performance visual tracking for UAV with automatic spatio-temporal regularization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11923–11932, 2020.

[45] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[46] Q. Liu, M. He, D. Xu, N. Ding, and Y. Wang. A mechanism for recognizing and suppressing the emergent behavior of UAV swarm. *Mathematical Problems in Engineering*, 2018:1–15, 2018.

[47] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems*, 30:1–12, 2017.

[48] LuGus Studios. *LiftOff - FPV Drone Racing.* LuGus Studios, 2018. `https://store.steampowered.com/app/410340/Liftoff_FPV_Drone_Racing/`. Accessed: 20.08.2023.

[49] G. Lykou, D. Moustakas, and D. Gritzalis. Defending airports from uas: A survey on cyber-attacks and counter-drone sensing technologies. *Sensors*, 20(12):3537, 2020.

[50] S. López, J.-A. Cervantes, S. Cervantes, J. Molina, and F. Cervantes. The plausibility of using unmanned aerial vehicles as a serious game for dealing with attention deficit-hyperactivity disorder. *Cognitive Systems Research*, 59:160–170, 2020.

[51] I. Millington. *AI for Games.* CRC Press, 2019.

[52] S. A. H. Mohsan, M. A. Khan, F. Noor, I. Ullah, and M. H. Alsharif. Towards the unmanned aerial vehicles (UAVs): A comprehensive review. *Drones*, 6(6):147, 2022.

[53] Monolith Productions. *F.E.A.R.* Vivendi Universal Games, 2005.

[54] M. Nagai, T. Chen, A. Ahmed, and R. Shibasaki. UAV borne mapping by multi sensor integration. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 37:1215–1221, 2008.

[55] M. A. Nowak and K. Sigmund. Evolutionary dynamics of biological games. *science*, 303(5659):793–799, 2004.

[56] A. Nowé, P. Vrancx, and Y.-M. De Hauwere. Game theory and multi-agent reinforcement learning. *Reinforcement Learning: State-of-the-Art*, pages 441–470, 2012.

[57] J. Olson and M. Rashid. Modern drone warfare: An ethical analysis. In *2013 American Society for Engineering Education Southwest Section Conference, http://se. asee. org/proceedings/ASEE2013/Papers2013/157. PDF [10 czerwca 2015]*, 2013.

[58] OpenAI. ChatGPT: A large language model by OpenAI. `https://openai.com/research/chatgpt`, 2021. Accessed on September 1, 2023. Version: GPT-3.5.

[59] J. Orkin. Three states and a plan: the ai of fear. In *Game developers conference*, volume 2006, page 4. CMP Game Group SanJose, California, 2006.

[60] G. Panice, S. Luongo, G. Gigante, D. Pascarella, C. Di Benedetto, A. Vozella, and A. Pescapè. A SVM-based detection approach for GPS spoofing attacks to UAV. In *2017 23rd International Conference on Automation and Computing (ICAC)*, pages 1–11. IEEE, 2017.

[61] S. Park, H. T. Kim, S. Lee, H. Joo, and H. Kim. Survey on anti-drone systems: Components, designs, and challenges. *IEEE Access*, 9:42635–42659, 2021.

[62] S. Parsons and M. Wooldridge. Game theory and decision theory in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 5:243–254, 2002.

[63] E. Politi, I. E. Panagiotopoulos, I. Varlamis, and G. Dimitrakopoulos. A survey of UAS technologies to enable beyond visual line of sight (BVLOS) operations. In *VEHITS*, pages 505–512, 2021.

[64] H. D. Purnomo and H.-M. Wee. Soccer game optimization: an innovative integration of evolutionary algorithm and swarm intelligence algorithm. In *Meta-heuristics optimization algorithms in engineering, business, economics, and finance*, pages 386–420. IGI Global, 2013.

[65] Riot Games. *Valorant*. Riot Games, 2020. `https://playvalorant.com/`. Accessed: 22.08.2023.

[66] W. H. Sandholm. *Population games and evolutionary dynamics.* MIT press, 2010.

[67] E. Semsar-Kazerooni and K. Khorasani. Multi-agent team cooperation: A game theory approach. *Automatica*, 45(10):2205–2213, 2009.

[68] K. Shao, Z. Tang, Y. Zhu, N. Li, and D. Zhao. A survey of deep reinforcement learning in video games. *arXiv preprint arXiv:1912.10944*, 2019.

[69] J. Simonjan, K. Harshina, and M. Schranz. Reinforcement learning based countermeasures against intelligent attacking UAV swarms. Unpublished. To be published in 2023/2024 as part of the proceedings for Wi-DroIT 2023., 2023.

[70] J. Simonjan, S. R. Probst, and M. Schranz. Inducing defenders to mislead an attacking UAV swarm. In *Proceedings of the IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 278–283. IEEE, 2022.

