Uncooperative Objects Effective Imaging through Flexible Flyaround Guidance

ESA’s 2021 Clean Space Industrial Days

20-24 September 2021

Andrea Brandonisio*, Michèle Lavagna

DIPARTIMENTO DI SCIENZE E TECNOLOGIE AEROSPAZIALI

***Speaker**

Outline

- Introduction
- Problem Architecture
- Problem Dynamics
- Decision Process Algorithm
- Main Test Campaign
- Conclusions

Outline

Introduction

→ Spacecraft **planning fly-around paths** via Deep Reinforcement Learning (**DRL**) fed by information derived from a set of images acquired during the ***close-proximity operations***.

→ optimize the **shape reconstruction** of an unknown and **uncooperative space object** through ***image processing***.

Imaging based sensor suite on board the chaser is the baseline for target data acquisition and drives the proper fly-around guidance synthesis in the mission design phase.

The problem fits into the broad domain of technologies development for on-orbit servicing (**OOS**) related **autonomous navigation**

State of the Art and Research Objectives

This work has been developed from our previous formulation of **A2C algorithm** to solve spacecraft path-planning for uncooperative and unknown space object mapping [1] with **linear** neural networks.

<u>**State of the Art**</u> about DRL applied on spacecraft path-planning:

- **Planetary landing** trajectory [2].
- **Planetary relative** autonomous navigation [3].
- **Small bodies** exploration [4], [5].

1. Analysis of <u>**Proximal Policy Optimization (PPO) performance**</u>.
2. Prove the feasibility of using <u>**Recurrent Networks**</u>.
3. Increase <u>**generalization**</u> capabilities to solve the problem for <u>**random initial conditions**</u>.

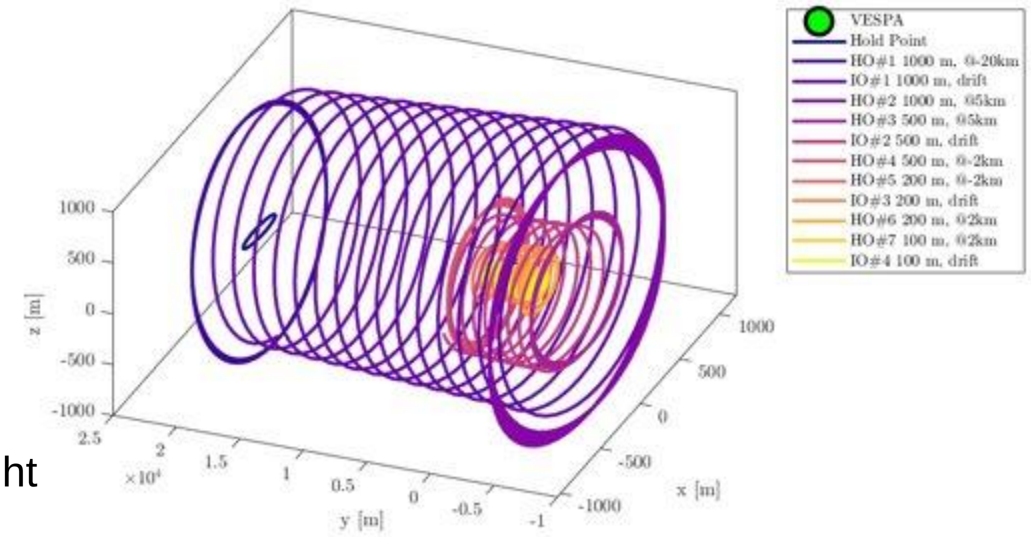
[1] Andrea Brandonisio, Michèle Lavagna and Davide Guzzetti. *“Deep Reinforcement Learning to Enhance Fly-aroundGuidance for Uncooperative Space Objects Smart Imaging”*. In: AAS/AIAA Astrodynamic Specialist Conference, Aug 2020.
[2] Brian Gaudet, Richard Linares, and Roberto Furfaro. *“Adaptive Guidance and Integrated Navigation with Reinforcement Meta-Learning”*. In: *Acta Astronautica 169*, 2020.
[3] Richard Linares et al. *“A Deep Learning Approach for Optical Autonomous Planetary Relative Terrain Navigation*”, 2017.
[4] David M. Chan and Ali-akbar Agha-mohammadi. *“Autonomous Imaging and Mapping of Small Bodies Using Deep Reinforcement Learning”*. In: *IEEE Aerospace Conference*, 2019.
[5] Margherita Piccinin and Michèle Lavagna. *“Deep Reinforcement Learning approach for Small Bodies Shape Reconstruction Enhancement”*. In: *AIAA Scitech 2020 Forum*, 2020.

Research Scenario: e.INSPECTOR*

<u>eINSPECTOR</u> is a **12U CubeSat** with two optical instruments:

- Visible camera
- Infrared camera

The **objective** of the mission is the VESPA **<u>image acquisition</u>** and the flight demonstration of new algorithms of **<u>relative navigation</u>** through image processing.



DIPARTIMENTO DI SCIENZE E TECNOLOGIE AEROSPAZIALI

* In development at the Department of Aerospace Science and Technologies of the Politecnico di Milano.

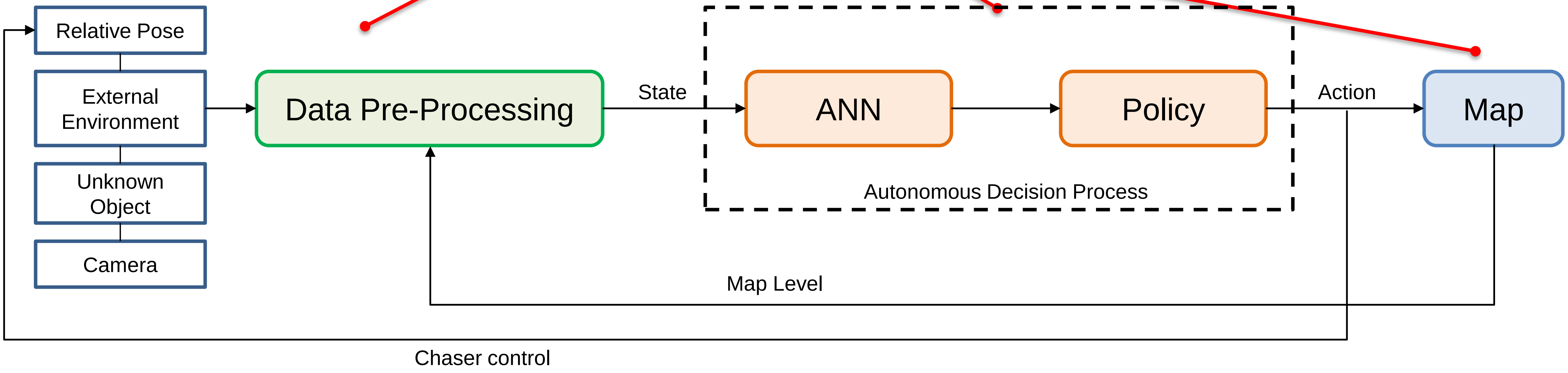
Outline

- Introduction
- Problem Architecture
- Problem Dynamics
- Decision Process Algorithm
- Main Test Campaign
- Conclusions

Problem Architecture

Three main parts:

1. **Input pre-processing**: environment information → feed the autonomous decision process.
2. **Autonomous decision process**: the DRL algorithm → spacecraft action.
3. **Output**: images taken by the spacecraft camera → target object overall map.

