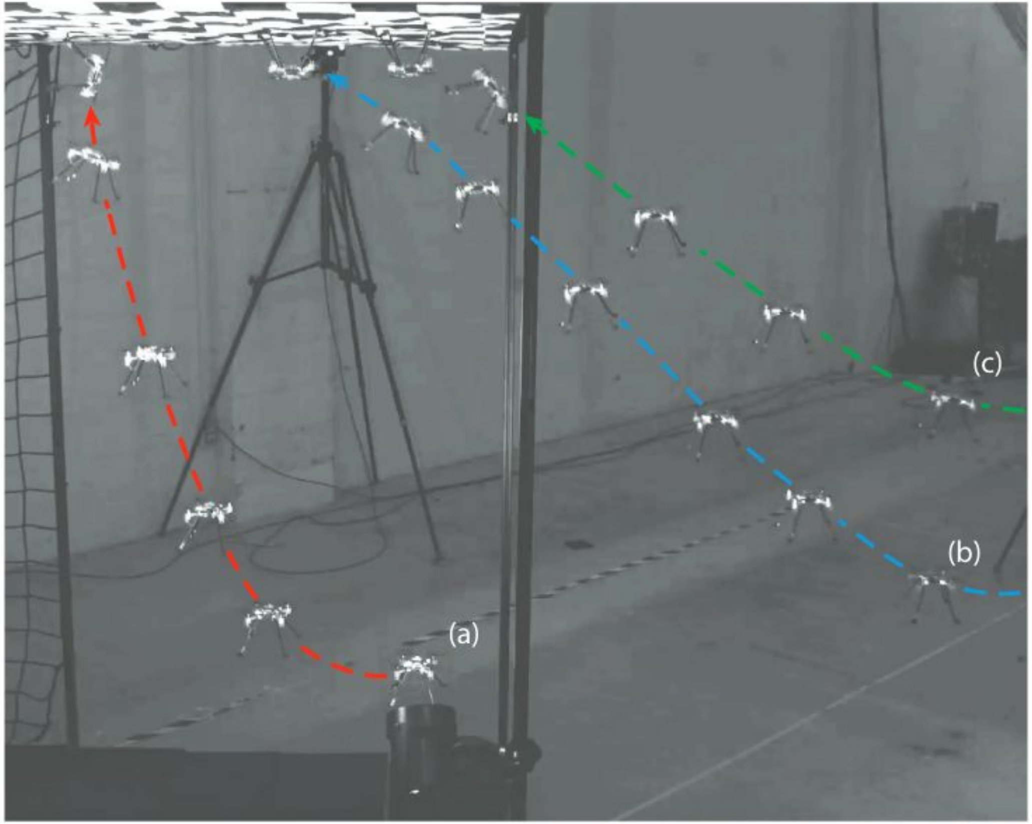
Inverted Landing in a Small Aerial Robot via Deep Reinforcement Learning for Triggering and Control of Rotational Maneuvers

Content

- Introduction
- Motivation
- Method
- Experiment
- Result
- Conclusion

Introduction

- UAV wildly used
- Landing



Motivation

- Inverted landing
    - Rapid & robust is challenging (depending on onboard sensing & computation)
    - Inspired by biological fliers such as bats, flies, and bees

- This work
    - Train a general, optimal control policy
    - Mapping the system’s emulated observational space to its motor command action space (triggering & rotational maneuvers)

