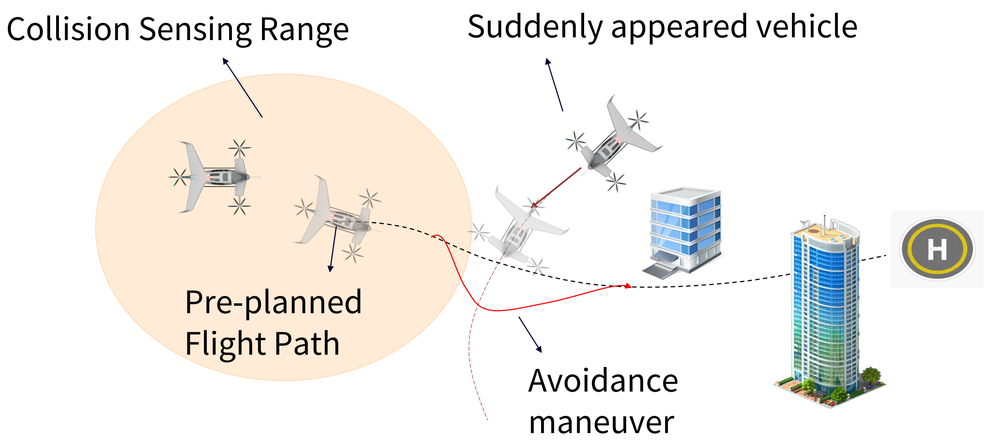
Federated Meta Reinforcement Learning for UAV Navigation in Urban Airspace

Lancaster University & Cranfield University

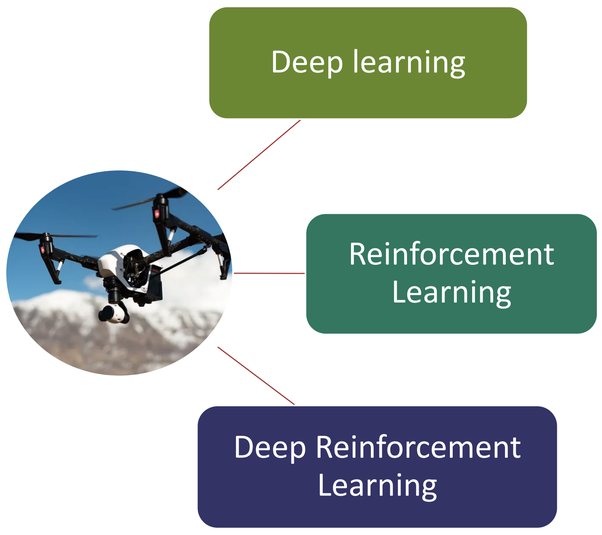
Research Fellows: Dr. Zhengxin Yu, Dr. Burak Yuksek
Investigators: Prof. Neeraj Suri, Prof. Gokhan Inalhan

Background: Autonomous Navigation

Autonomous navigation is a fundamental problem of mobile robots, which aims to o identify an optimal or suboptimal path from a starting point to a target point in a Two-Dimensional (2D) or Three-Dimensional (3D) environment while avoiding obstacles.

