QUT Centre for Robotics

Underwater Robotic Vision for Ecosystem Monitoring and Reef Restoration

Dr Scarlett Raine

Chief Investigator and Lecturer,
Queensland University of Technology, Centre for Robotics

ACKNOWLEDGEMENT OF TRADITIONAL OWNERS

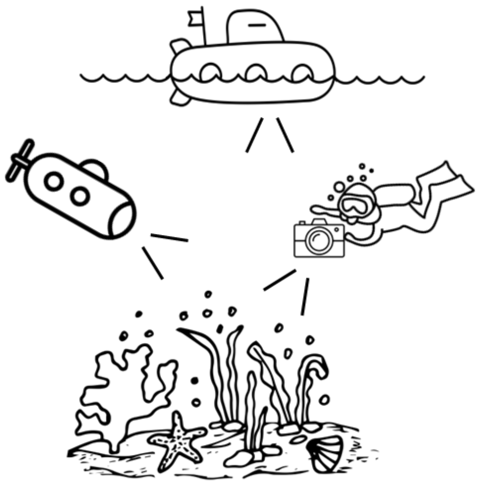
QUT acknowledges the Turrbal and Yugara, as the First Nations owners of the lands where QUT now stands. We pay respect to their Elders, lores, customs and creation spirits. We recognise that these lands have always been places of teaching, research and learning.

QUT acknowledges the important role Aboriginal and Torres Strait Islander people play within the QUT community.

Presentation Outline

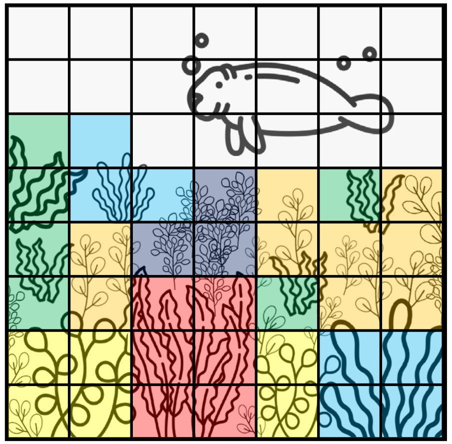
Introduction

Motivation

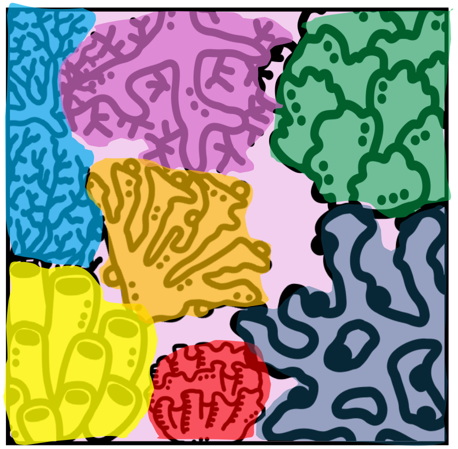


Bodies of Work

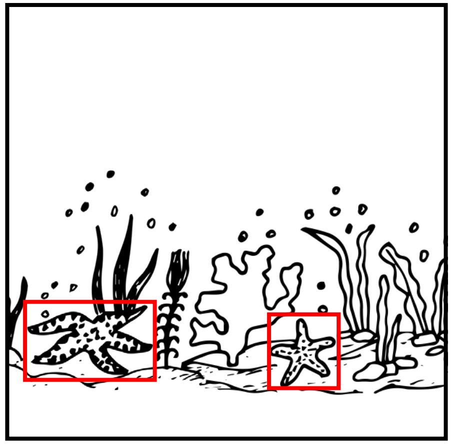
1. Seagrass: Coarse Segmentation



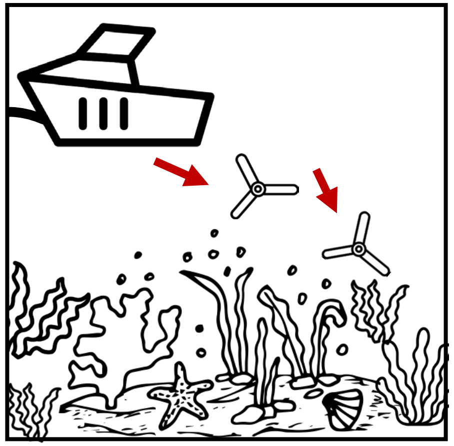
2. Coral: Segmentation



3. Underwater Object Detection



4. Large Scale Reef Restoration (with RRAP)

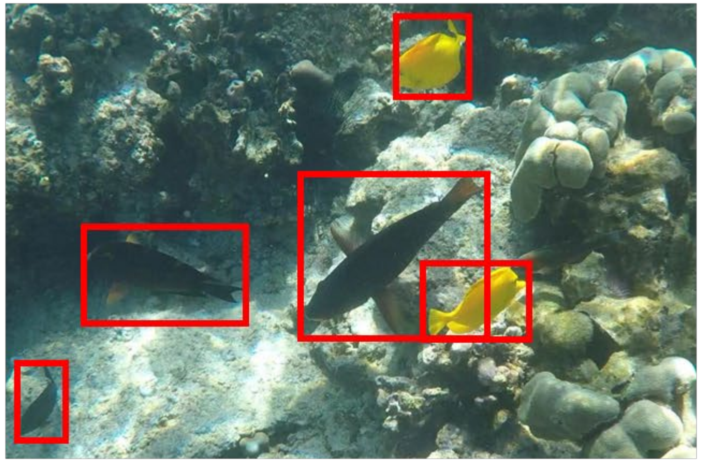


Future Work

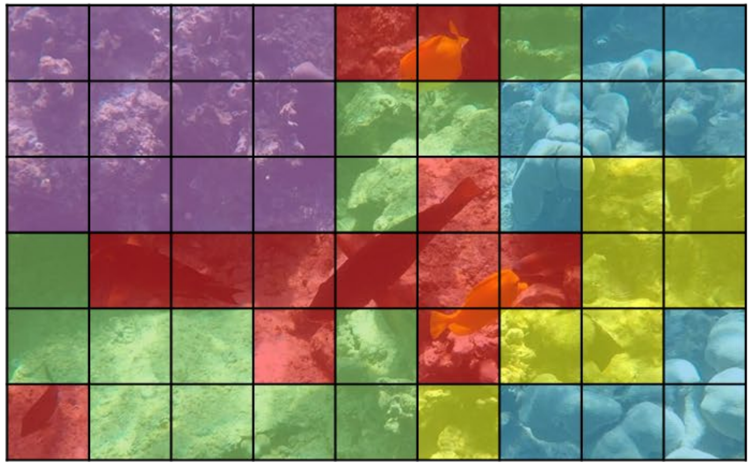
- Lecturer and Chief Investigator in QUT Centre for Robotics
- Expertise in Artificial Intelligence, Computer Vision and Robotics
- Passionate about environmental monitoring and conservation, with a particular focus on coral reefs, seagrass meadows and coastal ecosystems

"coral"

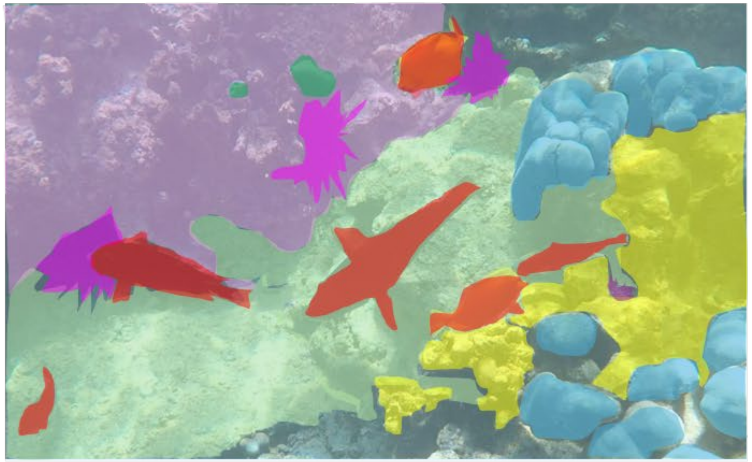
Classification



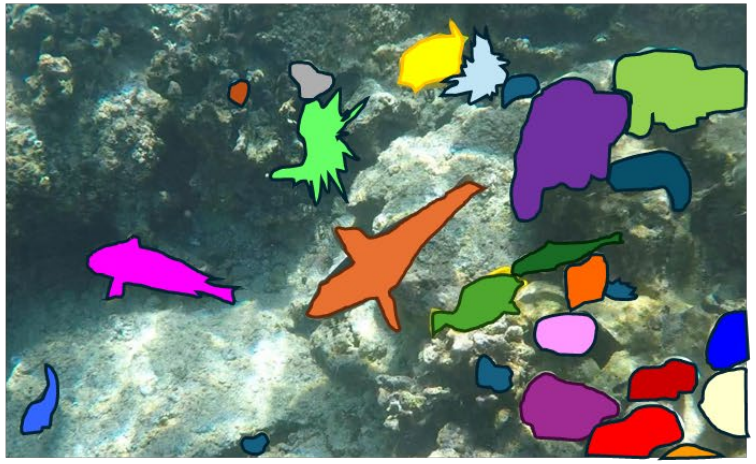
Object Detection



Coarse Segmentation



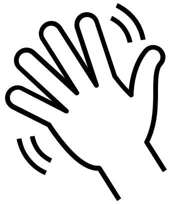
Semantic Segmentation



Instance Segmentation

Legend

	Coral Species A
	Coral Species B
	Coral Species C
	Coral Species D
	Fish
	Urchin
	Substrate
	Unknown



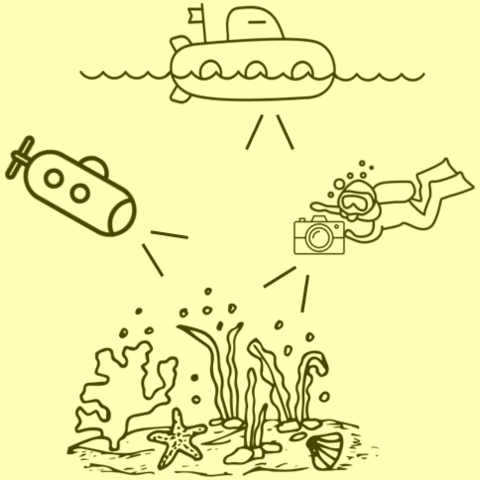
NOTE: Please interrupt if I use a word that you don't understand!