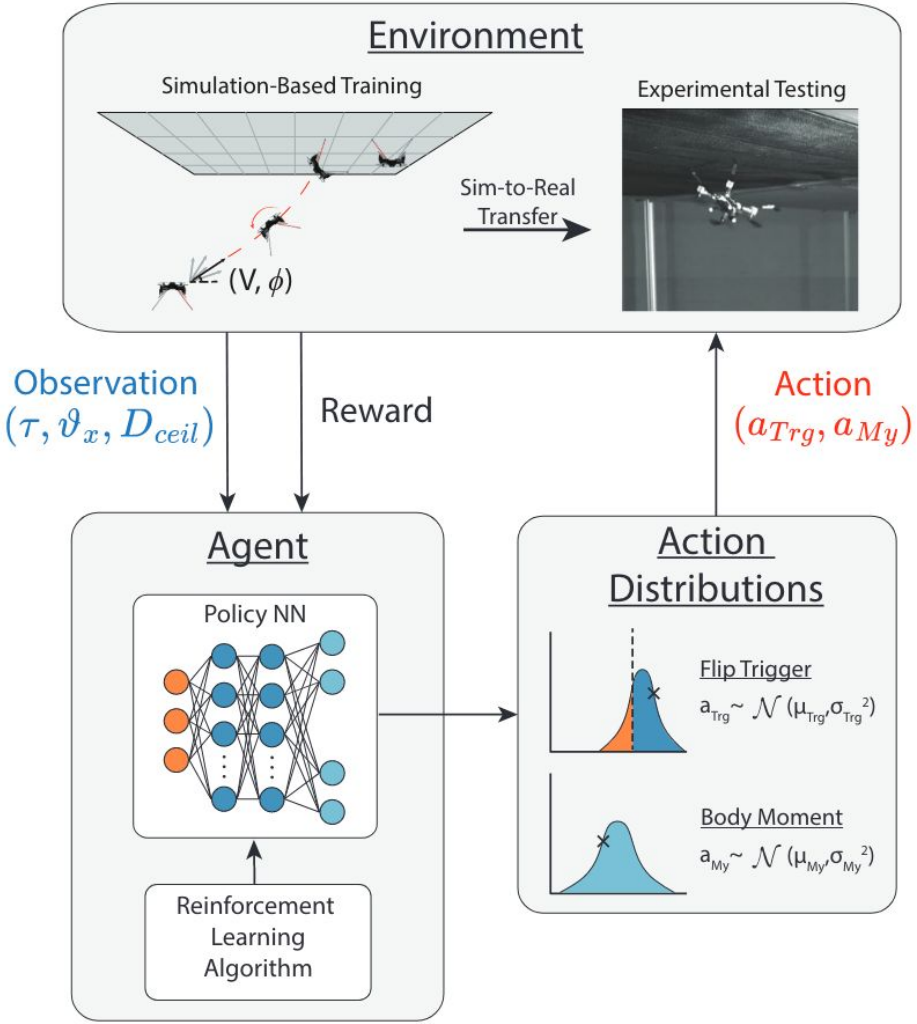
Method - Framework

Input

- Time-to-contact term (τ)
- Scaled velocity
- Distance to the ceiling

Output - Action distribution

- Flip trigger
- Body moment

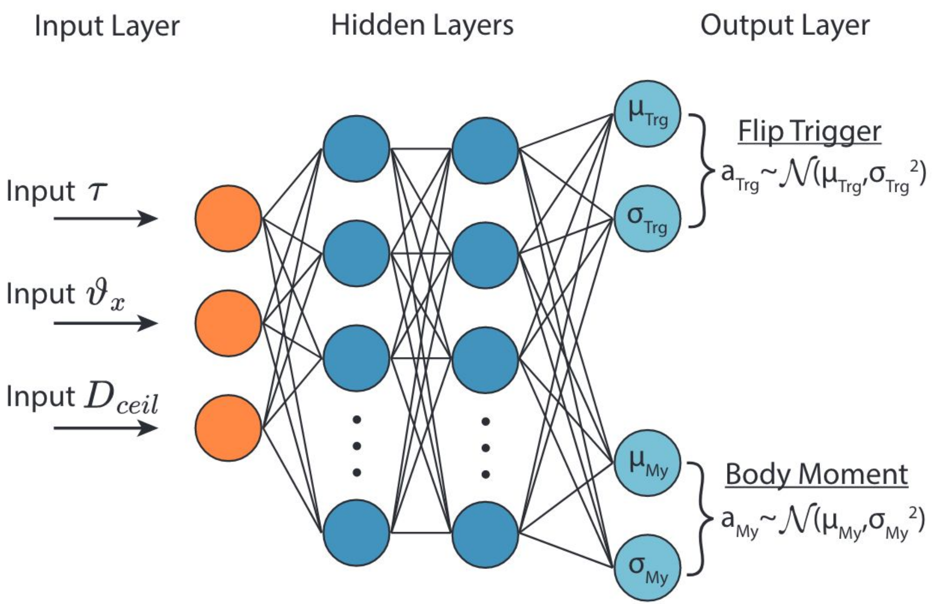


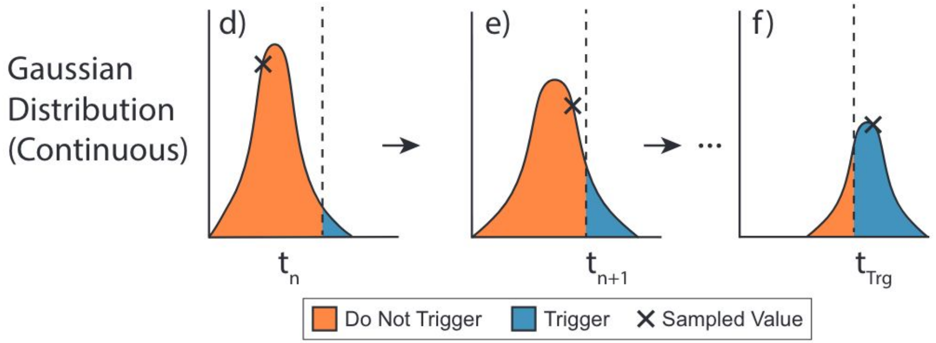
Method - Agent

Two action heads

- Event-triggering head
  - Triggers flip maneuver
- Rotation magnitude head
  - Predict body moment at the time of flip triggering

Sample actions





Method - Reward Design

- Soft Actor-Critic (SAC) algorithm
  - Maximize Objective function

$$J(\theta) = \mathbb{E}_{s_t \sim \rho^{\pi}, a_t \sim \pi_\theta} \left[R(s_t, a_t) + \beta H(\pi_\theta(s_t)) \right]$$

  - Reward R

$$R(s_t, a_t) = \sum_{k=0}^{T} \gamma^k r_{t+k+1},$$

    - All states but the triggering timestep, receive a reward of zero
    - All rewards from the final state are attributed to the action and observations that triggered the flip maneuver
  - Entropy regularization term: to encourage exploration

Method - Reward Design

- Four rewards (at triggering timestep)
  - Minimize the distance to the ceiling
  - Trigger the flip maneuver before colliding with the surface
  - Optimize the impact angle so the fore-legs contact first represented
  - Refine the above conditions to ensure reliable inverted landing
- Total reward for one episode

  r = 0.05 · rd + 0.1 · rτ + 0.2 · rθ + 0.65

$$r_d = clip\left(\frac{1}{|d_{min}|}, 0, c_0\right) \cdot \frac{1}{c_0},$$

$$r_\tau = clip\left(\frac{1}{|\tau_{trg} - 0.2|}, 0, c_1\right) \cdot \frac{1}{c_1},$$

$$r_\theta = \begin{cases} \frac{|\theta_{impact}|}{120^\circ} & 0^\circ \leq |\theta_{impact}| < 120^\circ \\ 1.0 & 120^\circ \leq |\theta_{impact}| \leq 180^\circ \end{cases},$$

$$r_{legs} = \begin{cases} 1.0 & N_{legs} = 3 \;||\; 4 \\ 0.5 & N_{legs} = 1 \;||\; 2 \\ 0 & N_{legs} = 0 \end{cases}.$$