[71] R. Smith. The long history of gaming in military training. *Simulation & Gaming*, 41(1):6–19, 2010.

[72] B. D. Song, J. Kim, and J. R. Morrison. Rolling horizon path planning of an autonomous system of UAVs for persistent cooperative service: Milp formulation and efficient heuristics. *Journal of Intelligent & Robotic Systems*, 84:241–258, 2016.

[73] R. J. Stone, P. B. Panfilov, and V. E. Shukshunov. Evolution of aerospace simulation: From immersive virtual reality to serious games. In *Proceedings of 5th International Conference on Recent Advances in Space Technologies - RAST2011*, pages 655–662, 2011.

[74] T. Susi, M. Johannesson, and P. Backlund. *Serious games: An overview.* Institutionen för kommunikation och information, 2007.

[75] J. Tang, H. Duan, and S. Lao. Swarm intelligence algorithms for multiple unmanned aerial vehicles collaboration: A comprehensive review. *Artificial Intelligence Review*, 56(5):4295–4327, 2023.

[76] K. S. Tekinbas and E. Zimmerman. *Rules of play: Game design fundamentals.* MIT press, 2003.

[77] S. Utubor. *Improving Detection of Attacks in Cyber-Physical Systems: Applying Gradient Boosting Based Machine Learning Techniques.* PhD thesis, The George Washington University, 2023.

[78] Valve. *Team Fortress 2.* Valve, 2007. `https://www.teamfortress.com/`. Accessed: 01.09.2023.

[79] Valve. *Counter-Strike: Global Offensive.* Valve, 2012. `https://store.steampowered.com/app/730/CounterStrike_Global_Offensive/`. Accessed: 01.09.2023.

[80] B. R. Van Voorst. Counter drone system, Sept. 14 2017. US Patent App. 15/443,143.

[81] B. Wang, S. Li, X. Gao, and T. Xie. UAV swarm confrontation using hierarchical multiagent reinforcement learning. *International Journal of Aerospace Engineering*, 2021.

[82] S. Wang, J. Chen, Z. Zhang, G. Wang, Y. Tan, and Y. Zheng. Construction of a virtual reality platform for UAV deep learning. In *2017 Chinese Automation Congress (CAC)*, pages 3912–3916. IEEE, 2017.

[83] L. Xiang and T. Xie. Research on UAV swarm confrontation task based on MADDPG algorithm. In *Proceedings of the 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pages 1513–1518. IEEE, 2020.

[84] Q. Yang, Y. Zhu, J. Zhang, S. Qiao, and J. Liu. UAV air combat autonomous maneuver decision based on ddpg algorithm. In *Proceedings of the IEEE 15th international conference on control and automation (ICCA)*, pages 37–42. IEEE, 2019.

[85] J. Yong-kang, C. Yong, and T. Daquan. Design of an UAV simulation training and assessment system based on unity3d. In *2017 IEEE International Conference on Unmanned Systems (ICUS)*, pages 163–167. IEEE, 2017.

[86] C. Yu, A. Velu, E. Vinitsky, J. Gao, Y. Wang, A. Bayen, and Y. Wu. The surprising effectiveness of ppo in cooperative multi-agent games. *Advances in Neural Information Processing Systems*, 35:24611–24624, 2022.

[87] G. Zhang, Y. Li, X. Xu, and H. Dai. Efficient training techniques for multi-agent reinforcement learning in combat tasks. *IEEE Access*, 7:109301–109310, 2019.

[88] T. Zhang, Q. Li, C.-s. Zhang, H.-w. Liang, P. Li, T.-m. Wang, S. Li, Y.-l. Zhu, and C. Wu. Current trends in the development of intelligent unmanned autonomous systems. *Frontiers of information technology & electronic engineering*, 18:68–85, 2017.

[89] C. Zhao, M. Shi, Z. Cai, and C. Chen. Detection of unmanned aerial vehicle signal based on gaussian mixture model. In *2017 12th International Conference on Computer Science and Education (ICCSE)*, pages 289–293. IEEE, 2017.

[90] M. Zhao, J. Zheng, and E. S. Liu. Server allocation for massively multiplayer online cloud games using evolutionary optimization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 17(2):1–23, 2021.

[91] S. Zhou, M. J. Phielipp, J. A. Sefair, S. I. Walker, and H. B. Amor. Clone swarms: Learning to predict and control multi-robot systems by imitation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4092–4099. IEEE, 2019.

[92] Y. Zhou, B. Rao, and W. Wang. UAV swarm intelligence: Recent advances and future trends. *IEEE Access*, 8:183856–183878, 2020.

Appendices

Appendix A

Open Source Code in GitHub Repository

The code from this project can be found in the following repository: `https://github.com/KseniaGa/MasterThesis`.