Presentation Outline

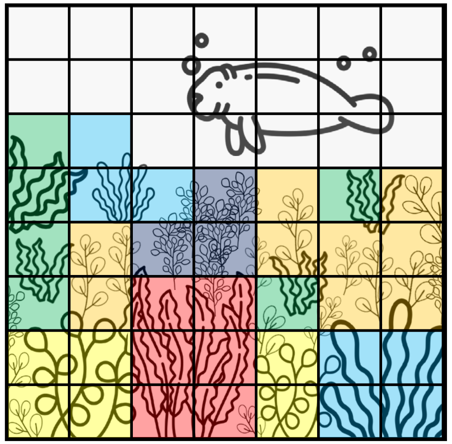
Bodies of Work

Intro

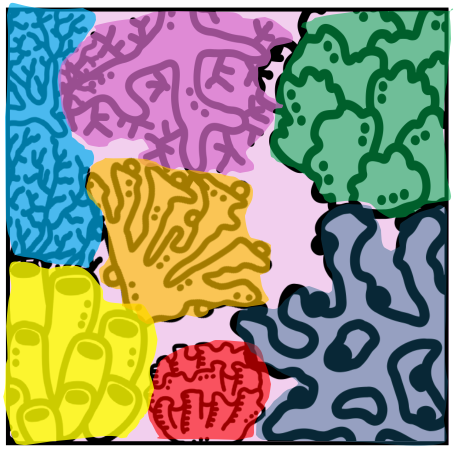
Motivation



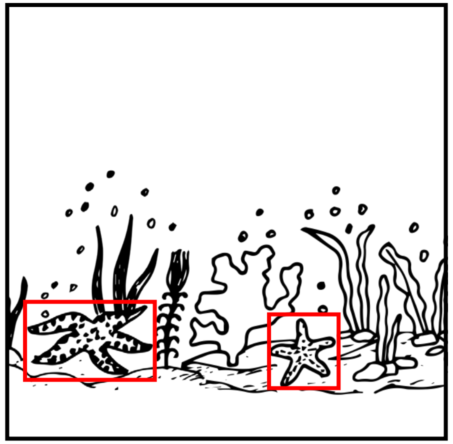
1. Seagrass: Coarse Segmentation



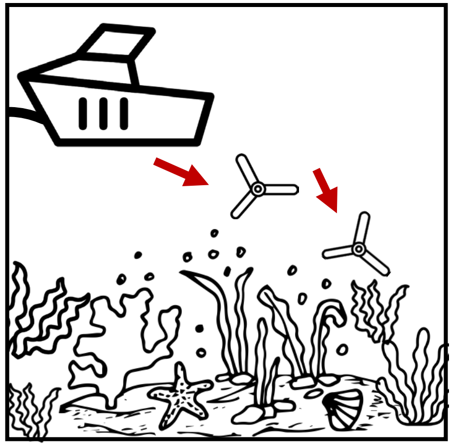
2. Coral: Segmentation



3. Underwater Object Detection



4. Large Scale Reef Restoration (with RRAP)

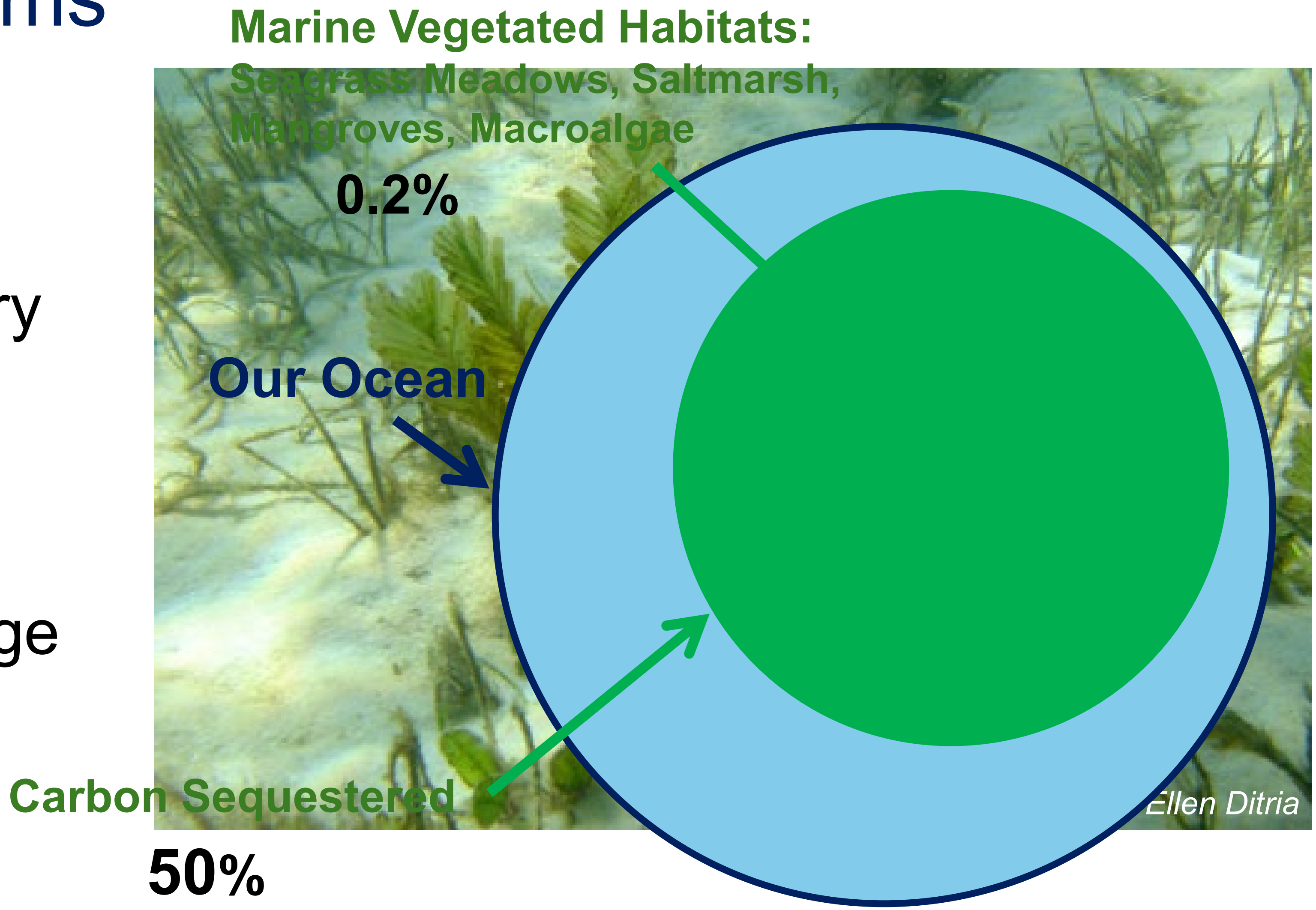


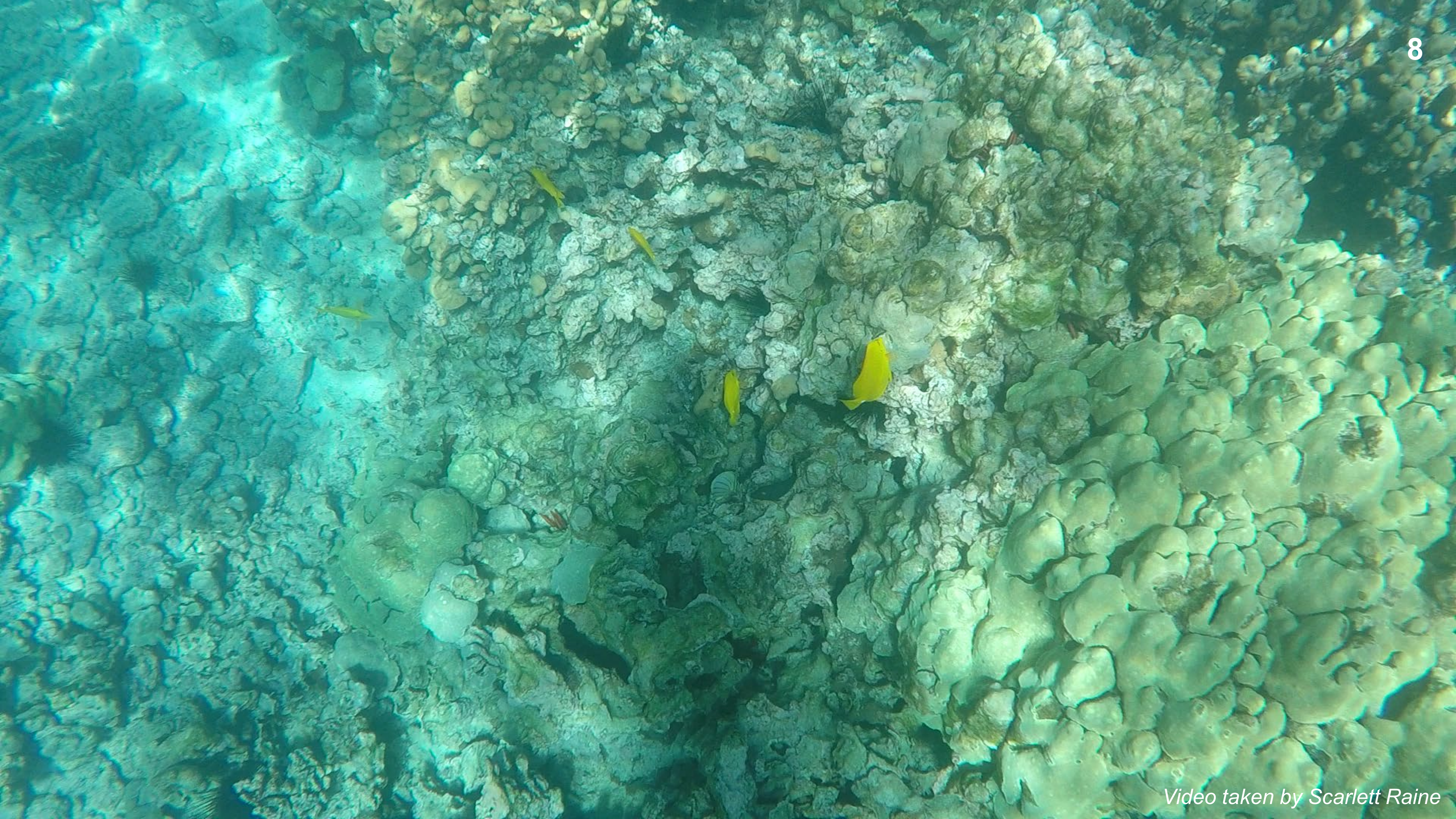
Future Work

Underwater Ecosystems

1. Seagrass Meadows

- Multi-functional role: support fisheries, nursery for marine species, source of food for dugongs
- Sequester and store large quantities of carbon