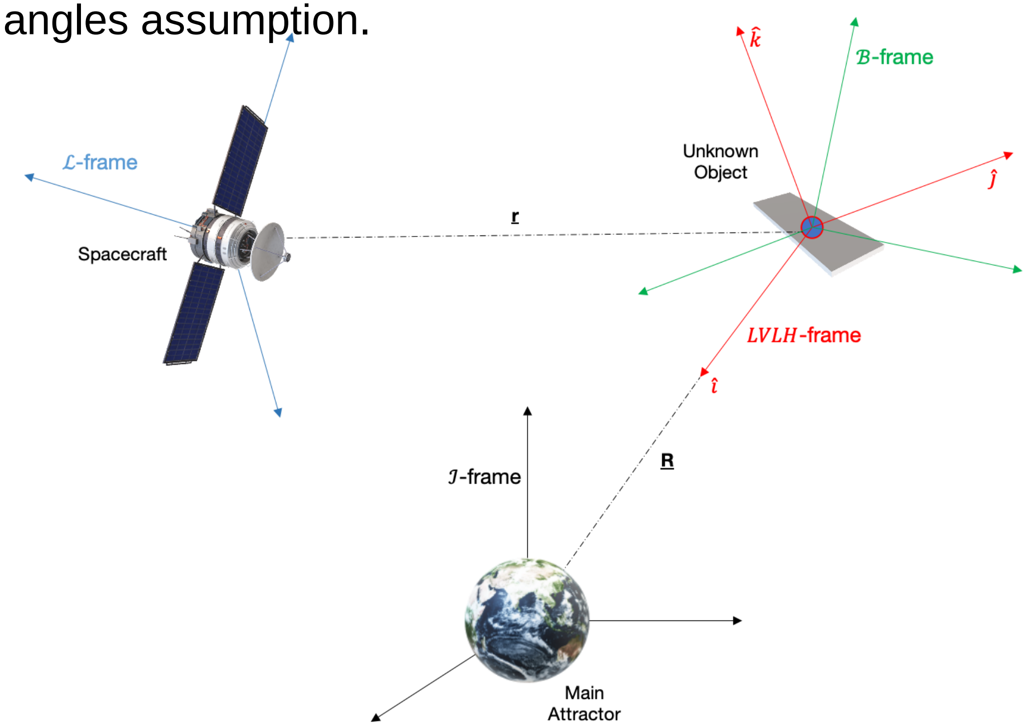


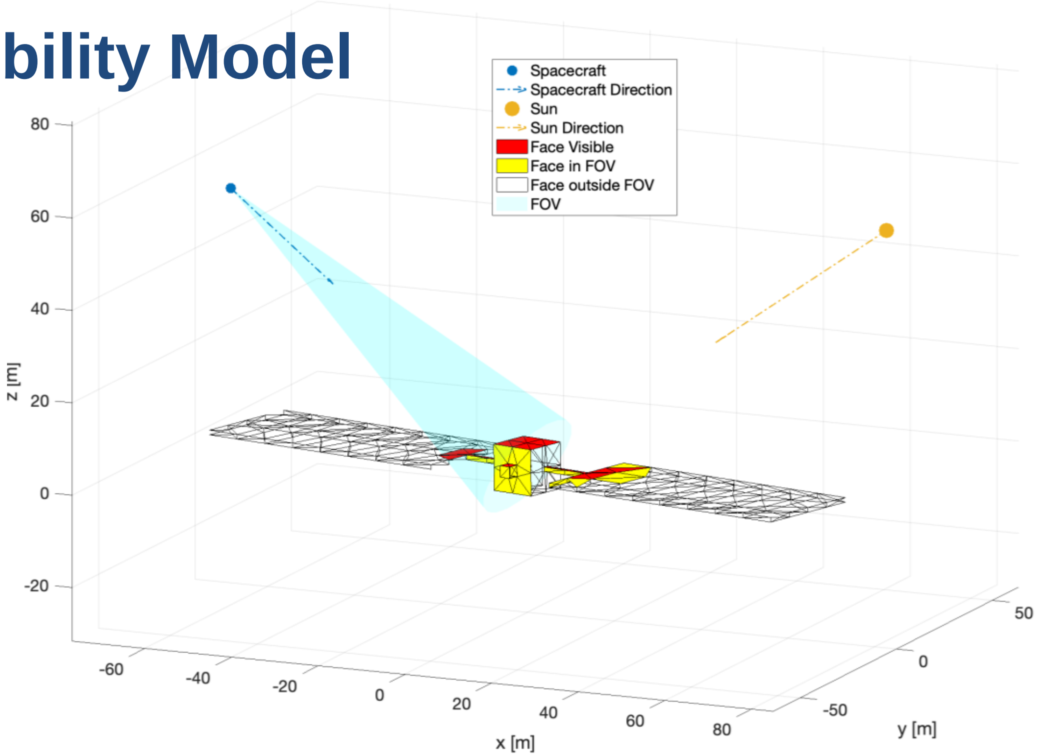
Outline

Translational-Rotational Dynamics and Visibility Model

Spacecraft-Object Relative dynamics:

- **Translational**: Eccentric Linearized 2° Order Differential Equations. Expressed in LVLH frame centred in target object c.m.
- **Rotational**: Euler's equations in LVLH frame, under small angles assumption.





Target Object: triangular mesh of polyhedron.
Strong assumption: camera points towards target center.

The **visibility** depends mainly on:

- The camera field of view (**FOV**)
- **Camera** incidence angle
- **Sun** illumination condition – incidence angle.

Outline

- Introduction
- Problem Architecture
- Problem Dynamics
- Decision Process Algorithm
- Main Test Campaign
- Conclusions

Autonomous Decision Process

Autonomous Navigation Problem

Formulated as an active **SLAM problem** (active Simultaneous Localization and Mapping problem)

Solved with Partially Observable Markov Decision Process (**POMDP**) characterized by:

- **State** space and **Action** space
- **Reward** space
- **Transition Probability**

objective

1

Deep Reinforcement Learning algorithm:
Proximal Policy Optimization

DRL Models (1)

DRL-based algorithm formulation **three main models**:
- State Space model
- Action Space model
- Reward model

State Model

12 elements: relative translational position and velocity, relative rotational position and velocity.

$$\mathcal{S} = \begin{Bmatrix} \underline{x} \\ \underline{\dot{x}} \\ \underline{\alpha} \\ \underline{\dot{\alpha}} \end{Bmatrix}$$

$\mathcal{S}$ $\longrightarrow$ **stereovision-based** methods.

