Information-Driven Visual Navigation for UAVs Using Deep Reinforcement Learning

Active Visual Place Recognition in GPS-Denied Environments

Ing. Vojtěch Bešťák
Supervisor: Ing. Václav Truhlařík
Faculty of Information Technology, CTU in Prague

1 The Problem: Perceptual Aliasing

- UAV localization in GPS-denied zones relies heavily on Visual Place Recognition (VPR).
- In repetitive terrain (fields, forests), distant locations produce identical images, causing severe **perceptual aliasing**.
- Standard planners fly blindly along shortest paths, leading to state uncertainty and tracking failure.

(a) Aerial Orthophoto **(b)** Visual Salience

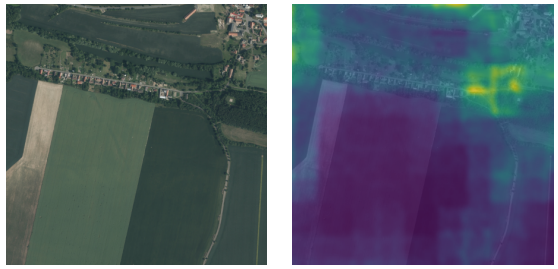

Figure 1: Repetitive terrain vs. salient visual landmarks.

2 The Solution & The Gap

- **Active Perception** solves this by using a reactive RL controller to steer the drone toward unique landmarks to maintain localization confidence.
- SOTA Vision Transformers (**SelaVPR++**) provide high VPR accuracy but are currently used only *passively*.
- *The Gap:* No prior work uses deep ViT saliency to actively steer continuous UAV trajectories.

3 System Formulation (POMDP)

Task formulated as a continuous POMDP for **Deep Reinforcement Learning**:

- *Hidden State (S_t):* True global coordinates and environmental map.
- *Observation (O_t):* Local RGB crop, SelaVPR++ saliency, trajectory history, and goal vector.
- *Multi-Objective Reward (R_t):* Blends dense goal-distance progress with a consumable visual saliency harvest.

4 Multi-Modal Architecture

- Fuses **4 input modalities**: RGB camera (I_t), SelaVPR++ saliency prior (M_t), visitation history (D_t), and goal telemetry (g_t).
- Trained via **PPO Actor-Critic** to output continuous velocity (v) and yaw rate (ω) commands.

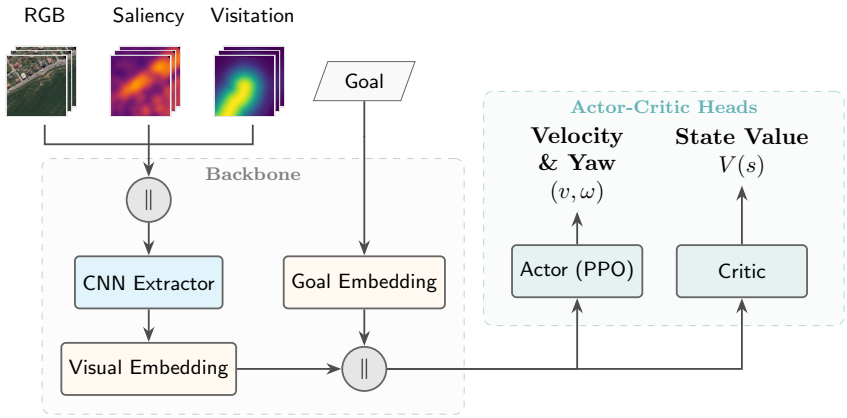

Figure 2: End-to-end multi-modal neural policy architecture.

5 Continuous Aerial Simulation

- Operates on authentic Czech orthophotos (ČÚZK) under second-order **unicycle kinematics** (inertia & yaw limits).
- Features a **Virtual EIS** module that decouples vehicle rotation from feature extraction to eliminate rotational aliasing.
- Treats visual saliency as a **consumable resource** within the camera footprint, forcing continuous active search.

(a) Aerial Orthophoto **(b)** Dynamic Depletion

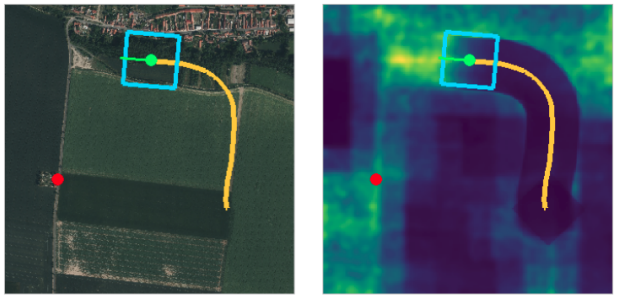

Figure 3: Consumable saliency footprint dynamics.

6 Results & Zero-Shot Generalization

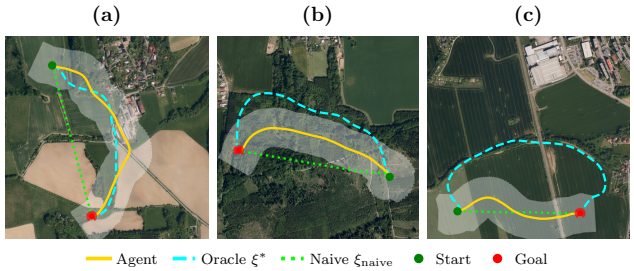

Figure 4: Agent actively detours to follow salient landmarks.

- **In-Domain:** Reduced **Average Localization Error (ALE)** by 1.6 m (closing 23.4% of the Oracle gap) with a 9.5% path detour.
- **Zero-Shot:** 100% mission success on unseen maps, closing **31.2% of the gap** to the Oracle without retraining.
- **Algorithmic Insight:** PPO outperforms Imitation Learning (DAgger, 53% success) by continuously correcting drift.

In-Domain Evaluation (500 episodes):

Method	Success	ALE [m] (↓)	Recall@1 (↑)
Naive Baseline	100%	53.35	0.658
Our Model	**99.4%**	**51.74**	**0.669**
Oracle Baseline	100%	46.46	0.708

7 Contributions & Impact

Publication & Open Source

- Validated active perception for VPR using Vision Transformers in continuous flight.
- **Target Publication:** V. Bešťák, V. Truhlařík. *Information-Driven Visual Navigation for UAVs Using Deep Reinforcement Learning.* In prep. for **IEEE ICRA 2027**.
- Serves as the foundation for upcoming **PhD research** at IMR CIIRC CTU.
- Custom simulator, training pipelines, and models are fully **open-sourced**.