Method - Sim2Real

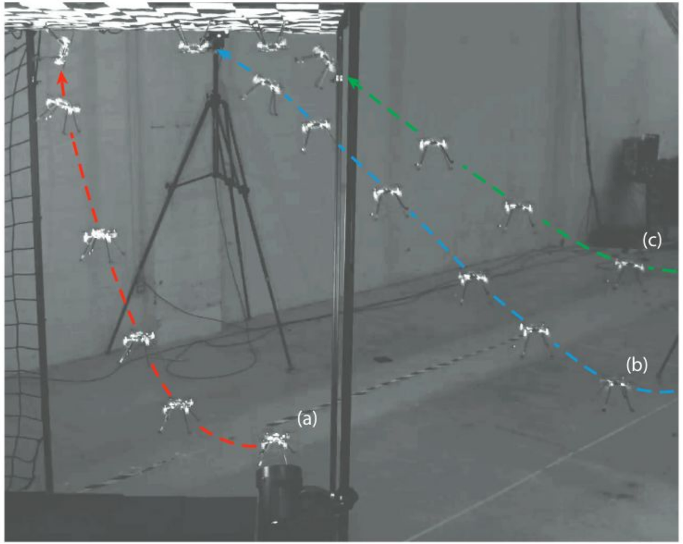
- Domain Randomization
  - Core idea
    - Randomize parameters in simulated environment (e.g., friction force (physical), object color (visual))
    - Ensure parameters in simulated environment include parameters in real environment with a high probability
  - Example
    - Require friction coefficient in the simulated environment between 0.1-1, while it may be 0.3 in real world
  - Advantage: allows migration to real world directly without further training
- Implementation
  - Sample system's mass & moment of inertia about the flip-axis from Gaussian distribution

Experiment

- Setup
  - Simulation environment: Gazebo
  - Randomized system mass & moment
  - Initial UAV state: specified approach velocity V and angle ϕ relative to the horizon
- Data collection
  - {observation, action} pairs
- Training
  - SAC (on-policy algorithm)
  - Back propogation

Result

- Predicted flight trajectory [result video]
- Convergence



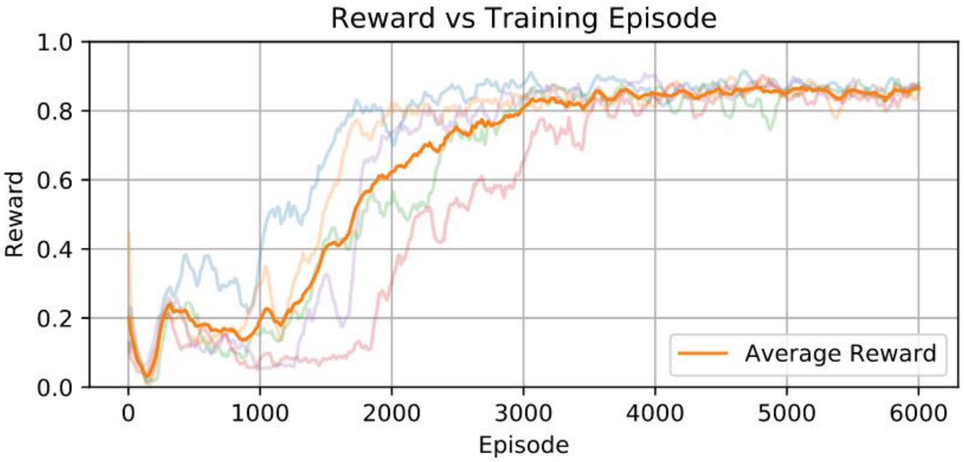


Fig. 4. Reward convergence plot showing consistent convergence to a generalized inverted landing policy within 3000 episodes/rollouts. The plot shows the convergence history for several training sessions and their average convergence behavior.