Action Model

7 actions: **thruster** fired along one of the six cartesian directions or null action.

$$\mathcal{A} = \{T_{x_+}, T_{x_-}, T_{y_+}, T_{y_-}, T_{z_+}, T_{z_-}, 0\}$$

Action $\longrightarrow$ translational relative dynamics equations

$$\begin{cases} \ddot{x} = \ddot{x}_{tr} + a_{x_+}\delta_{x_+} - a_{x_-}\delta_{x_-} \\ \ddot{y} = \ddot{y}_{tr} + a_{y_+}\delta_{y_+} - a_{y_-}\delta_{y_-} \\ \ddot{z} = \ddot{z}_{tr} + a_{z_+}\delta_{z_+} - a_{z_-}\delta_{z_-} \end{cases}$$

DRL Models (2)

Tasks requested to the agent:

1) **surviving** around the object;
2) taking good **quality images**;
3) obtaining the **best mapping** level;
4) **reducing** the number of control actions – **firings**.

Reward Model Scores:

- **Position score** R_d → Task 1
- **Time of flight** score R_t → Task 1
- **Sun incidence** score $R_\eta = \frac{1}{n}\sum_{j=0}^{n} r_{\eta,j}$ → Task 2
- **Camera incidence** score $R_\varepsilon = \frac{1}{n}\sum_{j=0}^{n} r_{\varepsilon,j}$ → Task 2
- **Map Level** score R_m → Task 3
- **Thrust Level** score R_f. → Task 4

Reward Model $R_{k,1}$	Reward Model $R_{k,2}$
$R_k = R_\eta + R_\varepsilon + R_d + R_t + R_m$	$R_k = R_\eta + R_\varepsilon + R_d + R_t + R_m + R_f$

ANN Configurations

Policy and Value Networks Linear Case

	Policy Network		Value Network	
Layer	Elements	Activation	Elements	Activation
1^{st} Hidden Layer	10*dim_obs	tanh	10*dim_obs	tanh
2^{nd} Hidden Layer	$\sqrt{n_{h1} * n_{n3}}$	tanh	$\sqrt{n_{h1} * n_{n3}}$	tanh
3^{rd} Hidden Layer	10*dim_act	Leaky-ReLU	10*dim_act	Leaky-ReLU
output	dim_act	softmax	dim_act	linear

objective

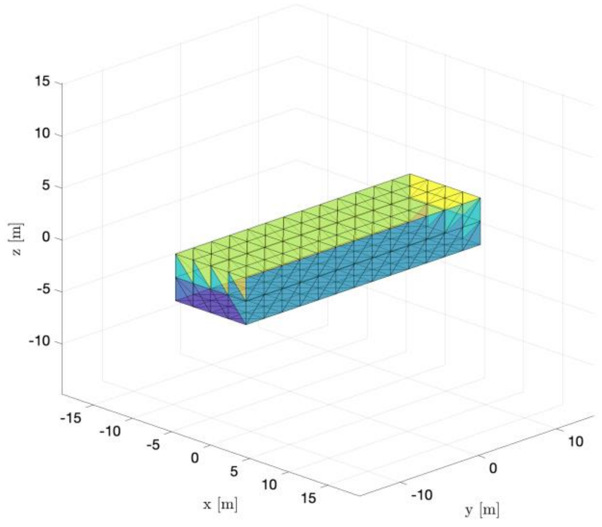
2

Policy and Value Networks Recurrent Case

	Policy Network		Value Network	
Layer	Elements	Activation	Elements	Activation
LSTM Layer	24	-	24	-
1^{st} Hidden Layer	64	ReLU	64	ReLU
2^{nd} Hidden Layer	32	ReLU	32	ReLU
output	dim_act	softmax	dim_act	linear

Outline

Main Test Campaign



- Simple **<u>parallelepiped</u>**: 344 faces.
- **<u>FOV</u>**: 10°
- **<u>Control Interval</u>**: 30s
- 25 correct **<u>photos</u>** / face

objective

3

<u>Reward Models</u>:

a. $\boldsymbol{R_{k,1}}$ → without thrust level score

b. $\boldsymbol{R_{k,2}}$ → with thrust level score

<u>Random Initial Conditions</u>:

a. RDM-A case:

→ Rotational position and velocity

→ Sun phase

b. RDM-B case:

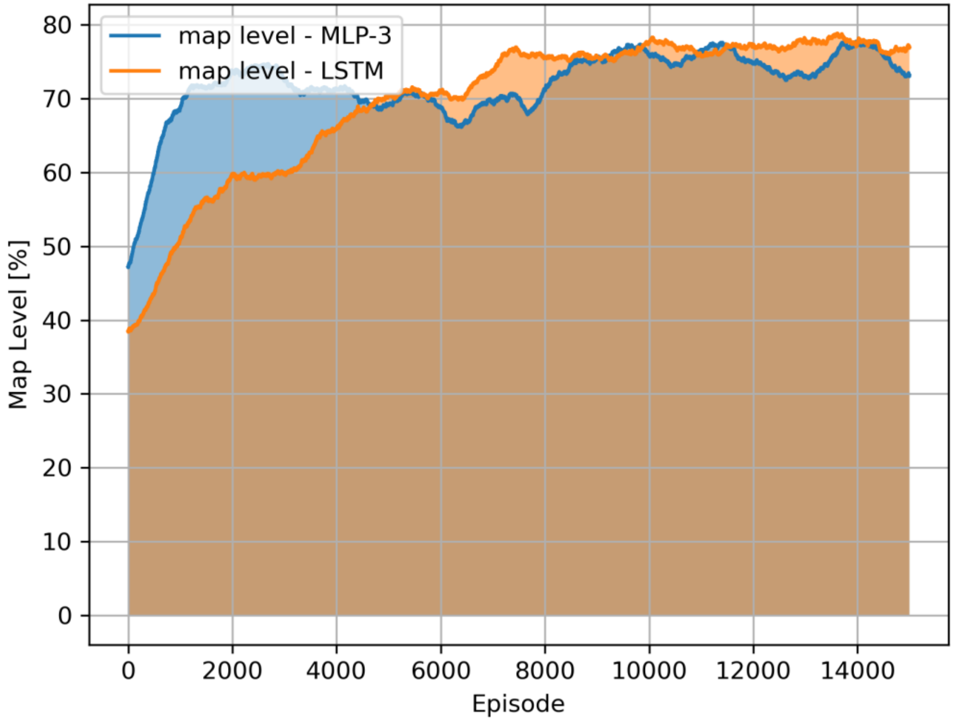
→ Relative rotational position and velocity

