Simulation Based Autonomous RPAS Navigation Using Reinforcement Learning

Agata Szeremeta (agatasz@my.yorku.ca) | **Costas Armenakis** (armenc@yorku.ca)
Geomatics Engineering, Department of Earth & Space Science and Engineering, Lassonde School of Engineering, York University

I. MOTIVATION

- Use of small Remotely Piloted Aircraft Systems (RPAS, also known as UAVs) has become widespread, particularly for low altitude photogrammetric mapping activities

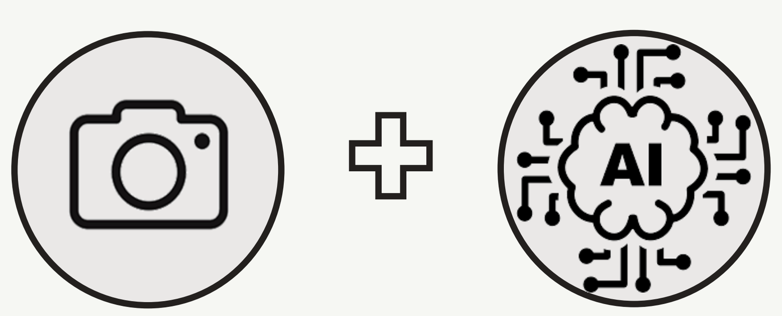
The success in using RPAS for aerial surveys depends on the need for their pilot operator to have a strong and complete knowledge of the flight environment, and good experience with the operation and mechanics of RPAS. Flying RPAS manually requires great amount of human involvement – including for mission/path planning and piloting, especially in GNSS-denied or poor environments.

- RPAS use is limited by navigation and sensor positioning challenges
- RPAS flight survey errors can introduce errors in final mapping product

II. RESEARCH OBJECTIVE

- Overcome limitations of RPAS navigation for mapping operations by integrating Artificial Intelligence (AI) and vision-based methods to enable RPAS to autonomously navigate along a specific trajectory

RPAS CAMERA



REINFORCEMENT LEARNING

- At present, focus is on following a road

III. RESEARCH CONTRIBUTIONS

Minimizes (human) pilot involvement and ensures quality of mapping products based on camera sensors, without need for additional navigation sensors on RPAS

Integrated AI and vision-based methods enable system to be used in GNSS-denied environments

Autonomous navigation contributes to successful and reliable use of RPAS beyond visual line of sight