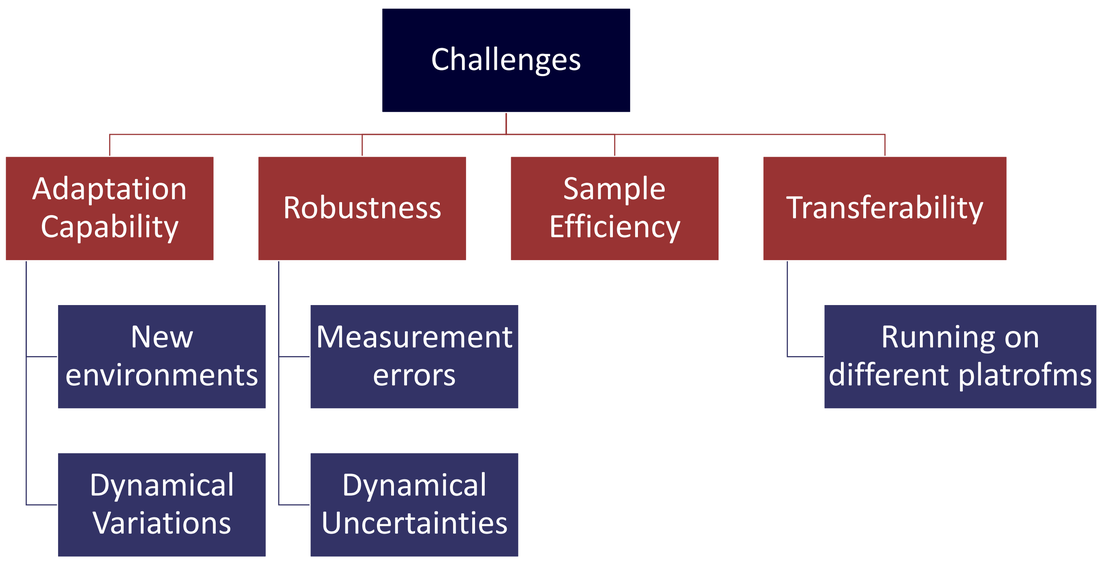


Learning-Based Navigation Methods: State-of-the-art



- **Pros:**
  - They can provide scalable solutions for large operation environments.
  - Require little manual-engineering.
- **Cons:**
  - Weak adaptability to new environment
  - Low sample efficiency

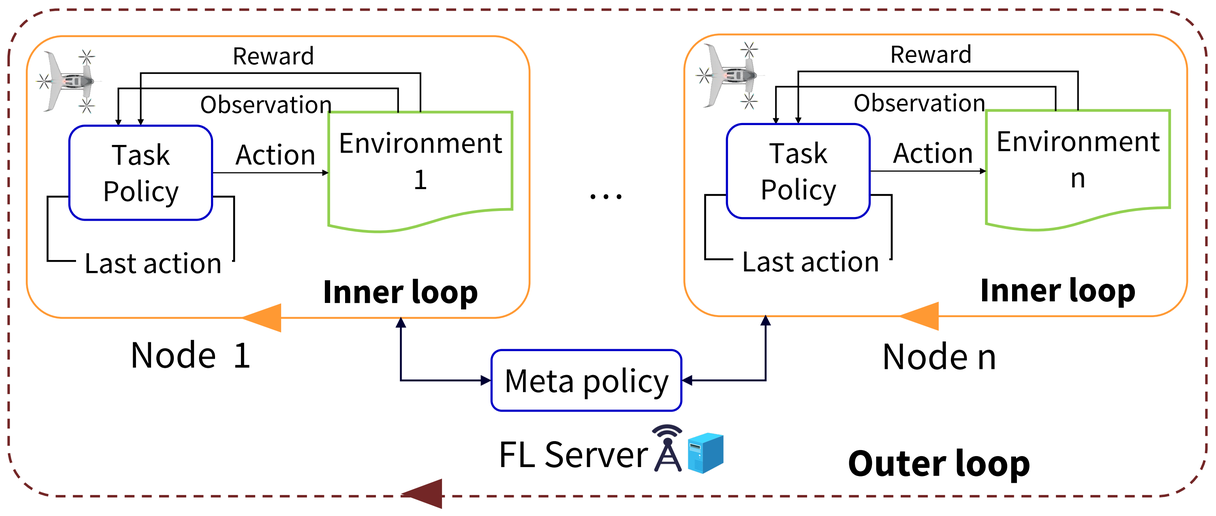
Challenges in AI-Based UAV Navigation



Design Goals

- Model generalizability across heterogeneous vehicles
- Fast adaptation to dynamic environments (Short training time for new environments)
- Lifelong learning ability
- Flight/operation safety during learning process