→ Sun phase

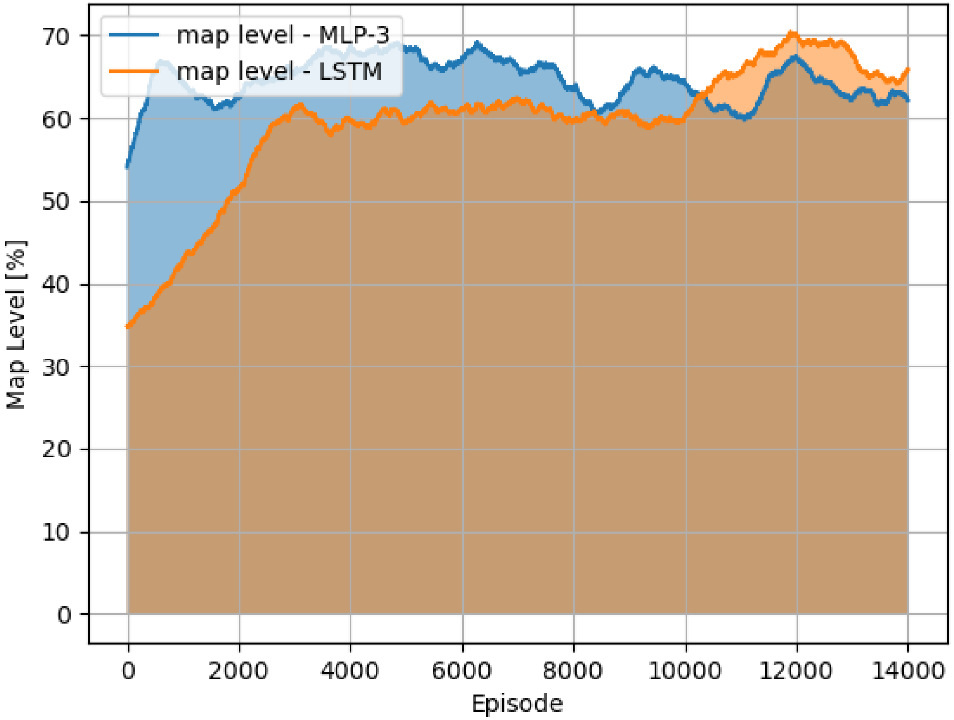
→ Relative position

Test Results: RDM-A

Reward Model $\boldsymbol{R_{k,1}}$ w/o thrust level score
+
Random Case A

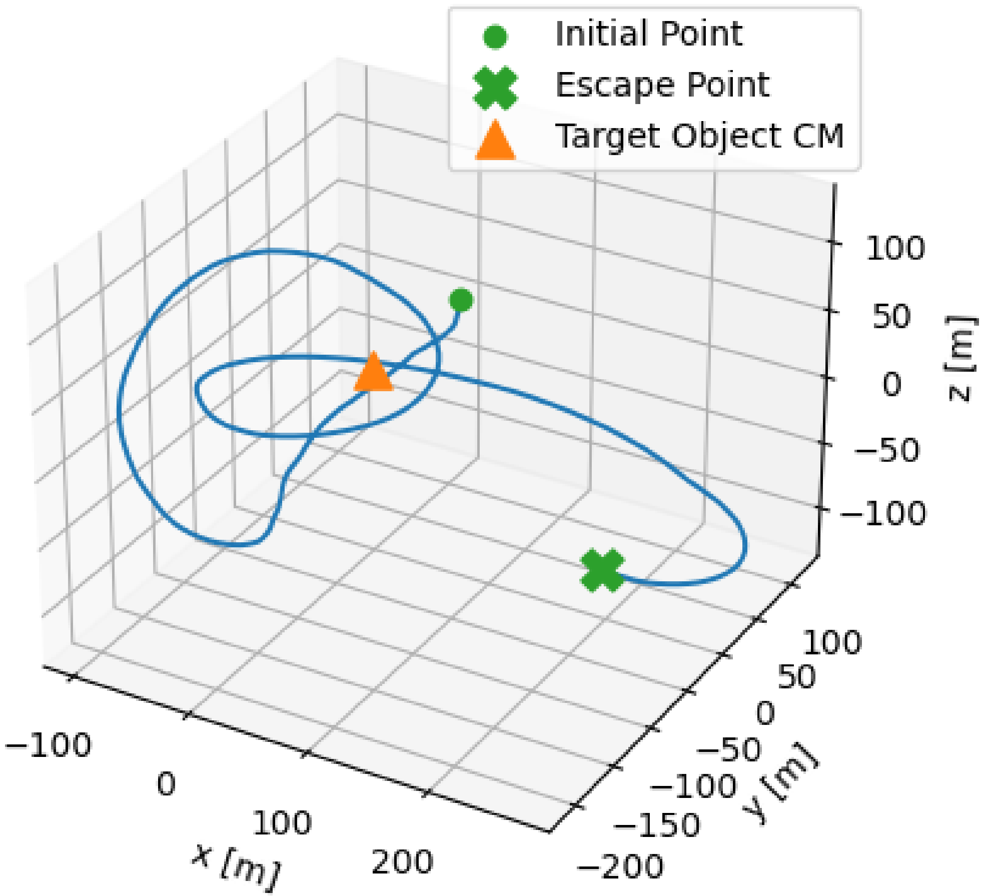


Reward Model $\boldsymbol{R_{k,2}}$ with thrust level score
+
Random Case A



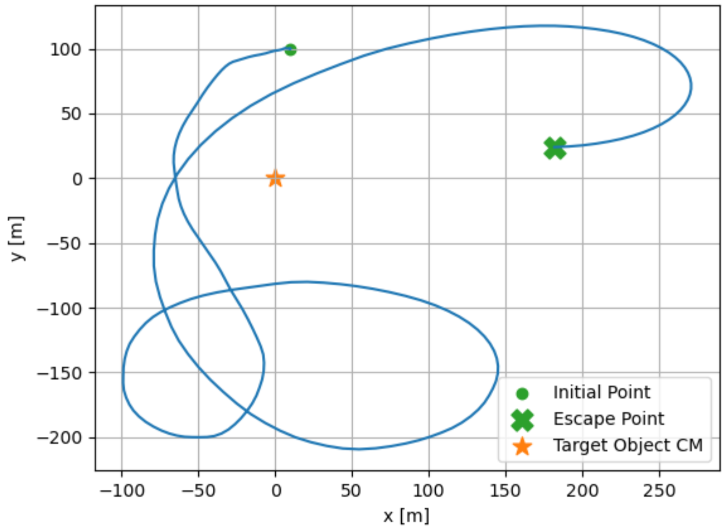
- MLP-3 **faster** than LSTM.
- LSTM more **stable.**
- Average **map level** of around **70%-80%.**

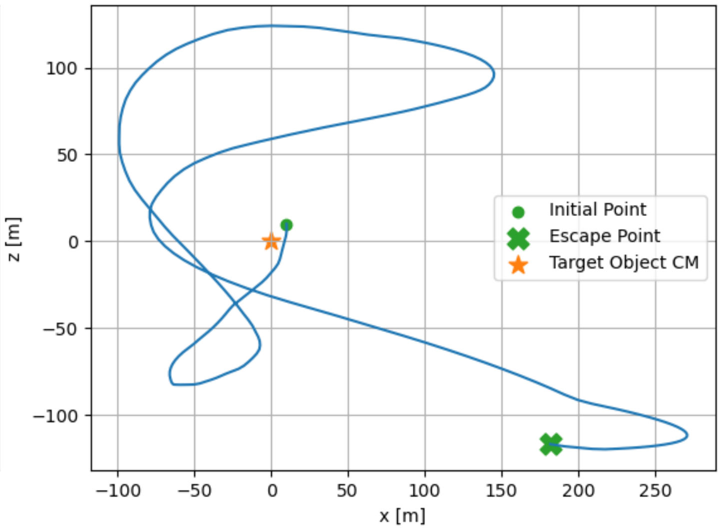
Test Results: RDM-A – Trajectory



Recurrent (LSTM) Policy Agent:

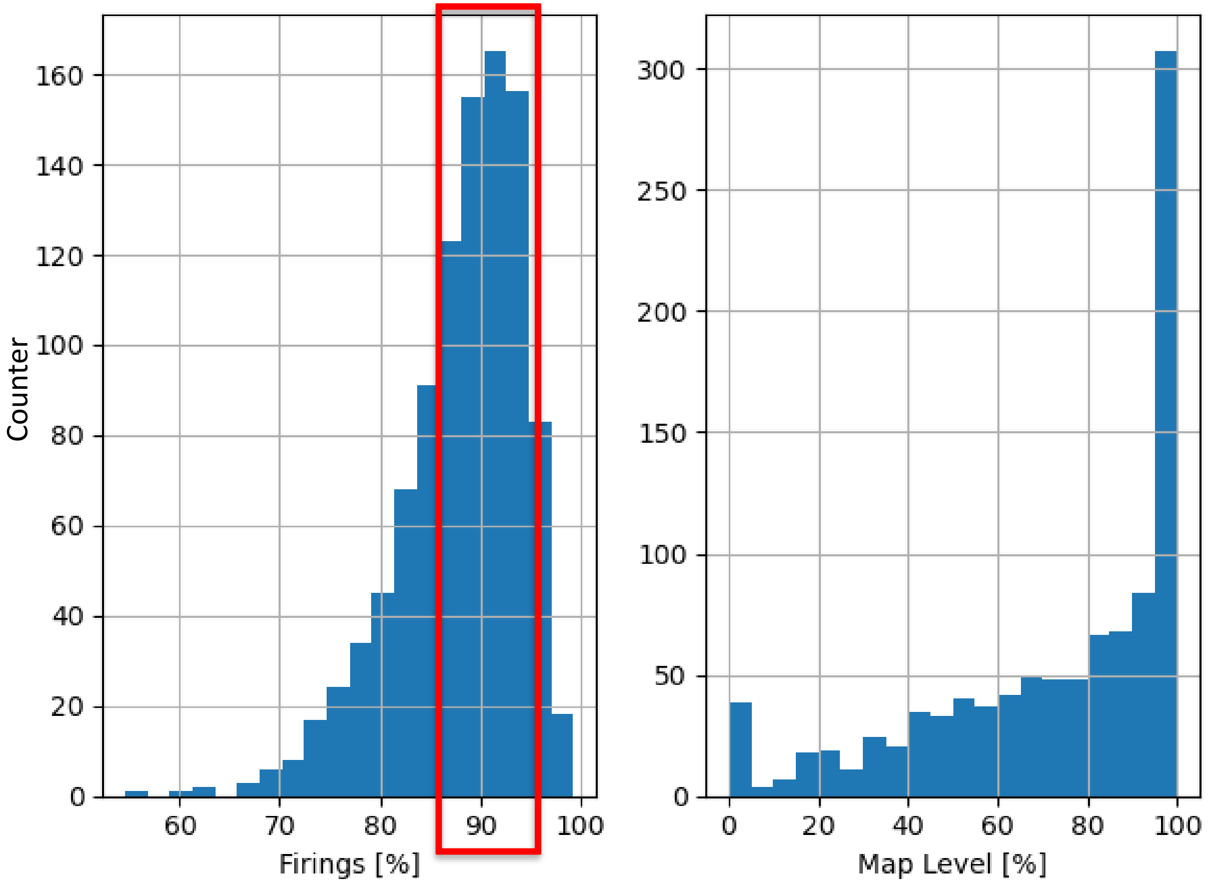
Initial Condition [m]	Map Level [%]	Time [h]	Reward Model
[10, 100, 10]	100	3.67	$R_{k,1}$