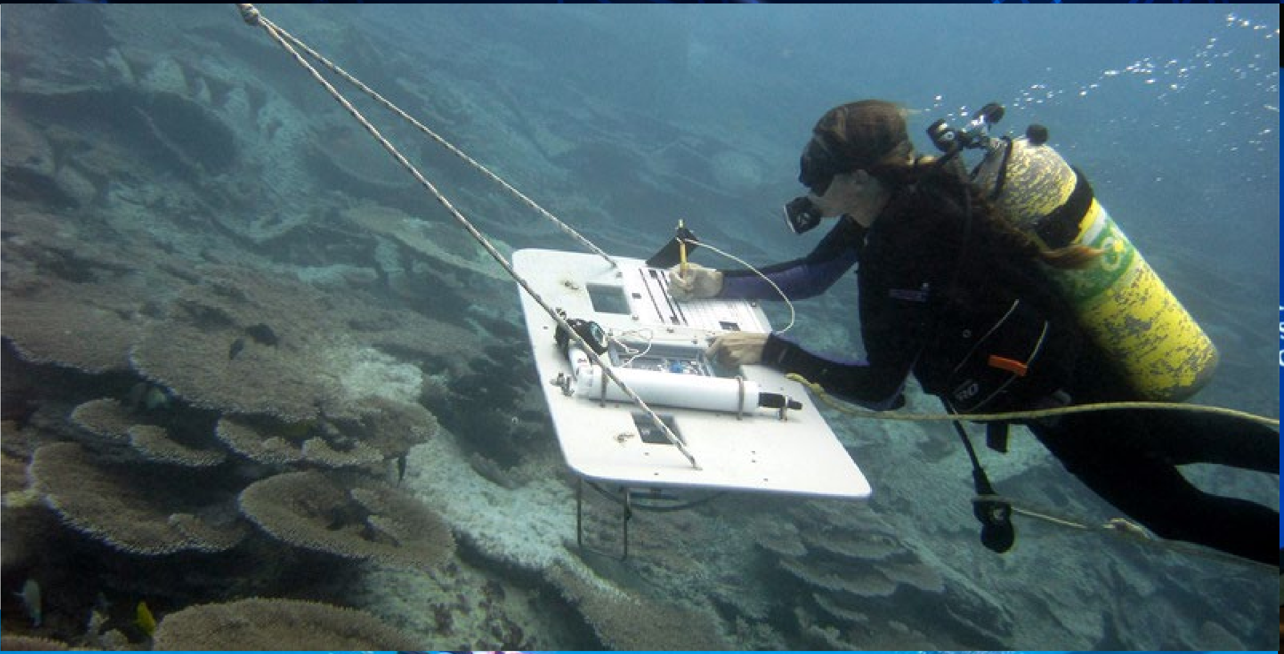
Video taken by Scarlett Raine

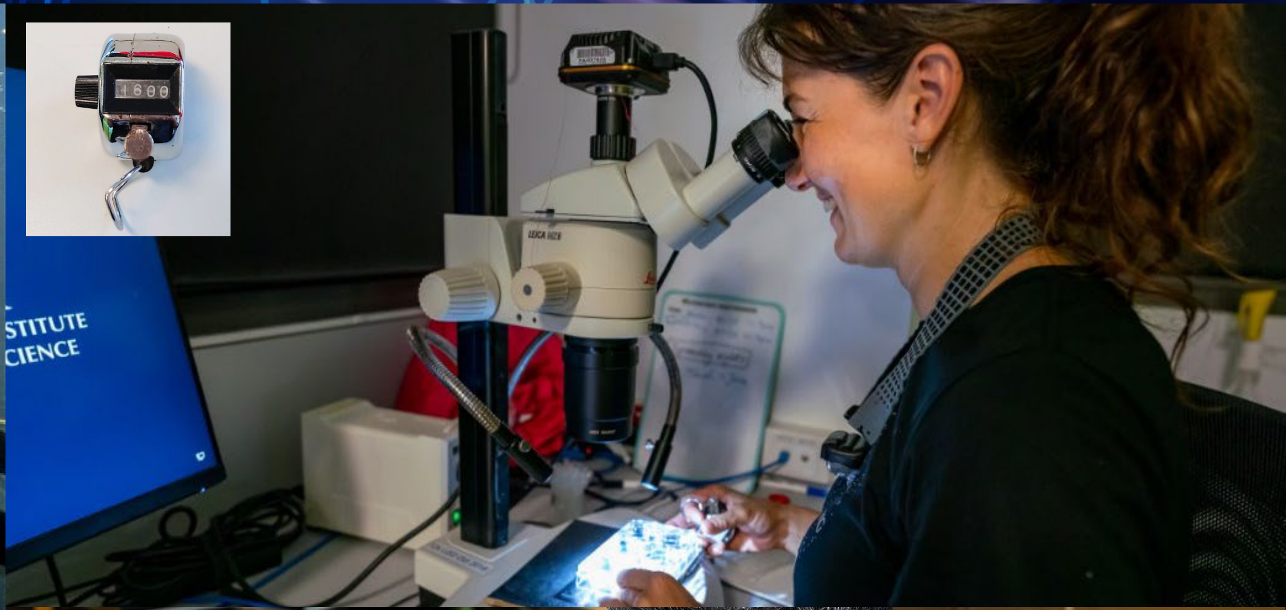
Underwater Ecosystems

2. Coral Reefs

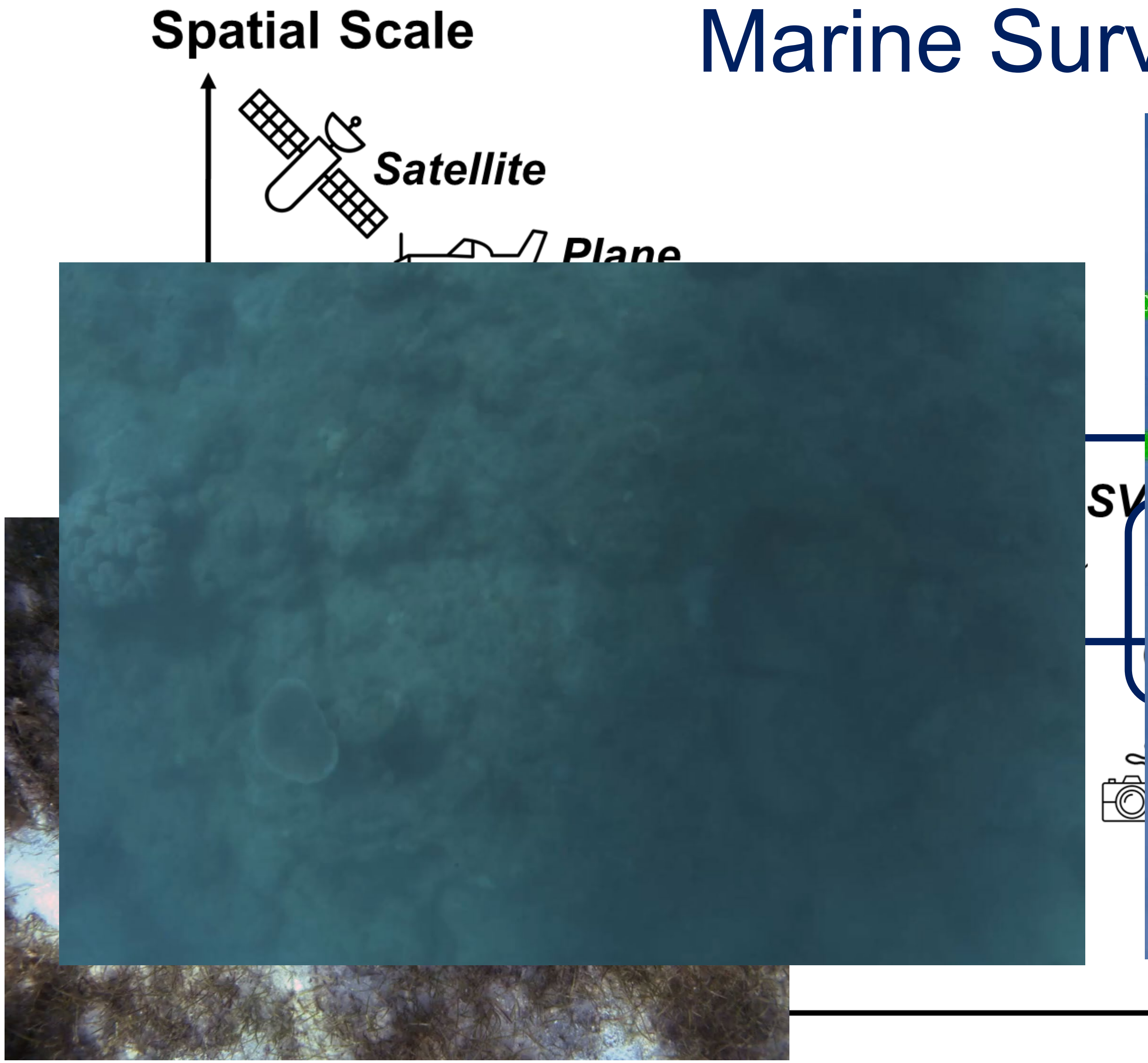
- Wealth of marine life and biodiversity, protect coastlines
- Support economies through tourism and fishing
- Under significant threat due to climate change

Image from NOAA Public Domain Library



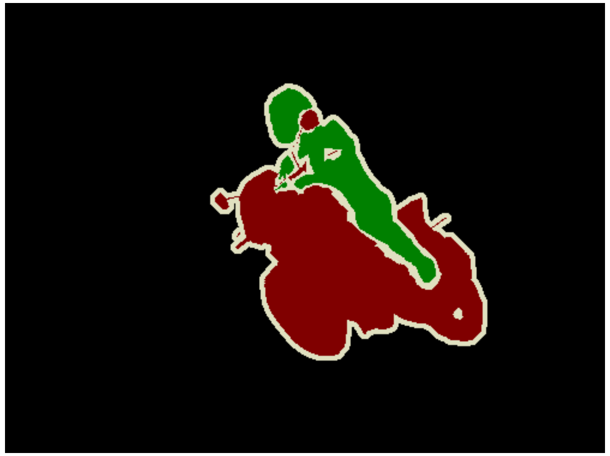


Marine Survey Methods



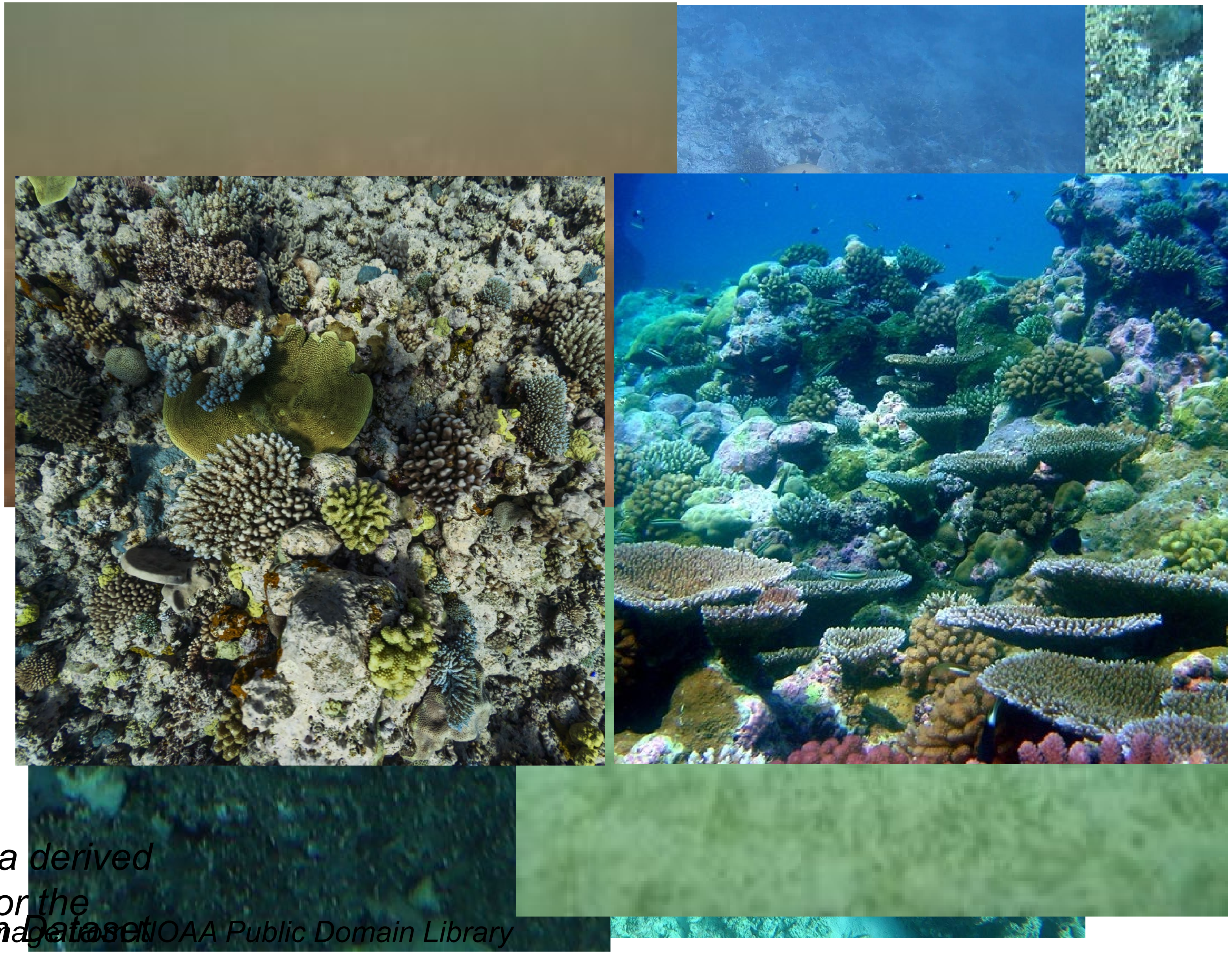
Typical Computer Vision Images

- Foreground/background separation
- Objects are defined
- Annotator agreement
- Can use crowd-sourcing to label