IV. APPROACH

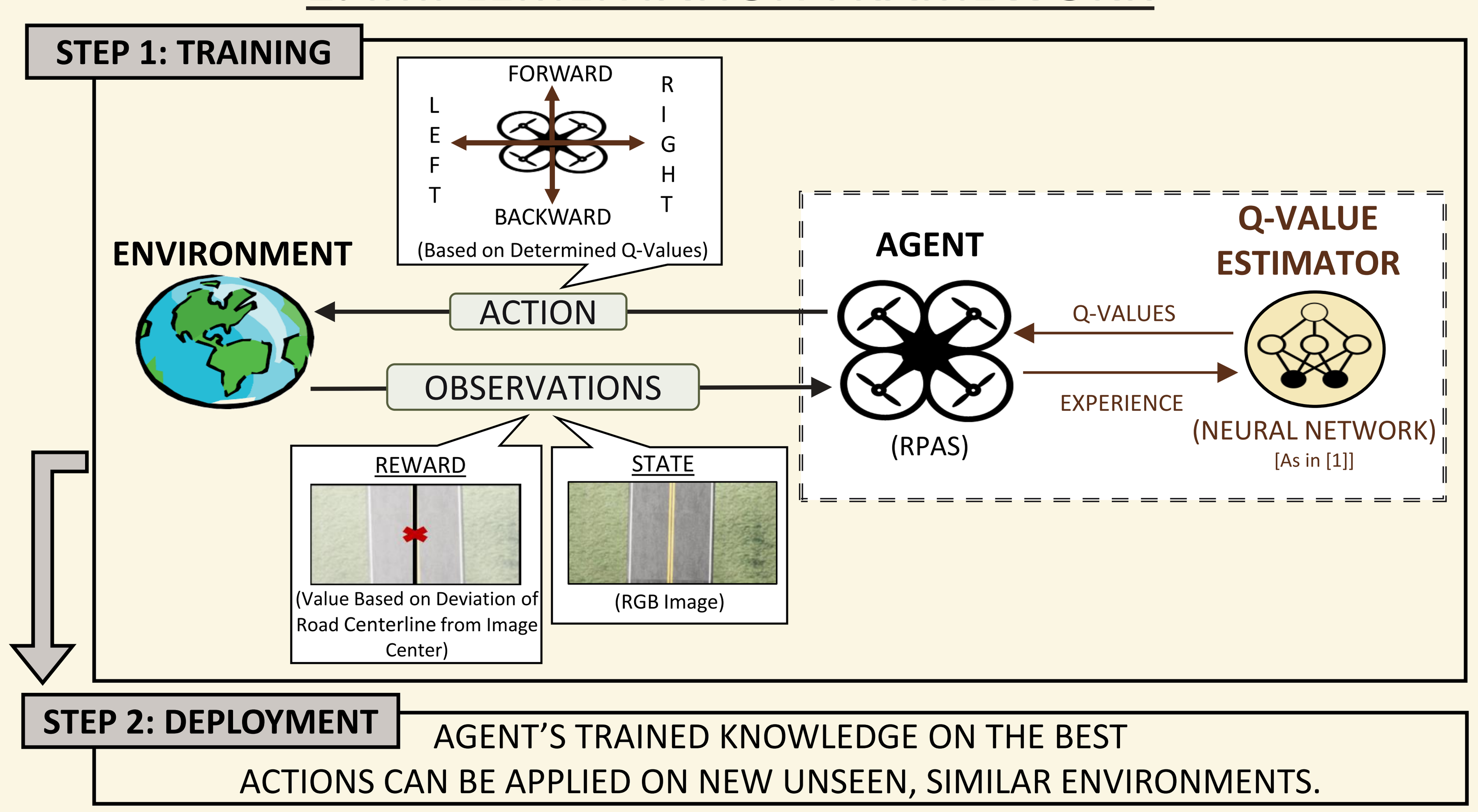
1. REINFORCEMENT LEARNING STRATEGY

- RPAS (also known as Agent) learns the best way to act in an environment through iterative trial and error experiences with it
- Agent receives rewards for actions taken in environment that work towards achieving its objective. Based on the received *reward (r)* signal, agent's neural network estimates the *Quality* (*Q-*) *Value* of taking certain *actions* (*a*) from *states* (*s*):

$$\underbrace{Q(s_t,a_t)}_{\text{The New Action Value}} = \underbrace{Q(s_t,a_t)}_{\text{The Old Value}} + \underbrace{\alpha}_{\text{The Learning Rate}} \times \left[\underbrace{r_{t+1} + \lambda \max_a Q(s_{t+1},a)}_{\text{The New Information}} - \underbrace{Q(s_t,a_t)}_{\text{The Old Information}}\right]$$