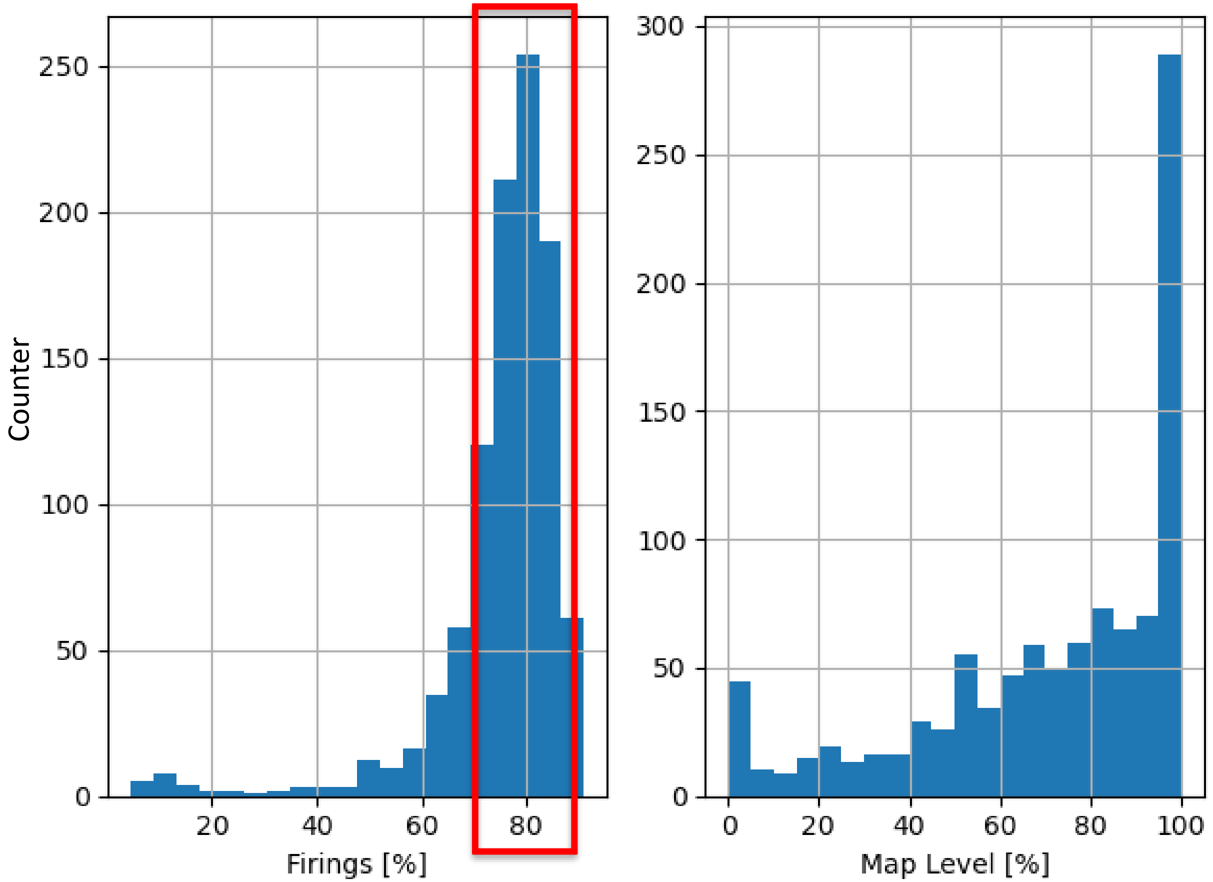


Test Results: RDM-A – Thrust Level Analysis

MLP-3 with **Reward Model $R_{k,1}$**



MLP-3 with **Reward Model $R_{k,2}$**

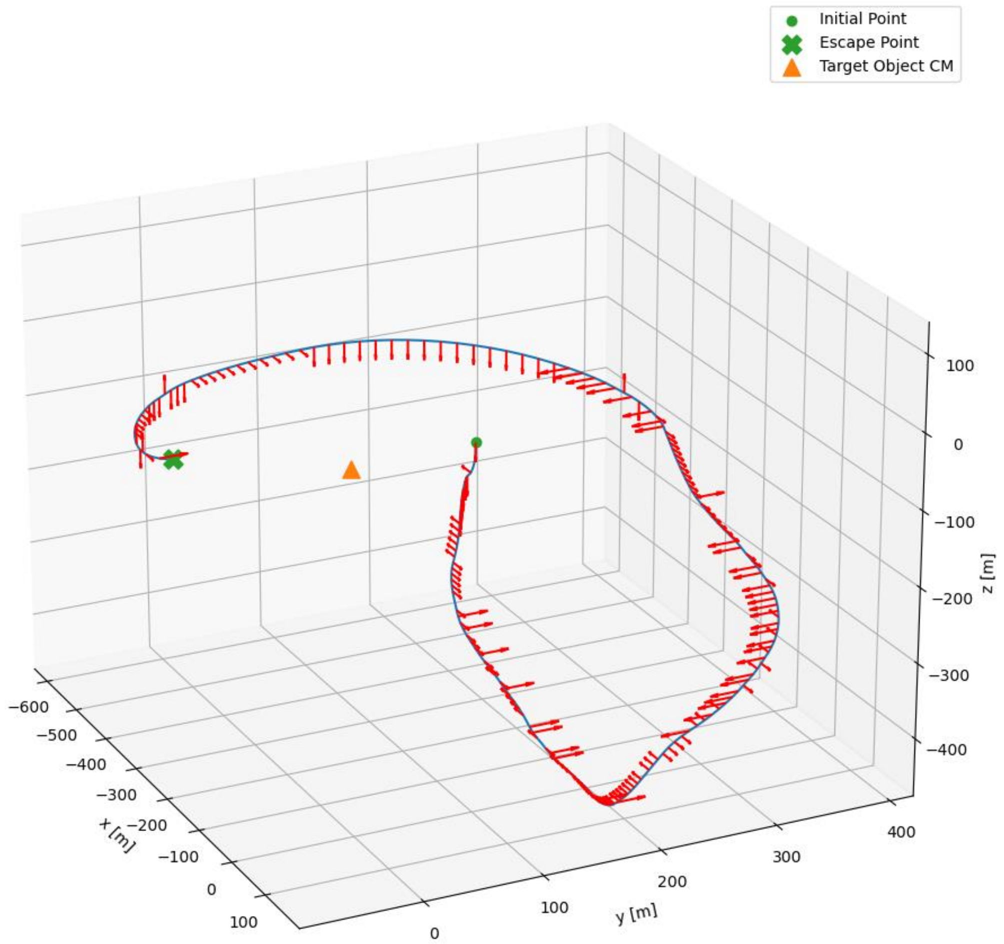


Test Results: RDM-A – Action Control Profile

Linear (MLP-3) Policy Agent:

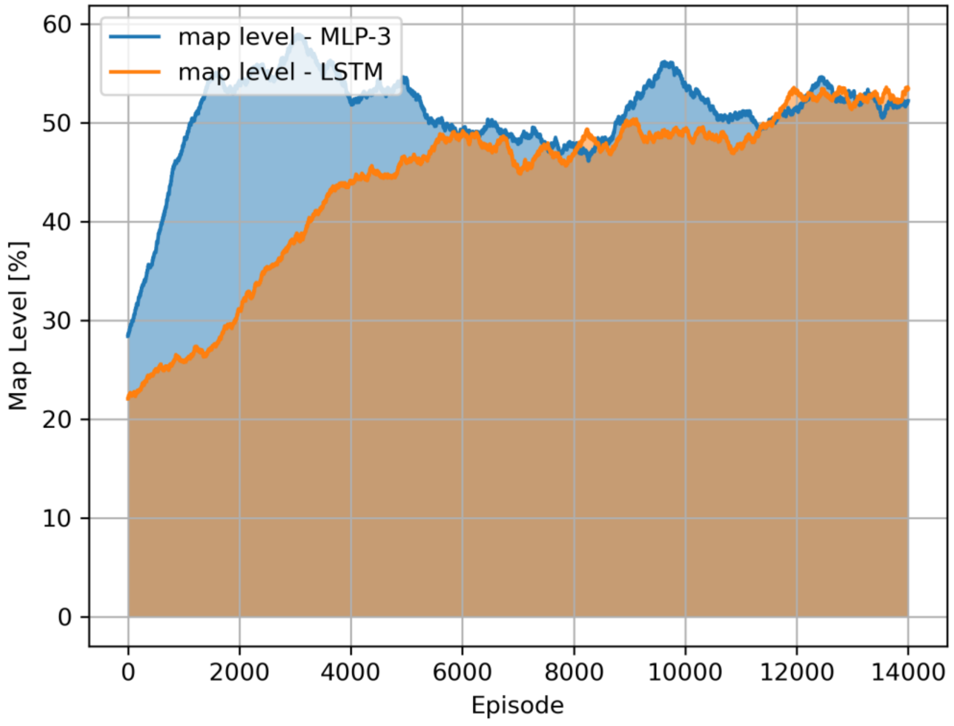
Initial Condition [m]	Map Level [%]	Time [h]
[10, 100, 10]	94	2.1