Underwater Imagery

- Variation in Underwater Image Characteristics:
  - Depth
  - Lighting
  - Camera Angle and Distance
  - Turbidity
  - Blur



Images from ‘Benthic and substrate cover data derived from a time series of photo-transect surveys for the Eastern Banks, Moreton Bay (Australia)

Images from CSIRO Deep Sea Coral and Rubble Dataset

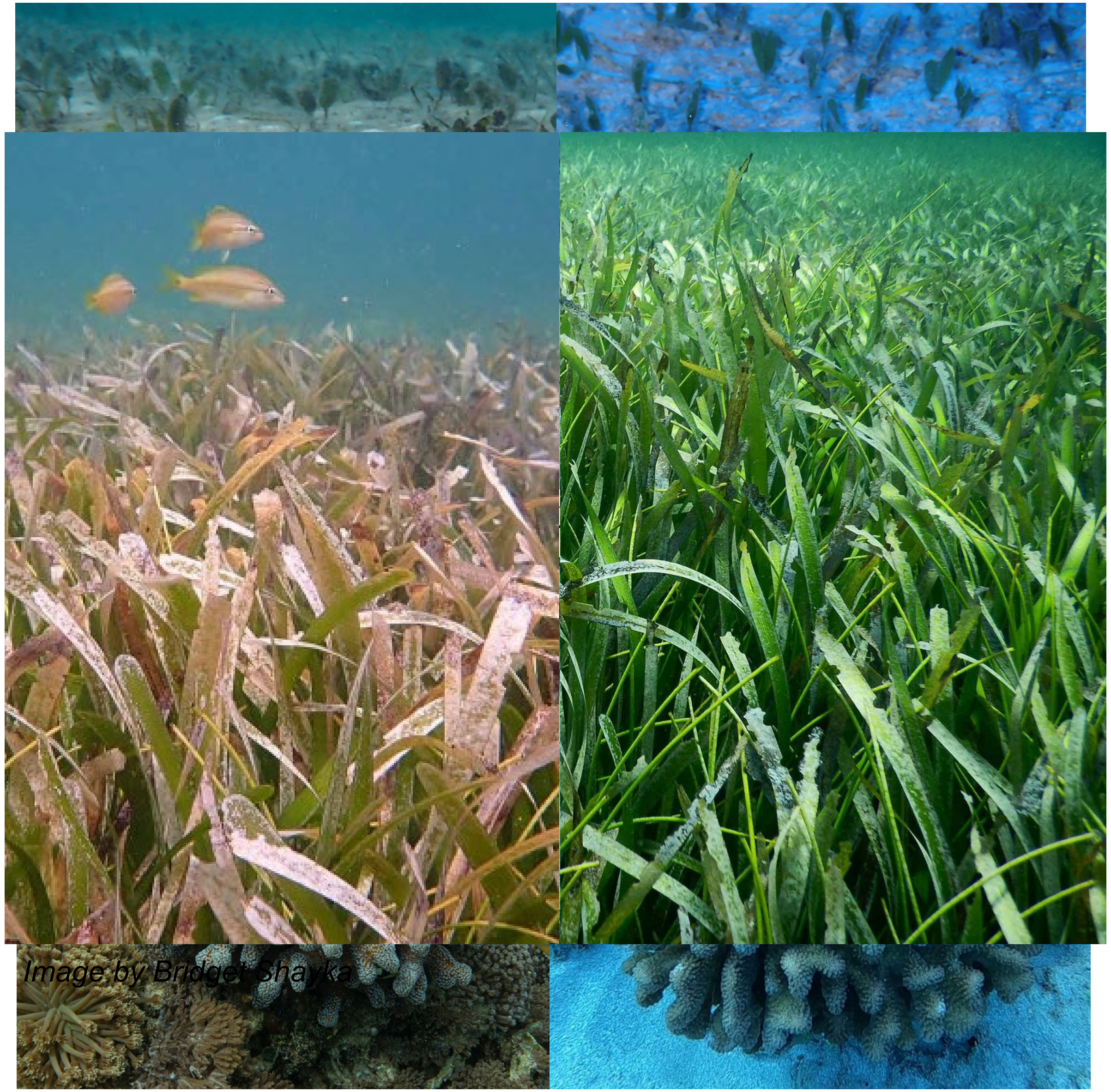
Images from CSIRO Crown of Thorns Starfish Dataset

Left image from XL Catlin Seaview Survey, right image from NOAA Public Domain Library

Underwater Imagery

- Species appearance varies due to:
  - Seasonality
  - Location
  - Environmental conditions
- Visual similarities between different species

- ***Underwater images must be labelled by domain experts***

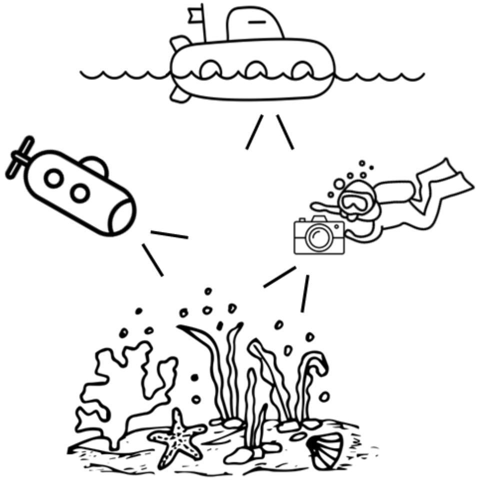


Presentation Outline

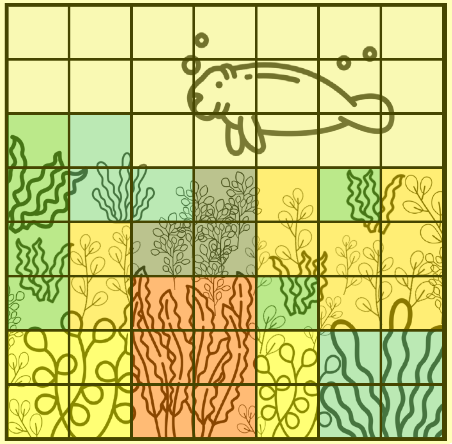
Bodies of Work

Intro

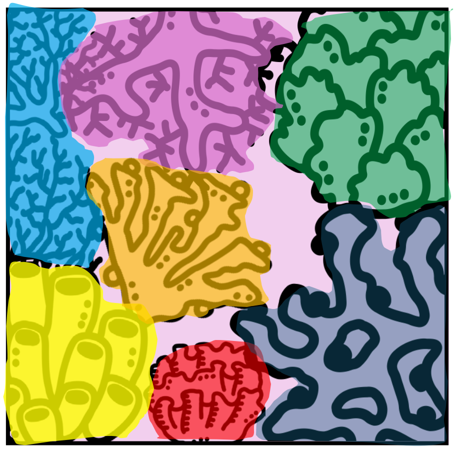
Motivation



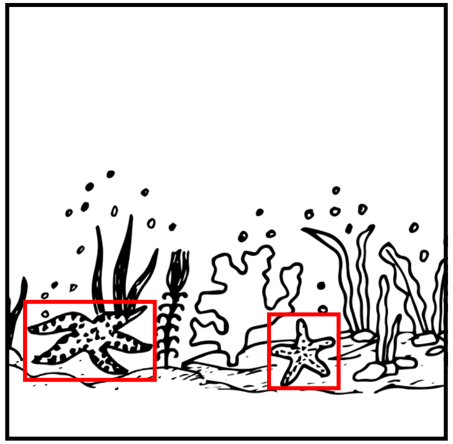
1. Seagrass: Coarse Segmentation



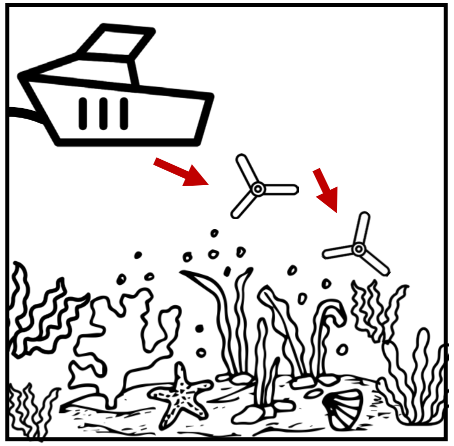
2. Coral: Segmentation



3. Underwater Object Detection



4. Large Scale Reef Restoration (with RRAP)

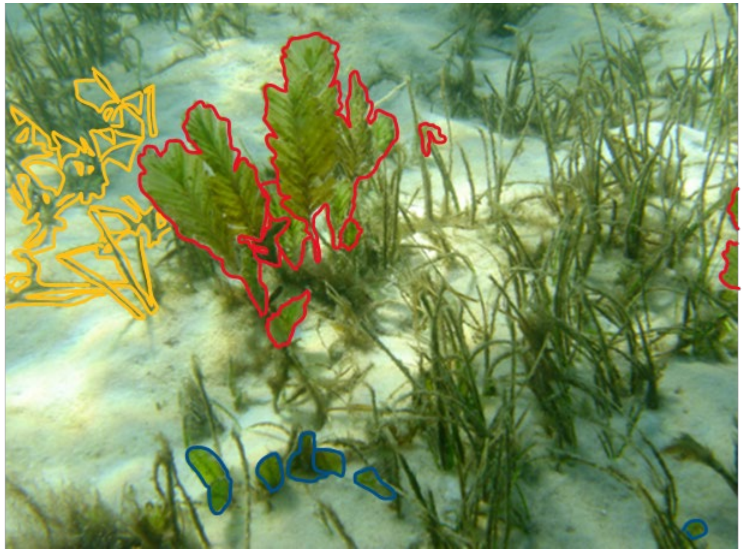


Future Work

Seagrass Mapping

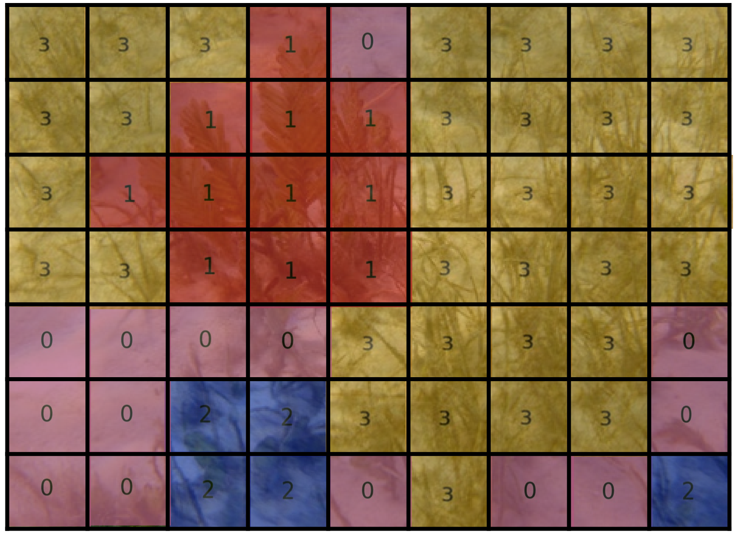
Image level label: “Background”, “Ferny”, “Rounded”, “Strappy”

