UNDERLOC

IMAGE BASED RELOCALIZATION AND ALIGNMENT FOR DYNAMIC UNDERWATER ENVIRONMENTS

QUT Centre for Robotics

ICVSS 2025 — Sicily ~ 6-12 July 2025 — International Computer Vision Summer School

Gorry B., Fischer T., Milford M., Fontan A., Queensland University of Technology (QUT)
beverley.gorry@hdr.qut.edu.au

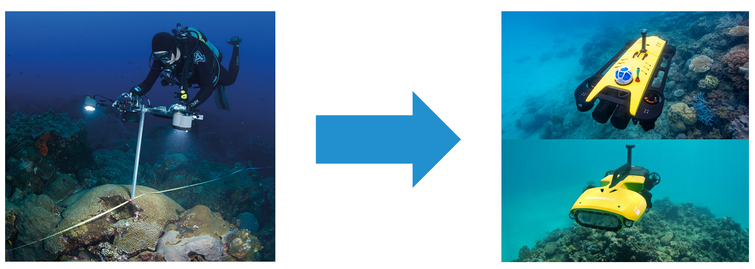
Abstract

Underwater ecosystem monitoring is crucial but challenging to automate due to the complexities of underwater imagery which hinder traditional visual methods. Our integrated **Visual Place Recognition (VPR)**, **feature matching**, and **image segmentation** pipeline enables robust identification of revisited areas and change analysis from **video-derived images**.

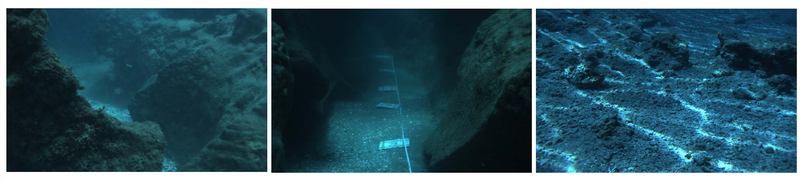
Furthermore, we introduce the SQUIDLE+ VPR Benchmark—a large-scale **underwater VPR benchmark** with diverse data from multiple robotic platforms.

Why Visual Place Recognition?

- VPR enables **effective, safe, and non-invasive monitoring** and conservation of vital yet threatened underwater ecosystems.



- Unique underwater conditions (turbidity, light attenuation, caustics, distortion) make reliable visual localization challenging.