- The highest Q-Values define the optimal actions to take

2. IMPLEMENTATION FRAMEWORK



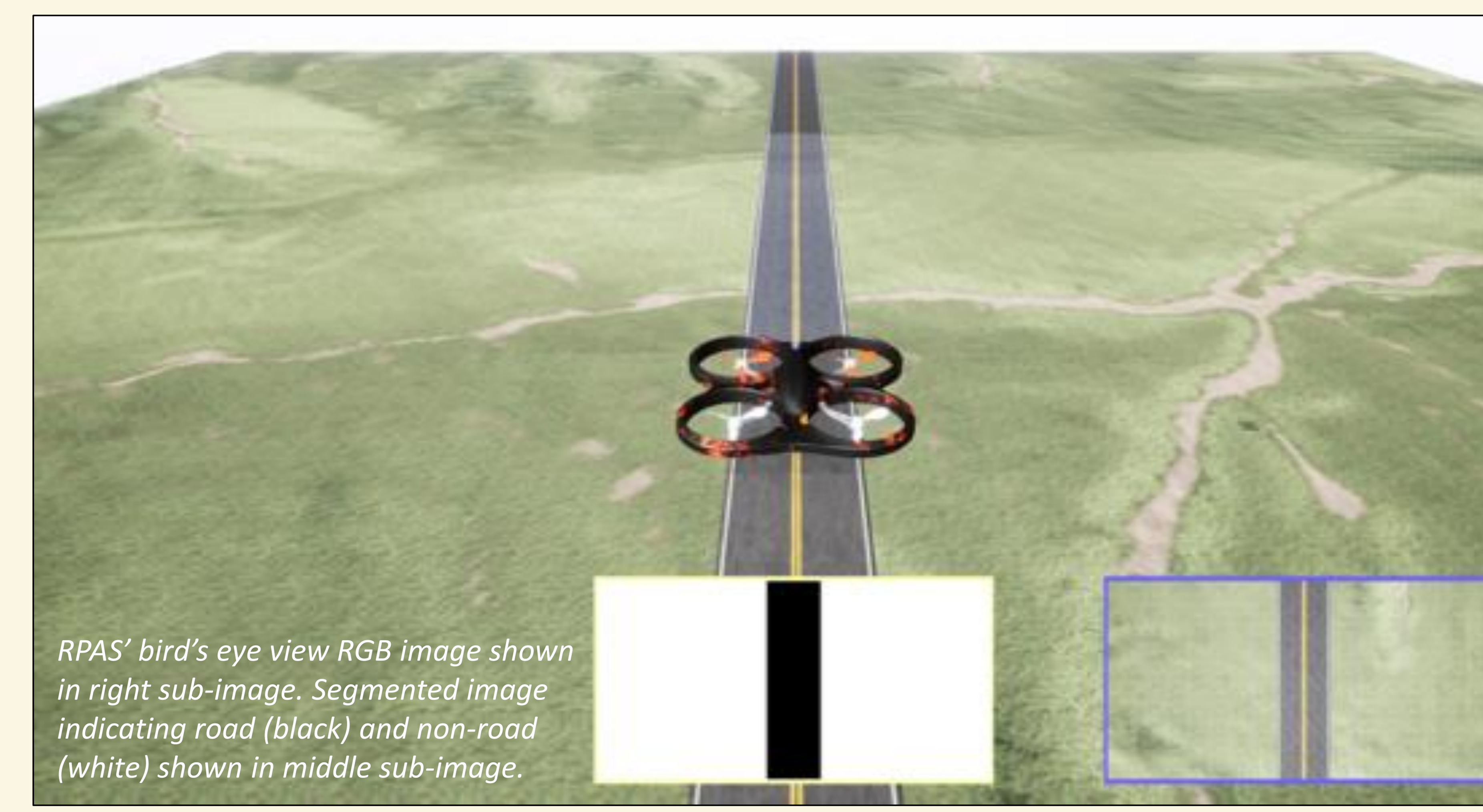
3. IMPLEMENTATION IN SIMULATED ENVIRONMENT

- Custom road environment created using Unreal Engine and Microsoft's AirSim simulator

RPAS' Training Goals in Simulated Environment:

Goal 1: Navigate from takeoff point to terminal point

Goal 2: Always have road in center of captured RGB bird's eye view image



RPAS' bird's eye view RGB image shown in right sub-image. Segmented image indicating road (black) and non-road (white) shown in middle sub-image.

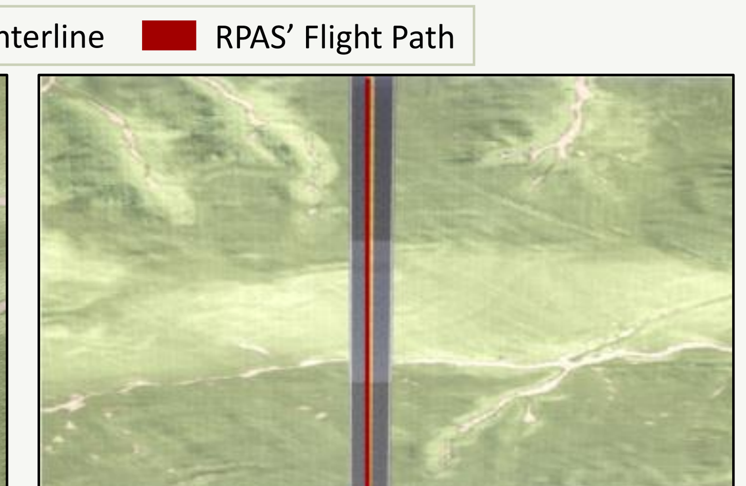
V. PRELIMINARY RESULTS

- RPAS successfully trained to fly along path following road feature in simulated environment

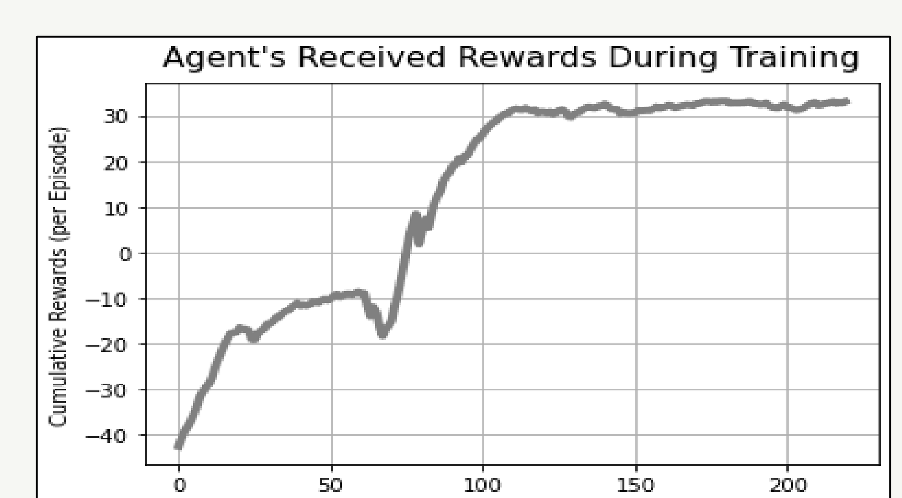
Total Training Time: ~18 hours Total Paths Travelled During Training: 220 Total Optimal Paths Taken During Training: 32

Legend: Road Centerline RPAS' Flight Path

Sample RPAS flight path during training.



RPAS flight path after training. As expected, RPAS navigates along road from start to end.



Smoothed rewards received over training time. Convergence indicates RPAS is taking optimal actions for navigating along road.

VI. FUTURE WORK

- Validate RPAS' obtained navigation knowledge against roads in different environments
- More versatile training, better deployment success
  - Continue training in more complex environments with more advanced RPAS dynamics (e.g., diverse terrain, different road shapes, and RPAS ascension/descension)
- Deployment in simulated actual environments

VII. REFERENCES

[1] Mnih et al. (2013). *Playing Atari with Deep Reinforcement Learning*.