Federated Meta Reinforcement Learning



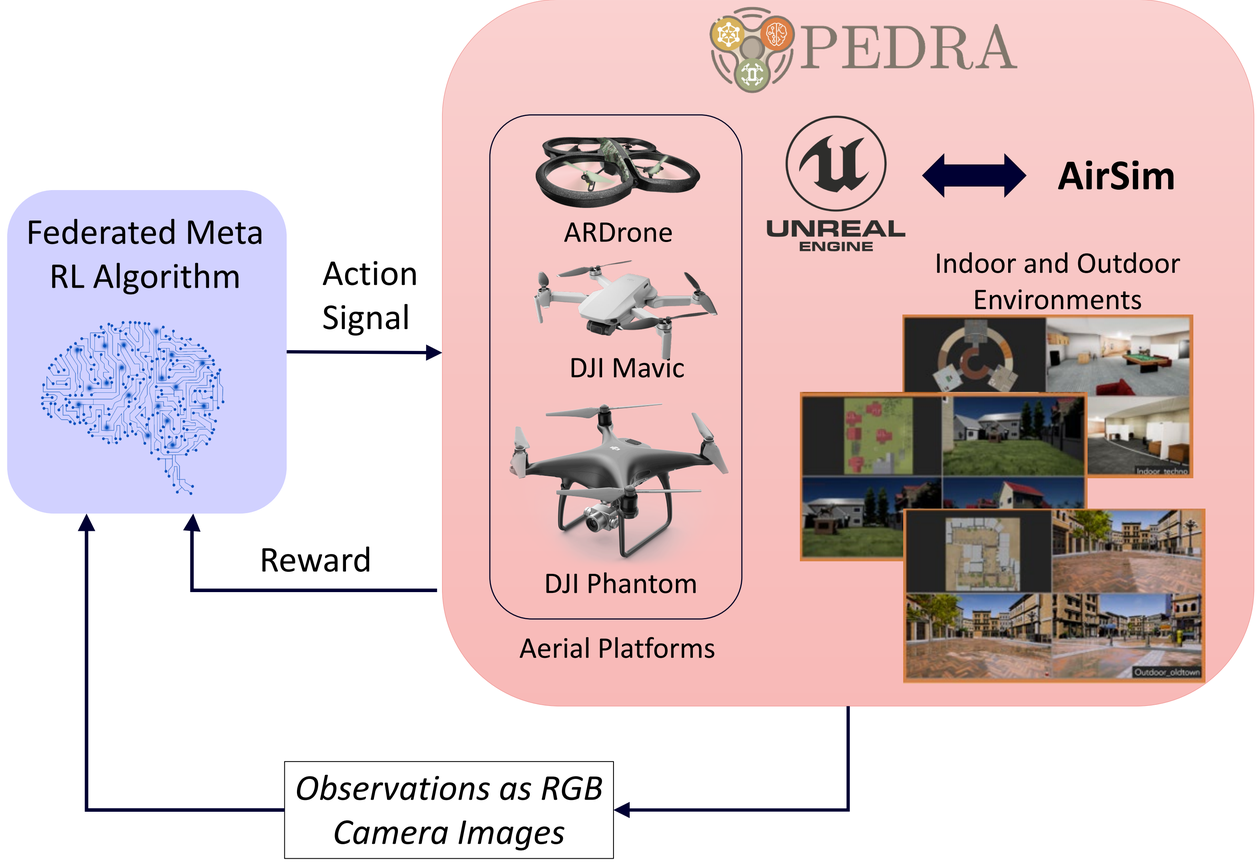
- Let the agent learn how to learn new tasks faster by reusing previous experience.
- Involve two learning loops of training:
  - ***Out loop*** learns common knowledge (represented by a neural network θ) from many tasks.
  - ***Inner loop*** learns policies based on the learned model θ.

Main differences from reinforcement learning:

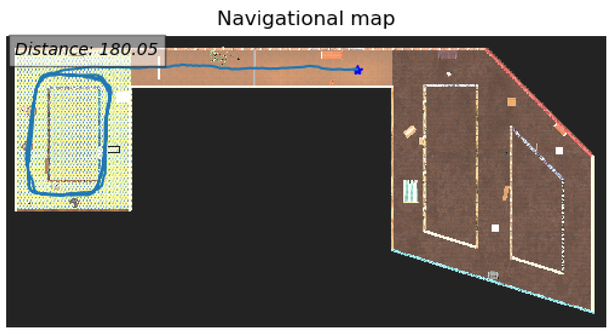
- In meta-RL, the current state s_t, last reward r_{t-1} and the last action a_{t-1} are all incorporated into the policy observation → $\pi_\theta(a_{t-1}, r_{t-1}, s_t)$.
- In RL, only the current state s_t is considered → $\pi_\theta(s_t)$.

Simulation Framework

Programmable Engine for Drone Reinforcement Learning Applications (PEDRA) [1] is utilized as simulation environment.



Preliminary Simulations in PEDRA



[1] Anwar, A., & Raychowdhury, A. (2020). Autonomous navigation via deep reinforcement learning for resource constraint edge nodes using transfer learning. *IEEE Access*, *8*, 26549-26560.

Ongoing and Future Works

- Implementation of the proposed algorithm will be completed.
- Adaptability and transferability will be evaluated in indoor and outdoor maps.
- The proposed application will be extended for the urban airspace scenarios.

This work is supported, in part, by the Engineering and Physical Sciences Research Council [grant number EP/V026763/1]