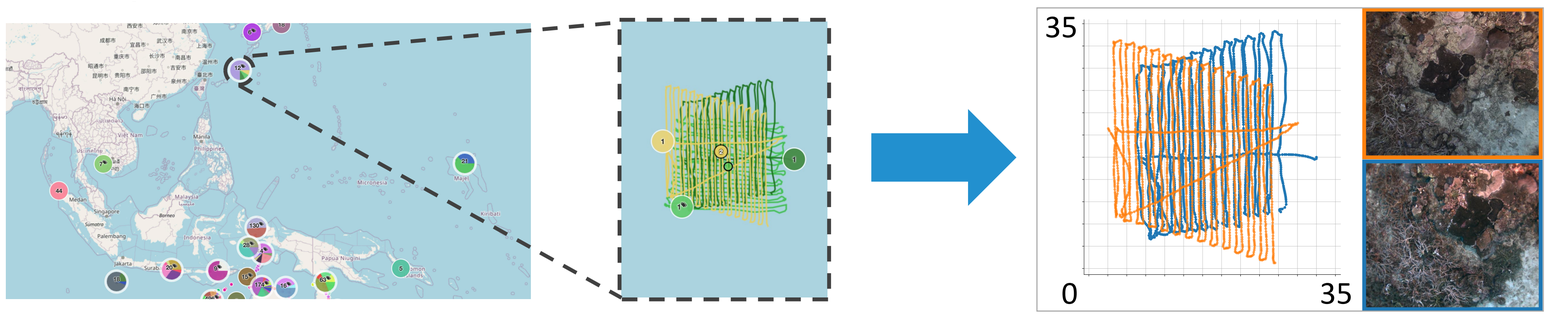


Turbidity | Light Attenuation | Caustics

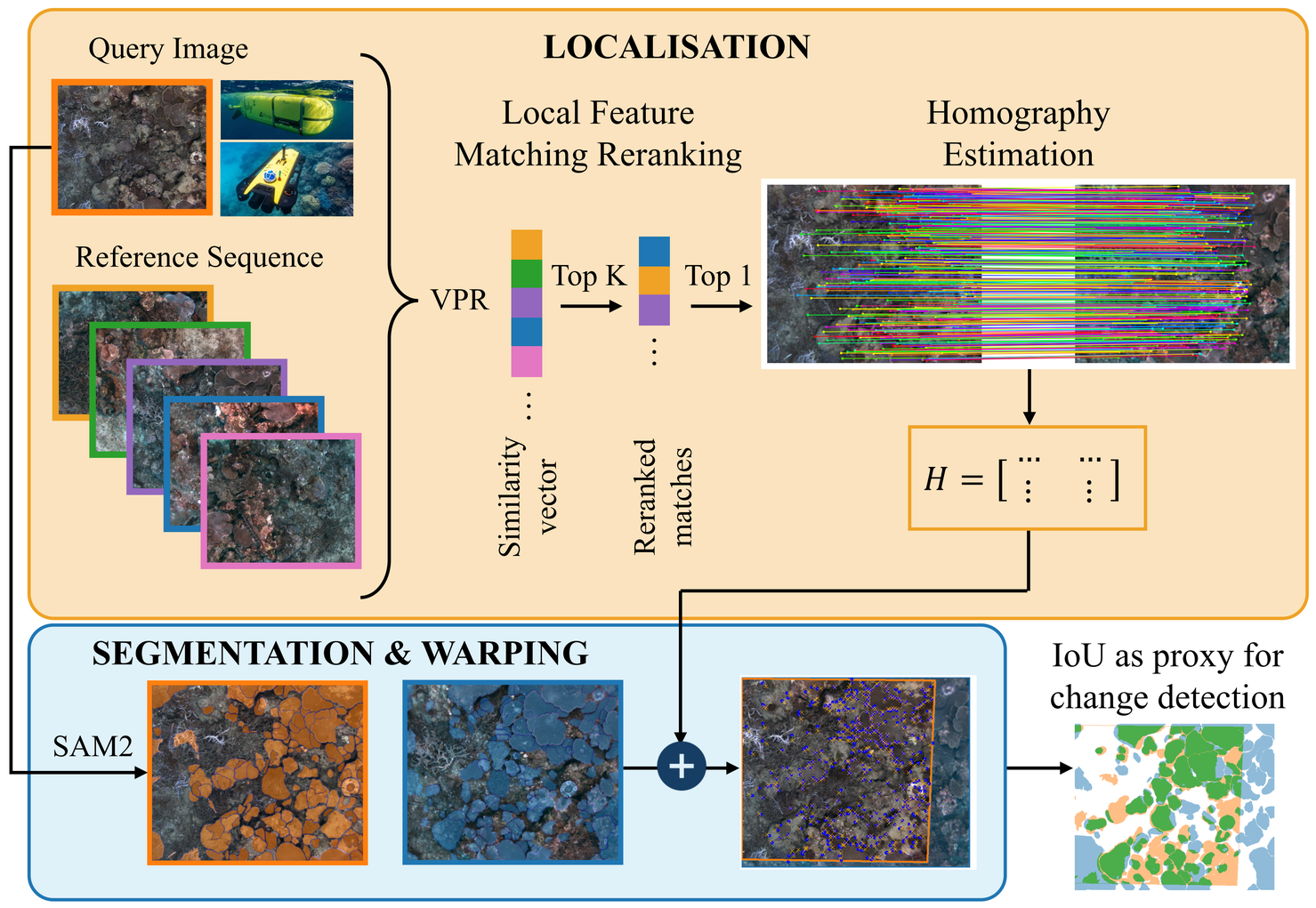
- Revisitation for tracking environmental changes and ensuring long-term ecosystem health.

SQUIDLE+ VPR Benchmark

- Current VPR benchmarks are largely based on terrestrial environments.
- We introduce a **new benchmark** for VPR in challenging underwater environments.
- 3 built-in sequences with code to **convert any SQUIDLE+ sequence** to our dataset format.
- Compatible dataset format with VSLAM Lab[1].



Technical Approach



UNDERLOC enables change detection using video streams from uncalibrated monocular cameras.

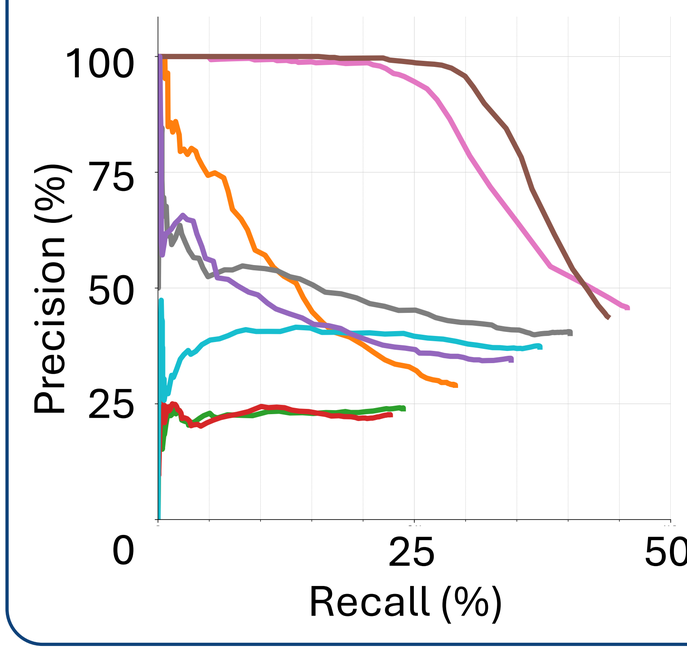
- VPR is used to find the **top K matches** per query.
- Matches are **reranked** and **filtered** using SuperPoint[2] keypoint correspondences and reprojection error.
- Semantic **segmentation masks** are extracted with SAM2[3] and **warped** using the homography matrix to enable change detection.

Find Us

Project Page

Github

Results



- Our hierarchical method achieves comparable performance to brute-force matching with **x100 speedup**.
- Robust performance in challenging underwater scenes.

	AUC-PR	Time per Query (ms)
Superpoint Brute-Force	0.39	27150
MegaLoc + Superpoint (Ours)	0.38	275

MIXVPR — COSPLACE — NETVLAD — ANYLOC — MEGALOC — CRICAVPR — SUPERPOINT-B — SUPERPOINT-H

References

[1] A. Fontan, T. Fischer, J. Civera, and M. Milford, "Vslam-lab: A comprehensive framework for visual slam methods and datasets," arXiv preprint arXiv:2504.04457, 2025.

[2] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 224–236.

[3] N. Ravi et al., "Sam 2: Segment anything in images and videos," 2024.

Acknowledgement

This research was partially supported by funding from ARC Laureate Fellowship FL210100156 to MM and ARC DECRA Fellowship DE240100149 to TF. The authors acknowledge continued support from the Queensland University of Technology (QUT) through the Centre for Robotics.