Segmentation



- Instances are poorly defined and overlapping
- Pixel-wise labels too costly and time-consuming

Coarse Seagrass Segmentation



- Estimates species composition and extents
- Sufficient resolution, considering scale of seagrass meadows

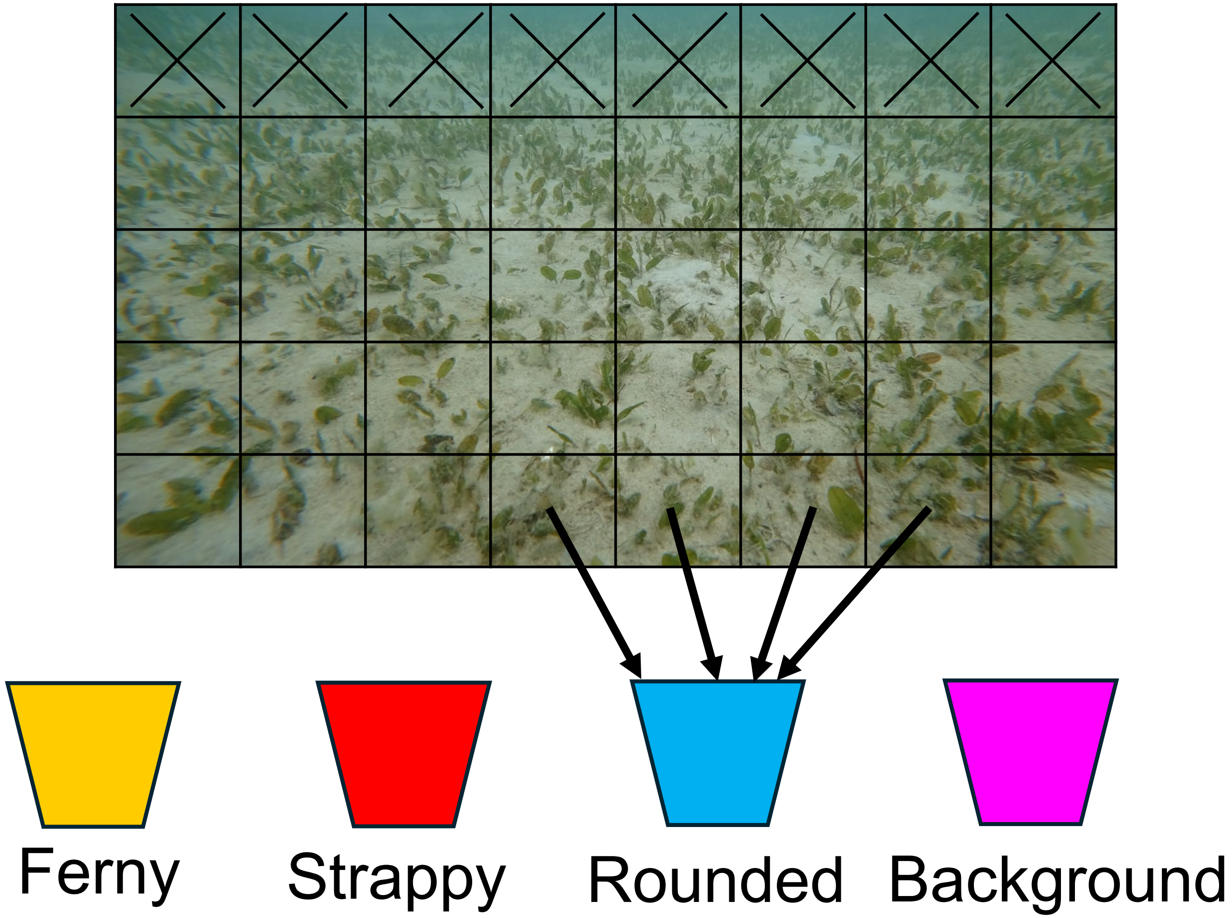
Legend: 0 Background 1 Ferny 2 Rounded 3 Strappy × Incorrect

Photo credit: Ellen Ditria.
From: https://catchmenttocoast.org/2020/06/04/how-can-we-best-assess-the-threats-and-status-of-connected-coastal-wetland-habitats/

DeepSeagrass

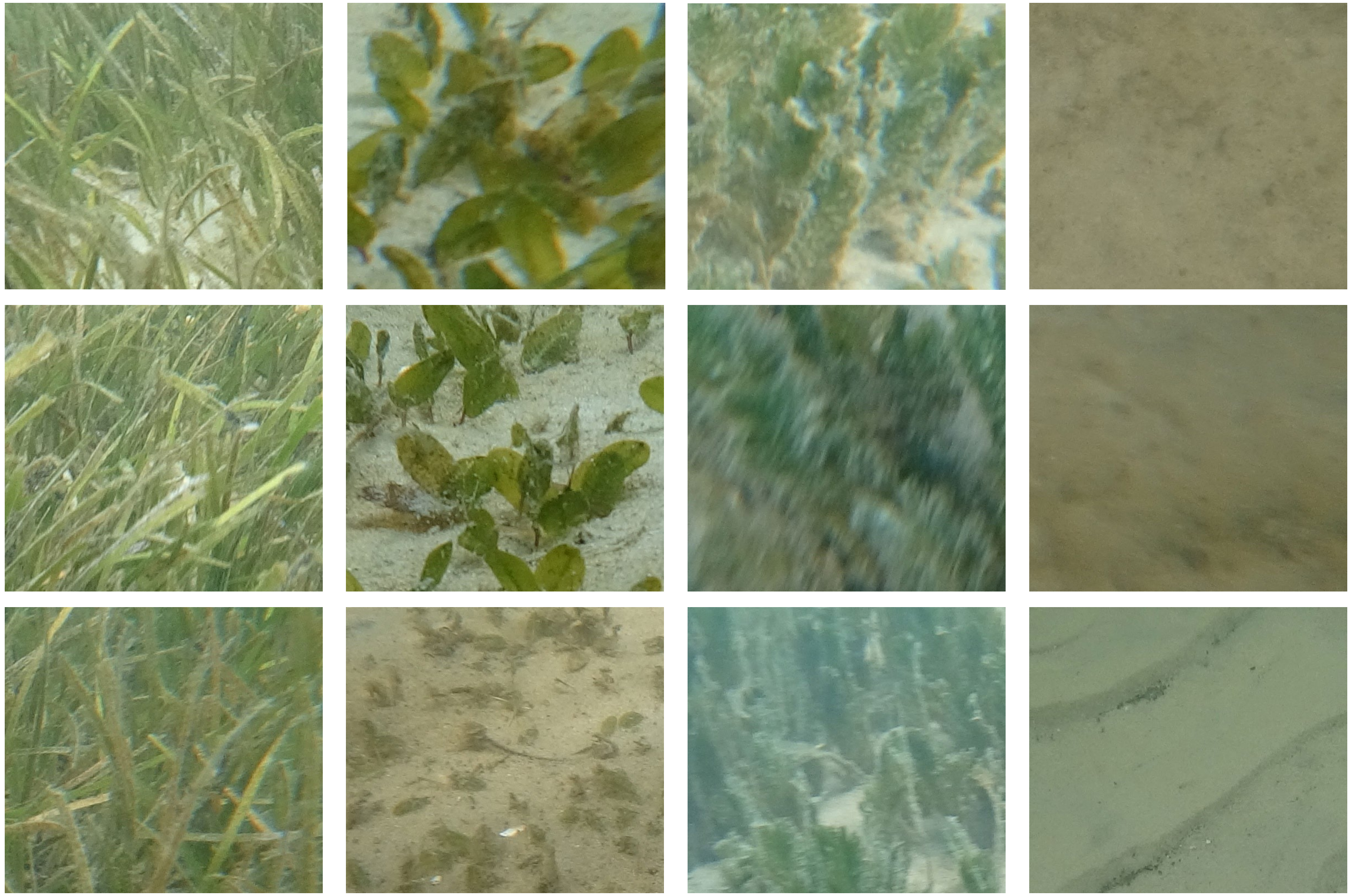
- Semi-automated dataset collection methodology transfers effort from annotation to collection time
- Collect images which contain only one species at a time
- All patches in each single species image can be labelled as the same class
- **66,946** image patches in 4 classes

Assign all patches into image-level class:



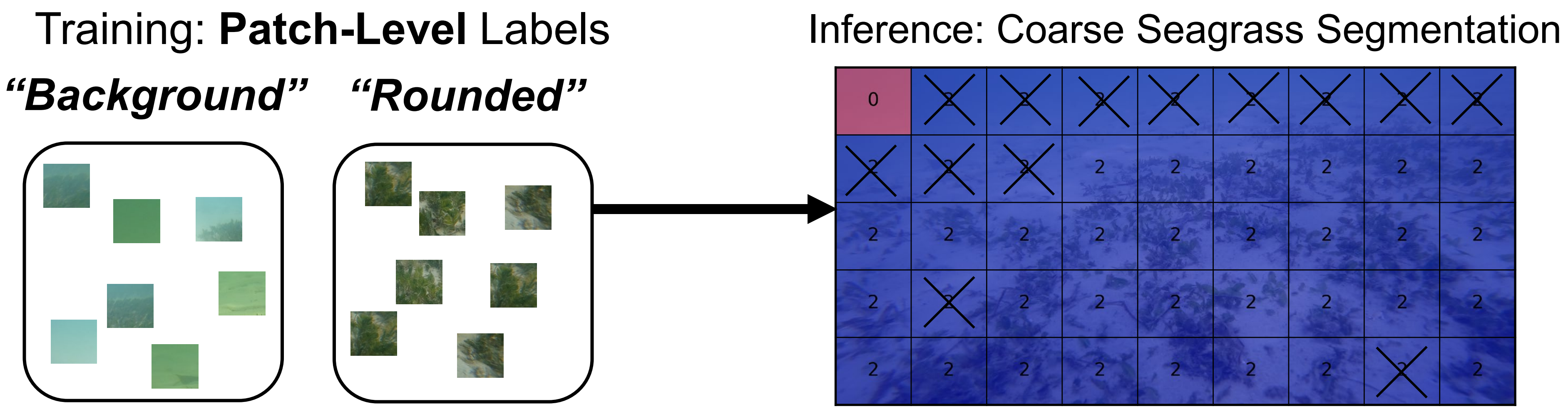
DeepSeagrass

- What features can we use to classify these types of seagrass?
  - Colour?
  - Texture?
  - Edges?
  - Deep features?