Reward Model	Firings [%]
$R_{k,2}$	75.4

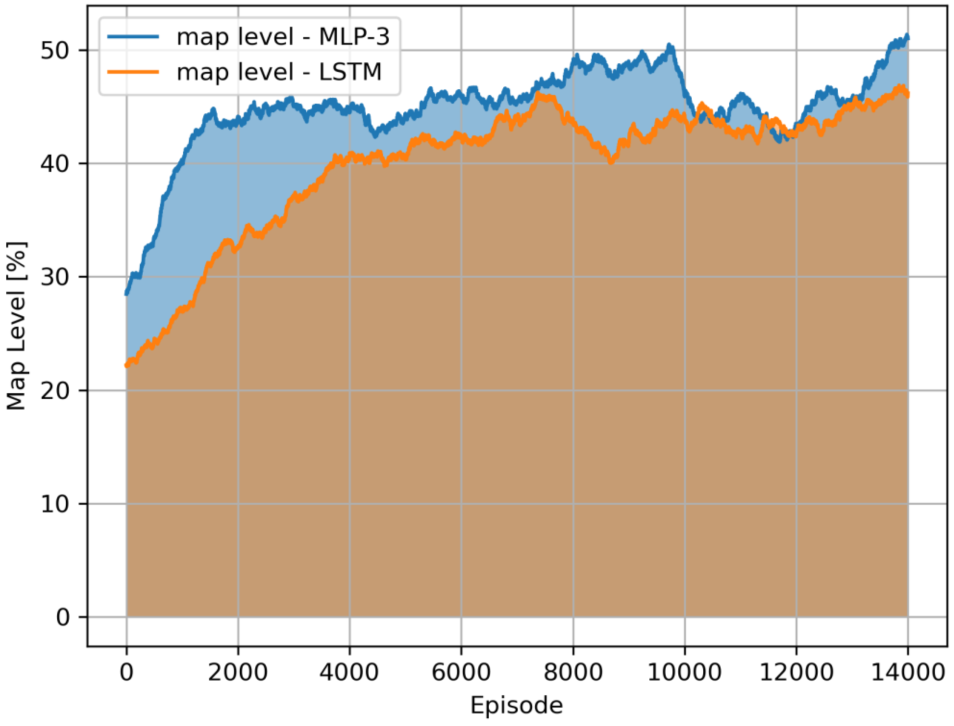


Test Results: RDM-B

Reward Model $R_{k,1}$ w/o thrust level score
+
Random Case B



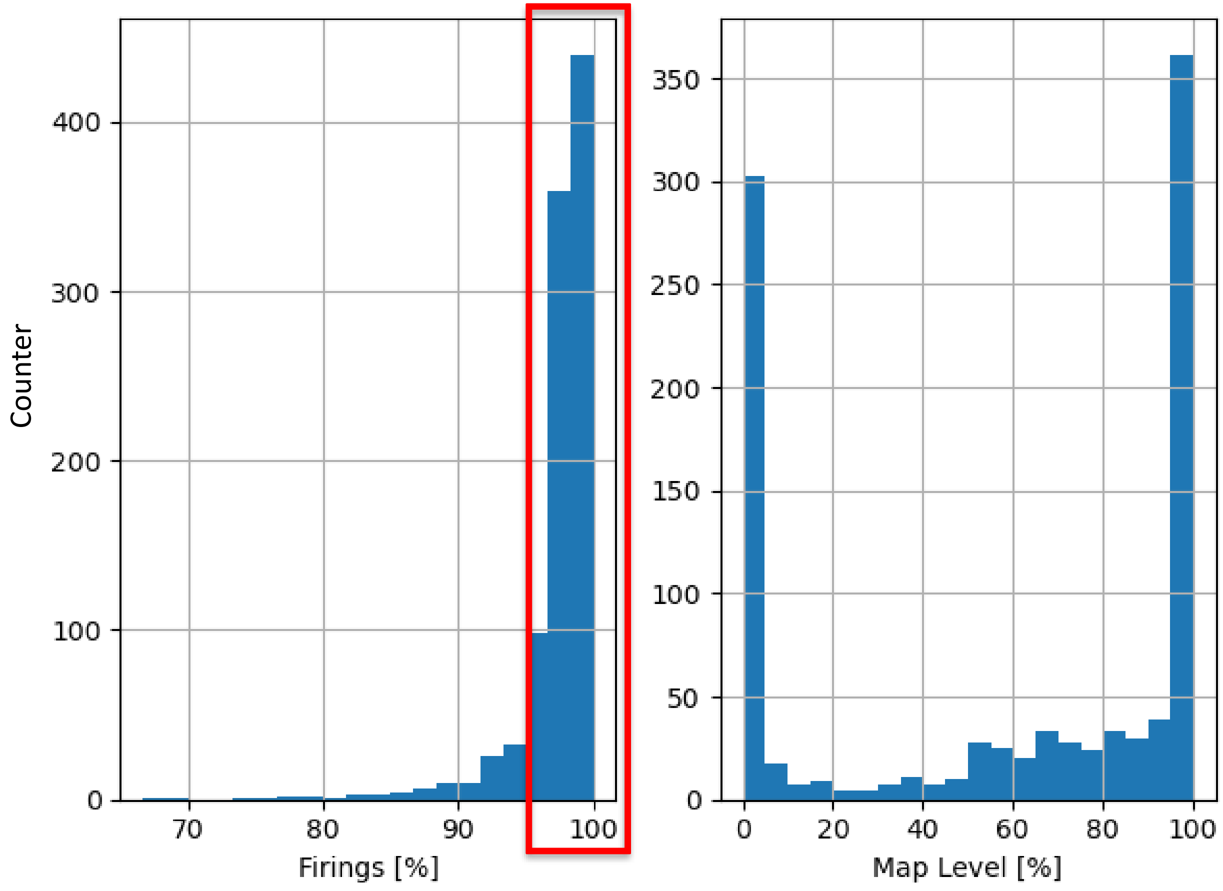
Reward Model $R_{k,2}$ with thrust level score
+
Random Case B



