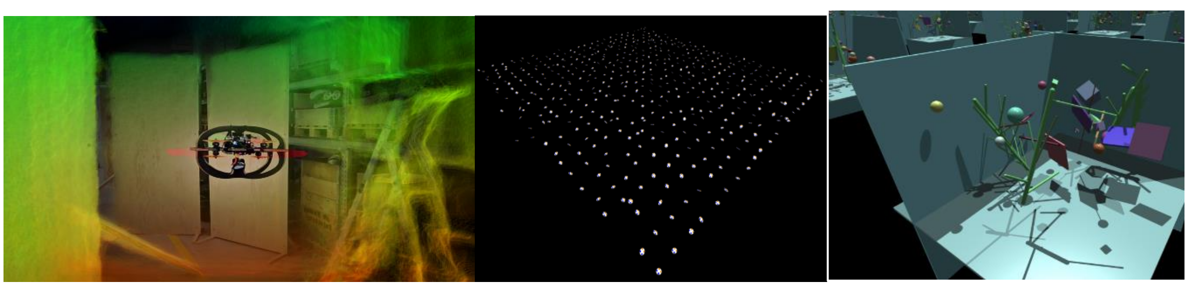
Deep Reinforcement Learning for Embedded Control Policies for Aerial Vehicles

Abstract: Unmanned aerial vehicles (UAVs) play an increasingly important role in industry and academia, being well suited for a wide variety of different tasks such as inspection and search and rescue missions. The area of small UAVs is dominated by designs developed from heuristics and empirical trial and error. While we employ sophisticated algorithms to generate optimal behavior strategies for these designs and don't rely on human intuition, we fail to extend this to the design process. Employing algorithmic frameworks to jointly optimize over behavior and design can mitigate the suboptimality resulting from the human bias. Deep Reinforcement Learning is a promising method for generating near optimal behavior for arbitrary robot morphologies. However, its efficacy relies heavily on the physical fidelity of the training environment [1]. Since the behavior of aerial robots is influenced by a variety of dynamics that are challenging to accurately model, The proposed master project aims to enhance the precision of the training environment's physical representation of the real world and investigate policy architectures to bolster the robustness of the learned policy and minimizing the sim2real gap. To validate the proposed changes the trained policies will be deployed on embedded platforms such as PX4/Betaflight/etc. to generate low-level motor commands by on-device inference. An important challenge in this project is the development of a pipeline to enable on-device inference for the trained networks e.g. by decreasing the inference time of the network via network quantization.



Tasks:

- Study and understand the basic Reinforcement Learning problem formulation, terminologies and common methods (Especially Proximal Policy Optimization (PPO)).
- Get familiar with the aerial-gym-simulator (github.com/ntnu-arl/aerial_gym_simulator) [6] for parallel GPU accelerated sample collection.
- Train a position tracking controller mapping position setpoints to low-level motor speed commands.
- Improvement of the policy architecture and the simulator with respect to sim2real transfer capabilities.
- Set up a pipeline to deploy the trained policies on embedded systems.
- Deploy the trained controller on a real-hardware platform in a motion-capture system

- Extensive experiments on physical systems to evaluate the effect of the developed methods and iteratively improve them.

Literature (indicative):

[1] W. Zhao, "Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey", 2020, IEEE Symposium Series on Computational Intelligence.
[2] J. Eschmann et al., "Learning To Fly in Seconds", 2023, arXiv Preprint.
[3] R. Ferede, "End-to-end Reinforcement Learning for Time-Optimal Quadcopter Flight", 2023, arXiv Preprint.
[4] A. Molchanov, T. Chen, W. Hoenig, J. A. Preiss, N. Ayanian, and G. S. Sukhatme, "Sim-to-(Multi)-Real: Transfer of Low-Level Robust Control Policies to Multiple Quadrotors," in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019, pp. 59–66.
[5] C.-H. Pi, K.-C. Hu, S. Cheng, and I.-C. Wu, "Low-level autonomous control and tracking of quadrotor using reinforcement learning", Control Engineering Practice, vol. 95, p. 104222, 2020.
[6] M. Kulkarni et al., "Aerial Gym -- Isaac Gym Simulator for Aerial Robots", 2023, ICRA 2023 workshop on The Role of Robotics Simulators for Unmanned Aerial Vehicles.
[7] Hwangbo, Jemin, et al. "Control of a quadrotor with reinforcement learning." IEEE Robotics and Automation Letters 2.4 (2017): 2096-2103.

Relevant Funded Project:

- **Title:** SPEAR: Spatial Perception & Embodied Autonomy Research
- **Funding Agency:** European Commission(Horizon Europe)

Main supervisor: Kostas Alexis, Professor, NTNU | **Co-supervisors:** Mihir Kulkarni, PhD Candidate, NTNU; Welf Rehberg, PhD Candidate, NTNU, Martin Jacquet, Postdoc, NTNU