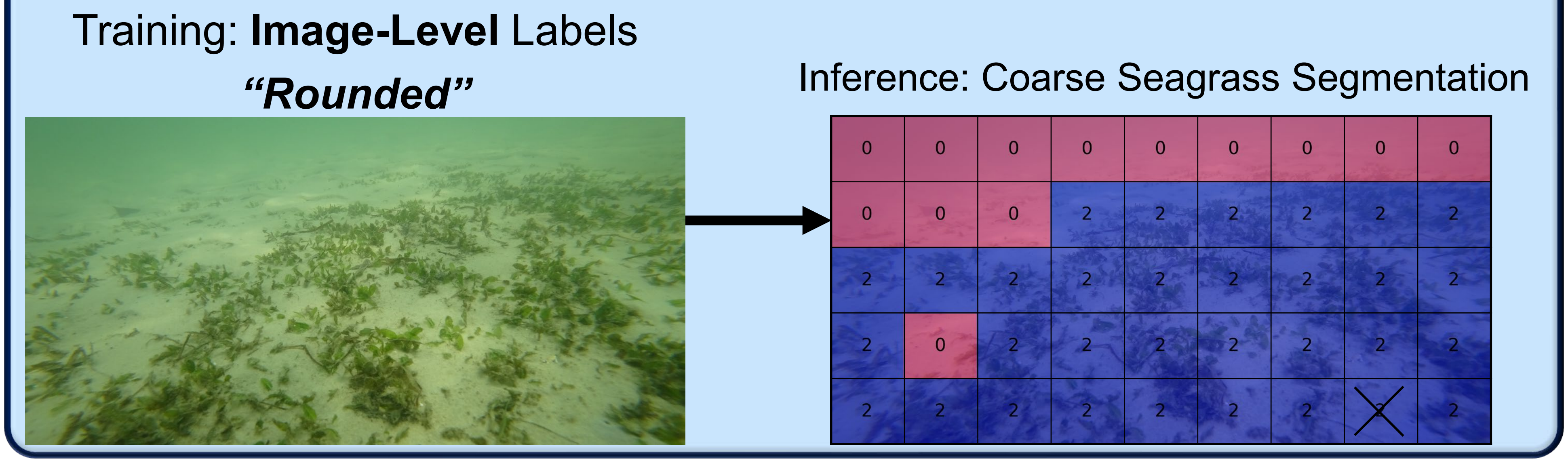


Strappy Rounded Ferny Background

Prior Approaches



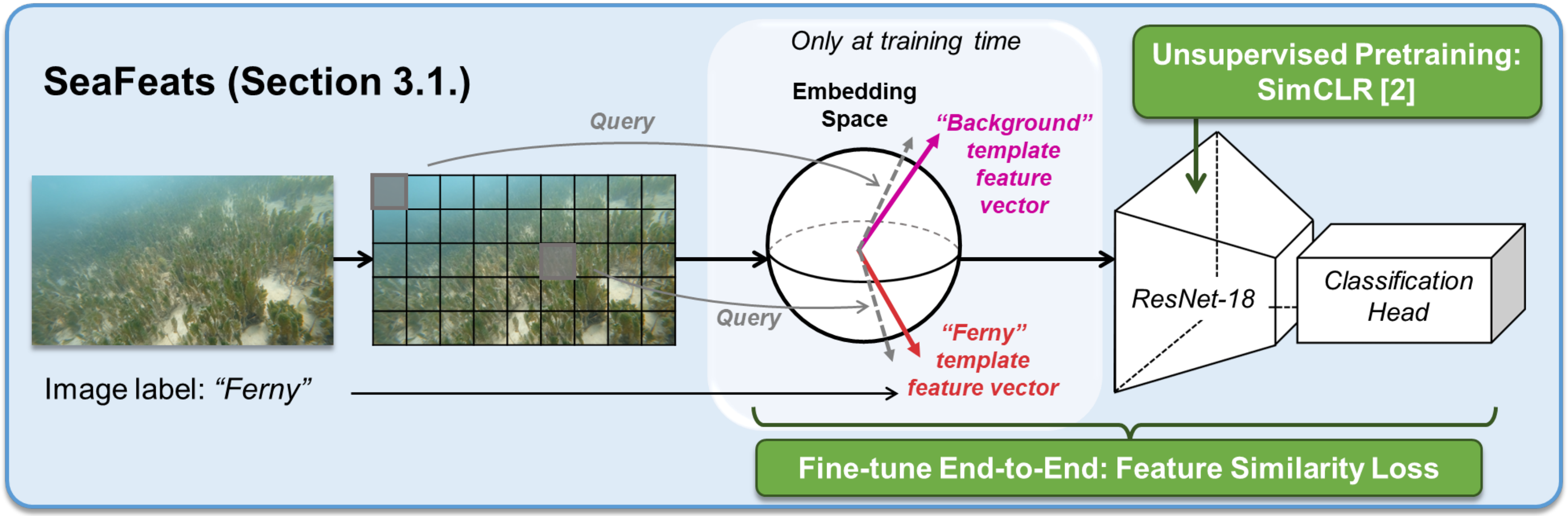
Our Approach



Reduces labels required by up to 96%

Legend: 0 Background 1 Ferny 2 Rounded 3 Strappy ╳ Incorrect

Our Framework



Every epoch:

$$\bar{\mathbf{v}}_c = \underbrace{\frac{1}{N_c} \sum_x \mathbb{1}_{[x=c]} \frac{\mathbf{v}_x}{||\mathbf{v}_x||_2}}_{\text{Average the feature vectors } v_x \text{ from a sample of } N_c \text{ patches}}$$

Template feature vector for class c

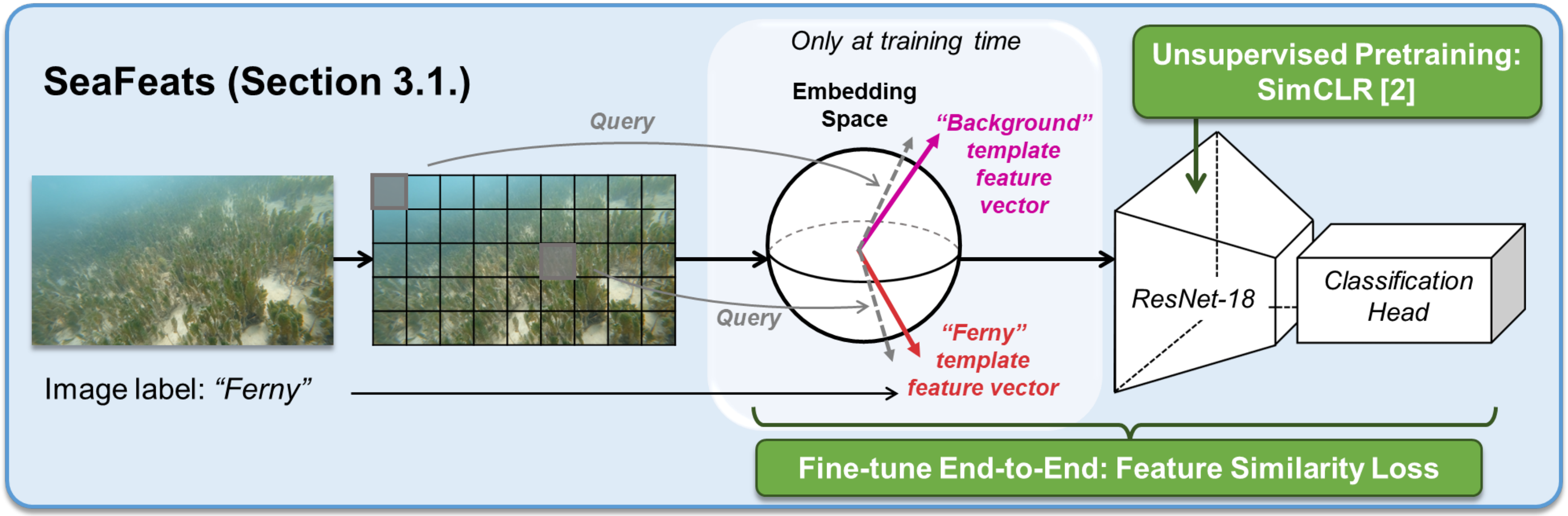
Feature vector $\boldsymbol{v}_x$ from ResNet-18 feature extractor

The class of patch feature $\boldsymbol{v}_x$ must be labelled at the image-level as class c

Average the feature vectors $\boldsymbol{v}_x$ from a sample of N_c patches

SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework



Feature vector $\boldsymbol{v}_x$ for current query patch

$$\mathrm{sim}(\mathbf{v}_x, \bar{\mathbf{v}}_c) = \frac{\mathbf{v}_x \cdot \bar{\mathbf{v}}_c}{||\mathbf{v}_x|| \; ||\bar{\mathbf{v}}_c||}$$

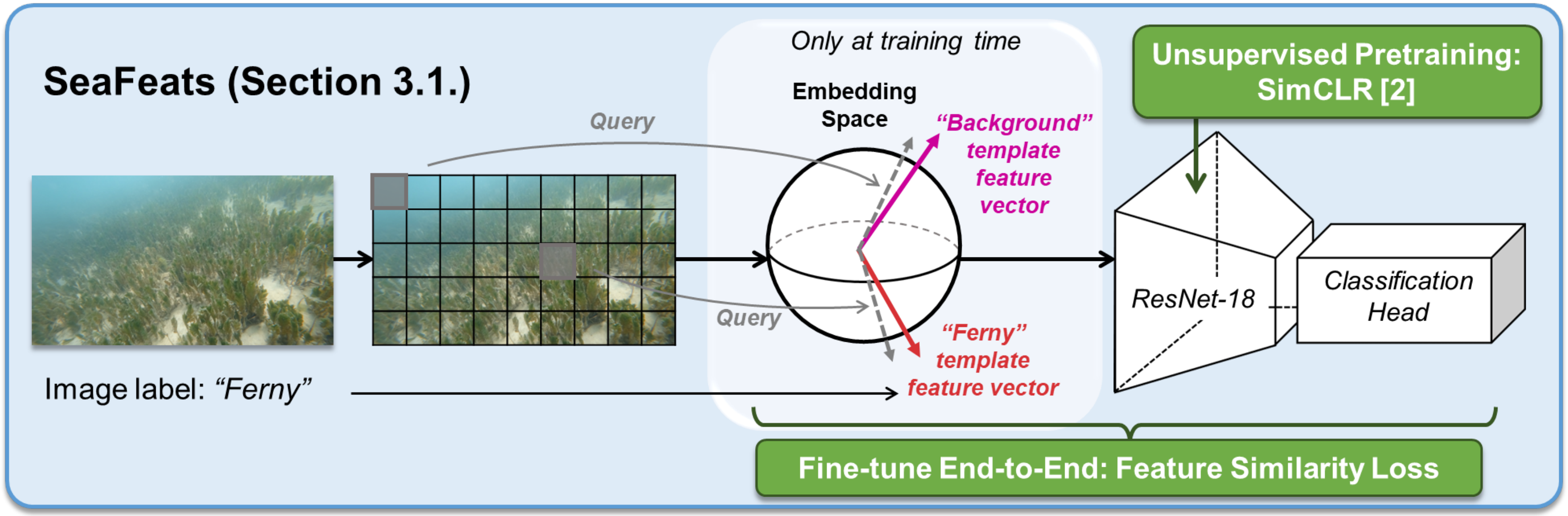
Template feature vector $\overline{\boldsymbol{v}}_c$ for class c

$$p_x = \begin{cases} 0 & \mathrm{sim}(\mathbf{v}_x, \bar{\mathbf{v}}_0) > \mathrm{sim}(\mathbf{v}_x, \bar{\mathbf{v}}_c) \\ c & \text{otherwise} \end{cases}$$