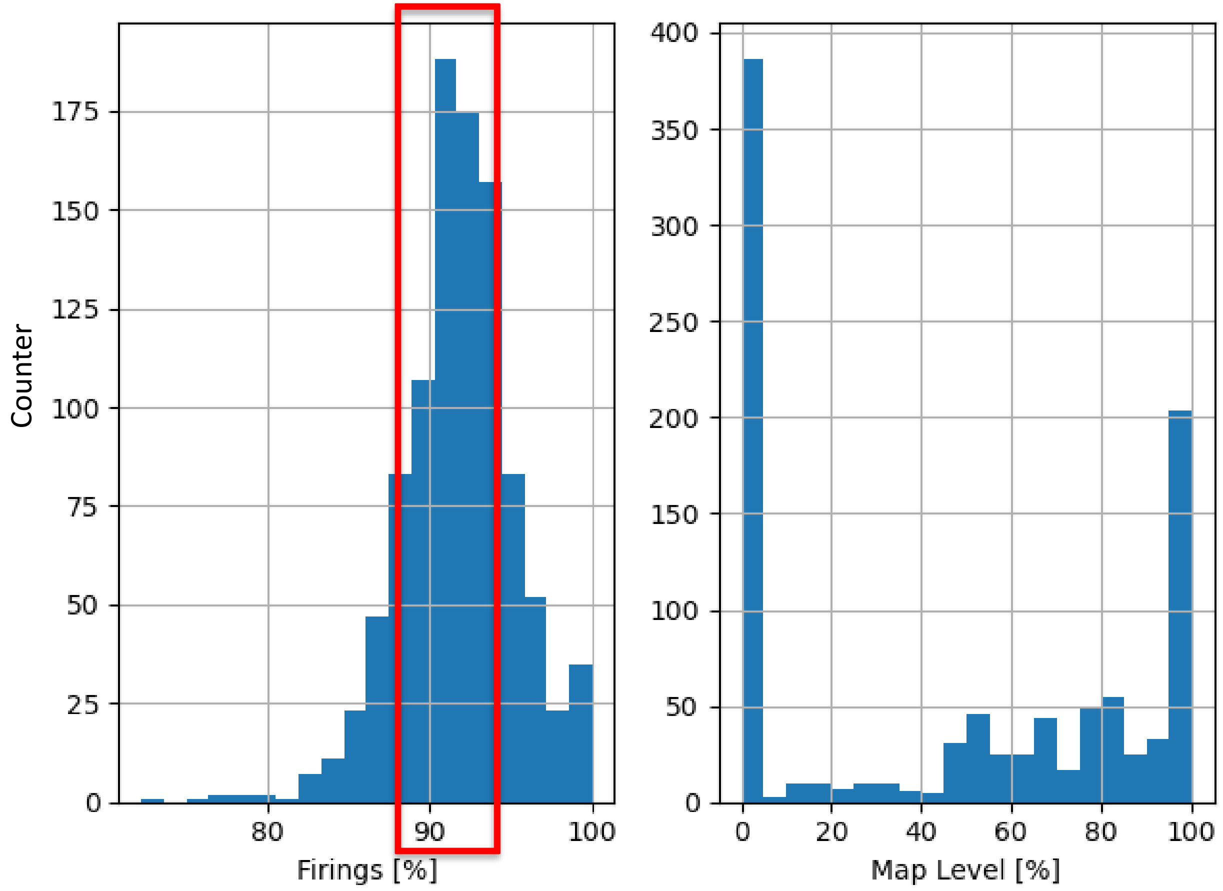
- MLP-3 **faster** than LSTM.
- LSTM more **stable.**
- Average **map level** of around **45%-50%.**

Test Results: RDM-B – Thrust Level Analysis

LSTM with **Reward Model** $R_{k,1}$

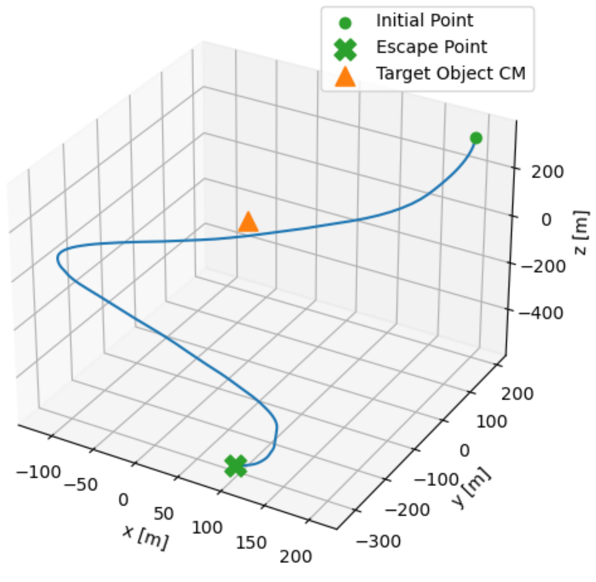


LSTM with **Reward Model** $R_{k,2}$

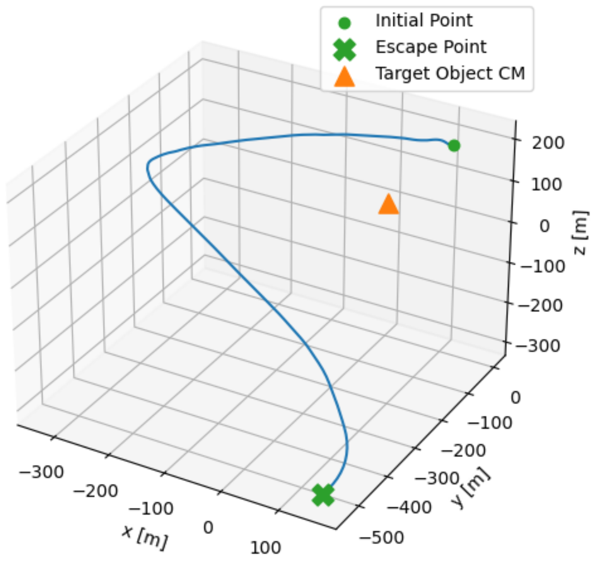


Test Results: RDM-B - Trajectories

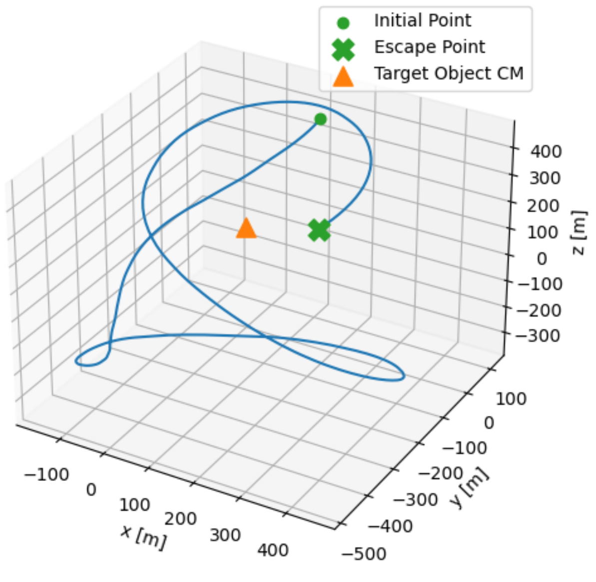
Recurrent (LSTM) Policy Agent:



Map Level [%]	Time [h]
50.7	1.06



Map Level [%]	Time [h]
84.3	1.27



Map Level [%]	Time [h]
76.1	2.21

Outline

- Introduction
- Problem Architecture
- Problem Dynamics
- Decision Process Algorithm
- Main Test Campaign
- Conclusions

Conclusions

<u>DRL</u> is confirmed applicable to autonomous **<u>path-planning</u>** for this OOS scenario both with ***linear*** and ***recurrent*** policy.

<u>PPO</u> is the most valid algorithm to solve this problem with **<u>randomized initial condition</u>**. (PPO overcomes A2C.)

<u>Future Developments</u>

- Spacecraft attitude dynamics 6DoF control
- Control interval and action modulation
- Continuous control
- State vector uncertainties

Increasing complexity of the **State** and **Action** spaces

Increasing robusteness of the ANN and DRL algorithm

Thank you for the attention.
Any questions?

POLITECNICO
MILANO 1863

Uncooperative Objects Effective Imaging through Flexible Flyaround Guidance

ESA's 2021 Clean Space Industrial Days
20-24 September 2021

Andrea Brandonisio, Michèle Lavagna

DIPARTIMENTO DI SCIENZE E TECNOLOGIE AEROSPAZIALI