Result

- Success factors
  - Higher flight speed
  - Lower fligt angle

- Sim vs Real

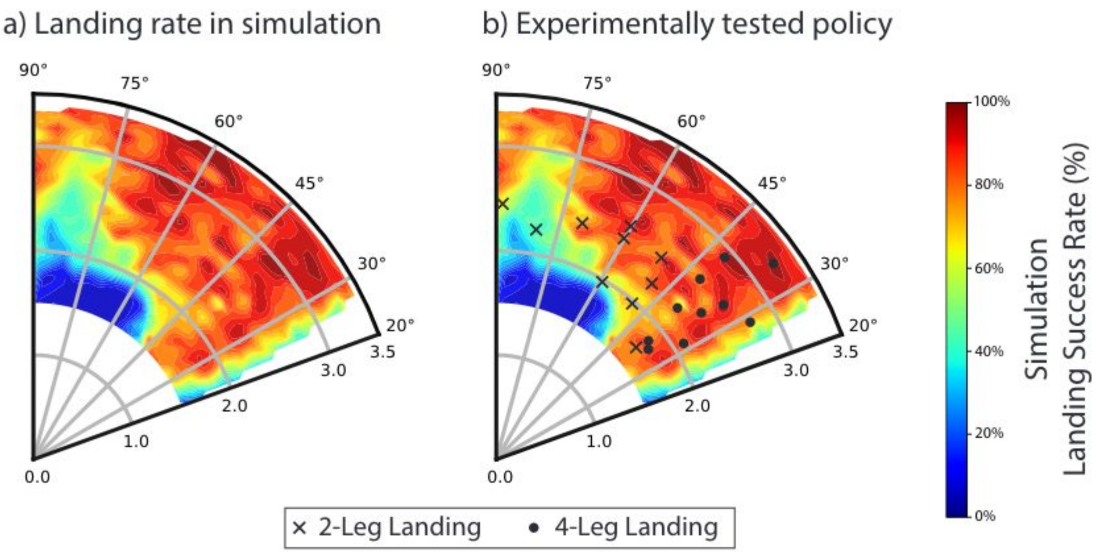


Fig. 5. (a) Plot showing the landing success rate of the trained policy over the simulated flight conditions, $V \in [1.5 - 3.5]$ m/s and $\phi \in [25^\circ - 90^\circ]$. (b) Markers indicating the success of the simulation-based policy transferred to experimental trials.

Result

- Map landing rate
  - Velocity
  - Triggering sensory-space

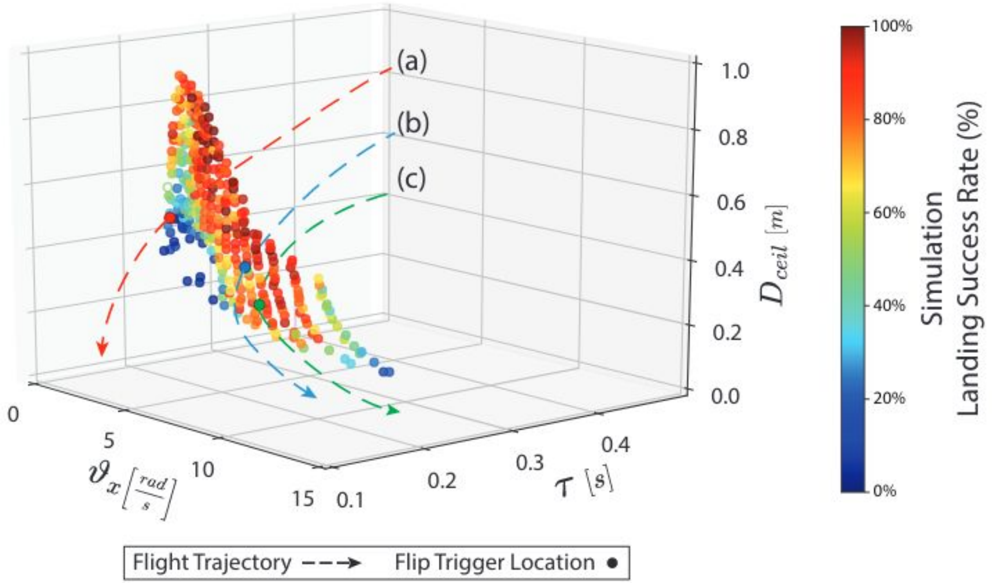


Fig. 6. Example flight trajectories for experimental inverted landing tests (Top) and their corresponding trajectories through observation space (Bottom). (a) Near vertical flight trajectory (2.25 m/s, 70°) with two-leg landing. (b) Angled flight trajectory (2.75 m/s, 40°) with successful four-leg landing. (c) Angled flight trajectory (2.50 m/s, 40°) with successful four-leg landing.

Conclusion

- Train a generalized inverted landing policy for small UAV
  - RL
  - Determing flip action from sensor observations

- Extension
  - Landing strategies
  - Path-planning and trajectory optimization problems