Pseudo-label for patch x

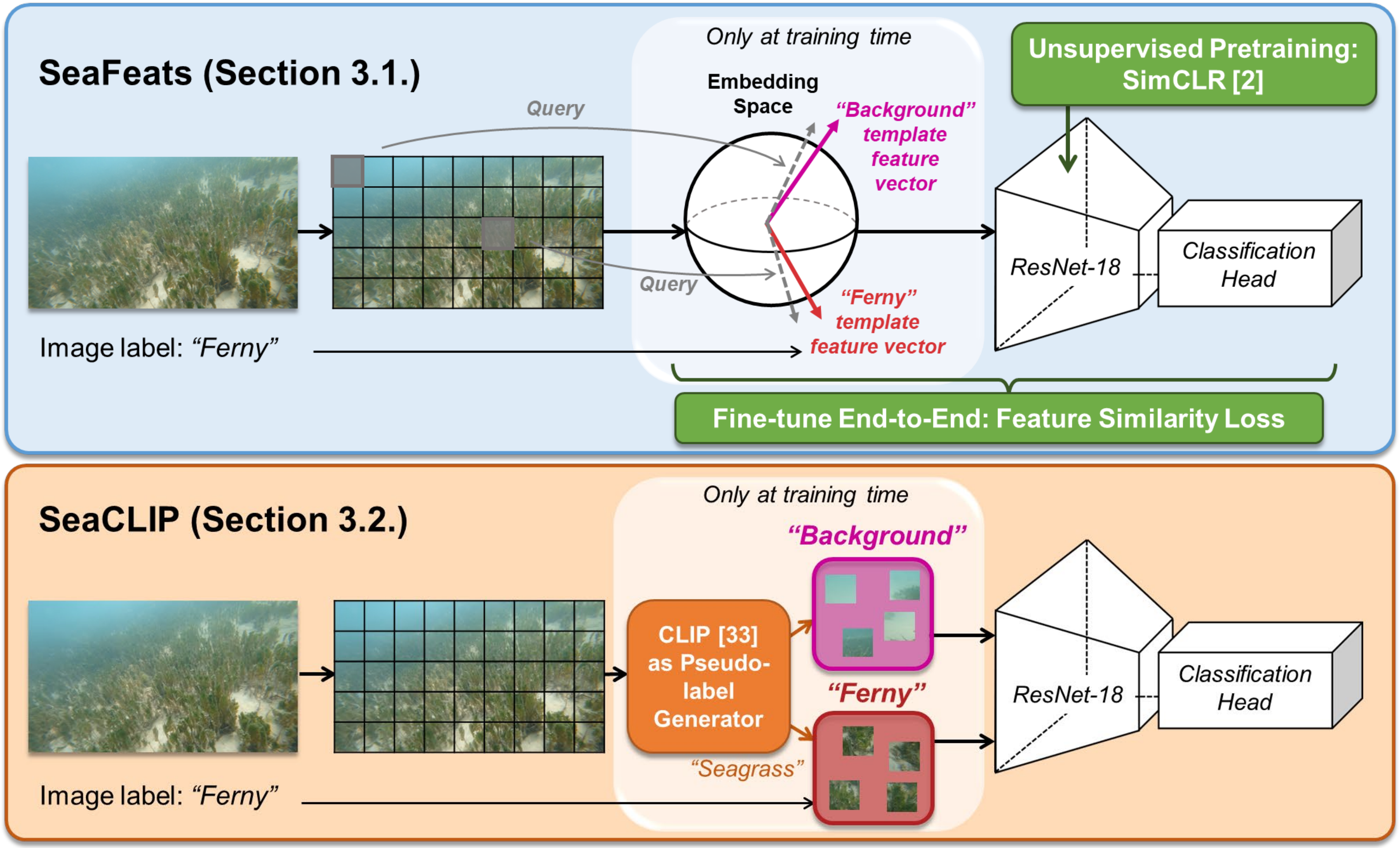
SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework



SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework



SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework

Patch

[taken from image labelled as “Ferny”]

+

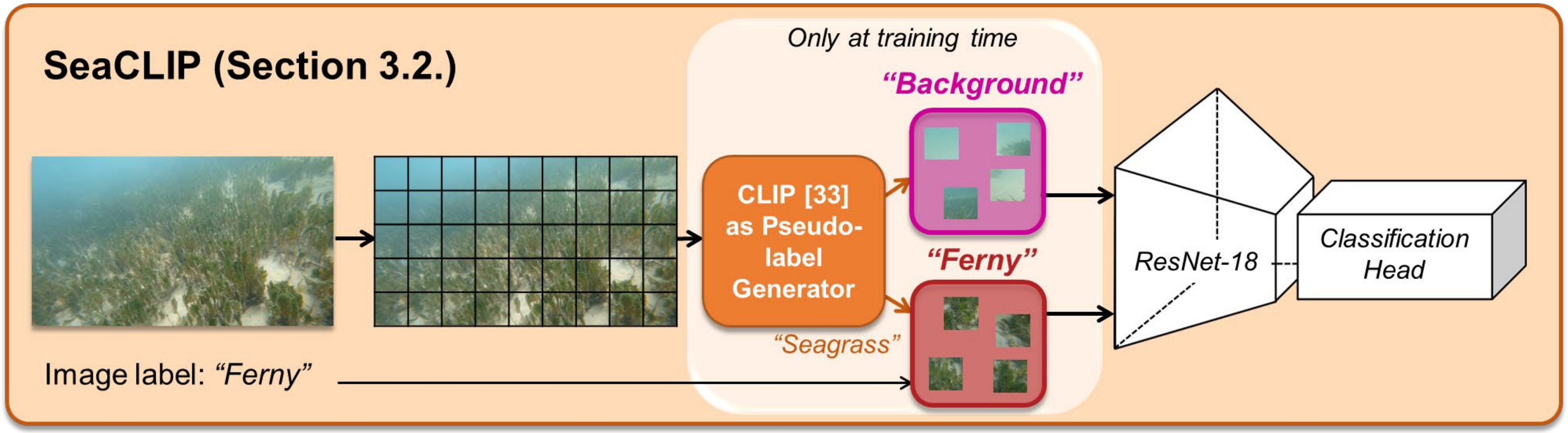
“image containing seagrass”
“image of sand”
“image of water”

Natural language prompts

→ CLIP →

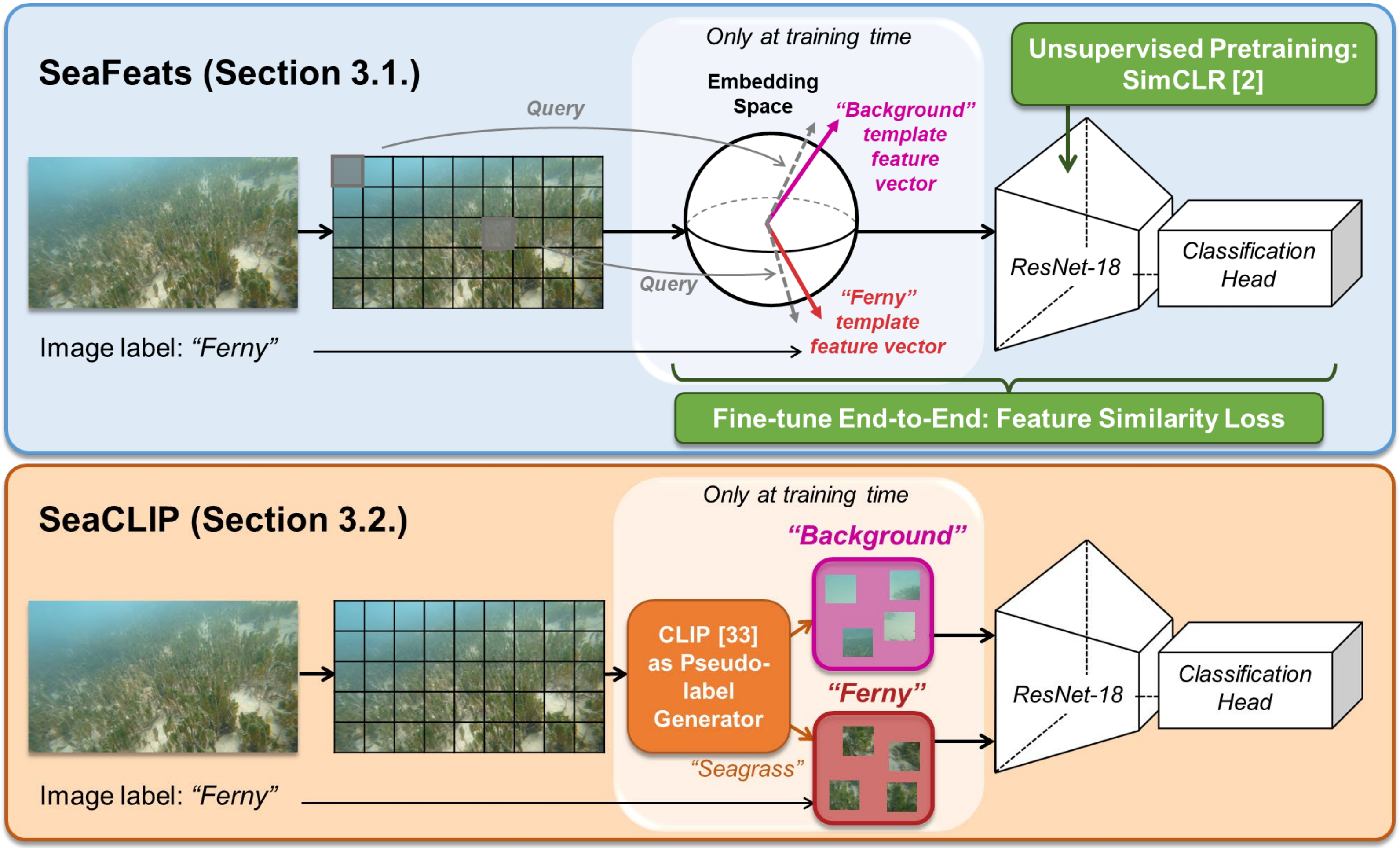
0.8 → *“image containing seagrass”* → Pseudo-label patch as “Ferny”
0.1
0.1

CLIP probabilities that each prompt matches the image



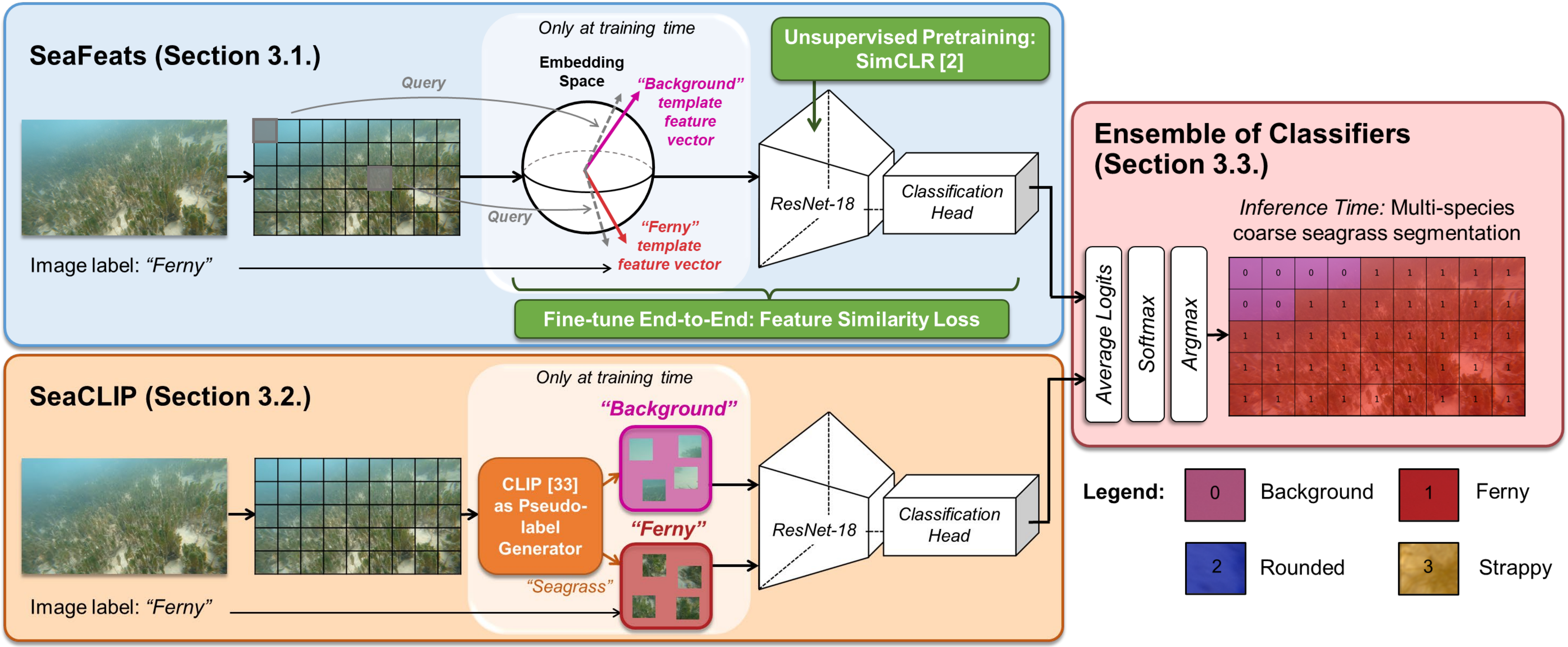
SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework



SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

Our Framework

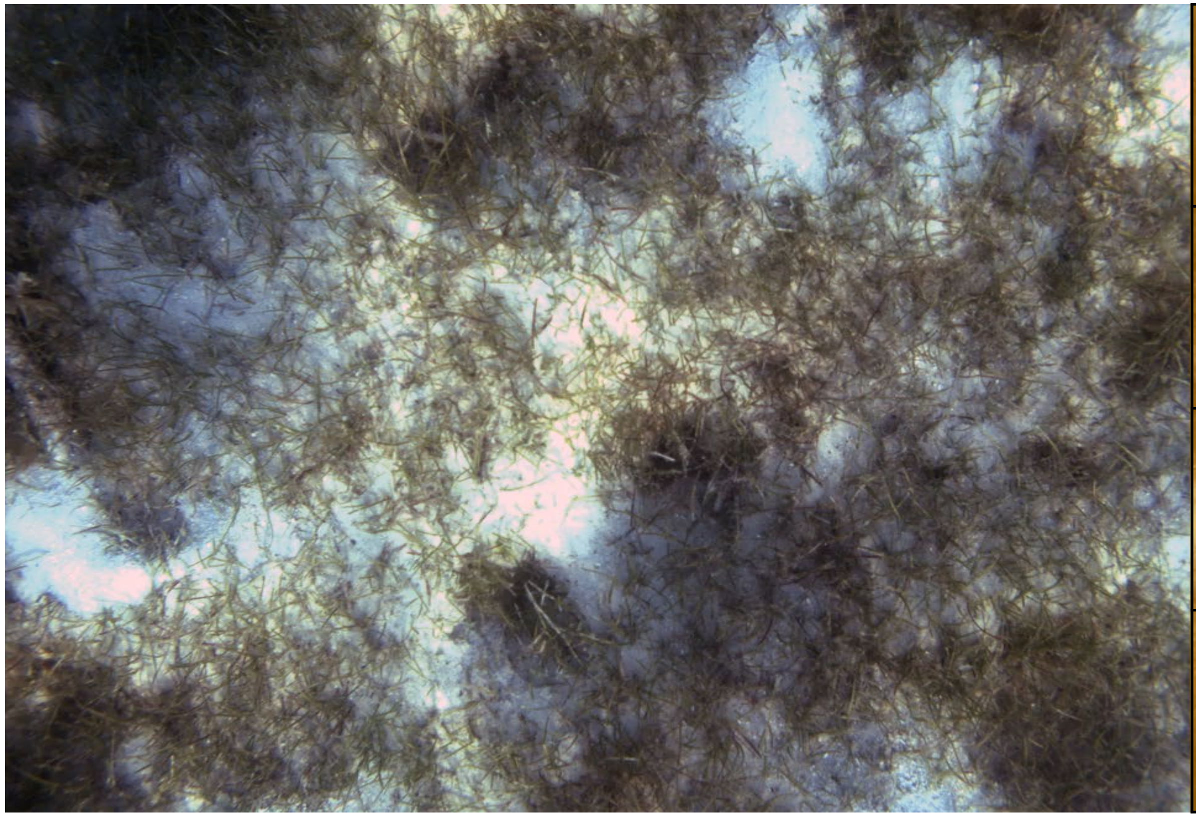


SimCLR [2]: Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Int. Conf. Mach. Learn., pages 1597–1607, 2020.
CLIP [33]: Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In Int. Conf. Mach. Learn., pages 8748–8763, 2021.

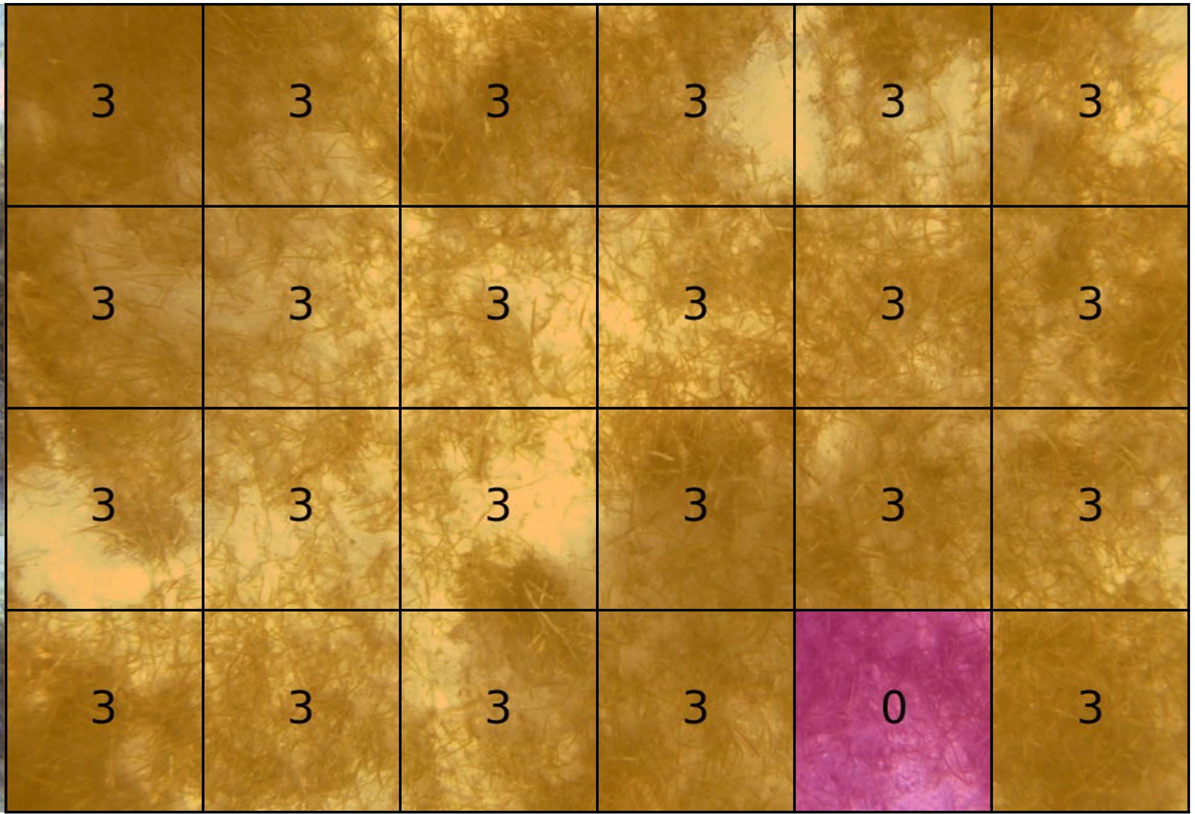
Inference on Imagery Collected by the 'FloatyBoat' Autonomous Surface Vehicle

Image and video source: Reconfigurable Robots for Scaling Reef Restoration [26]

Query Image



Model Inference

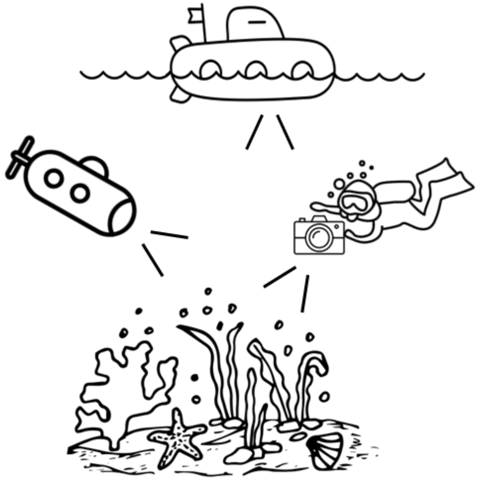


Presentation Outline

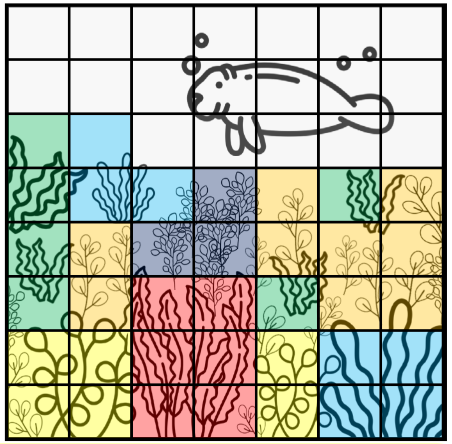
Bodies of Work

Intro

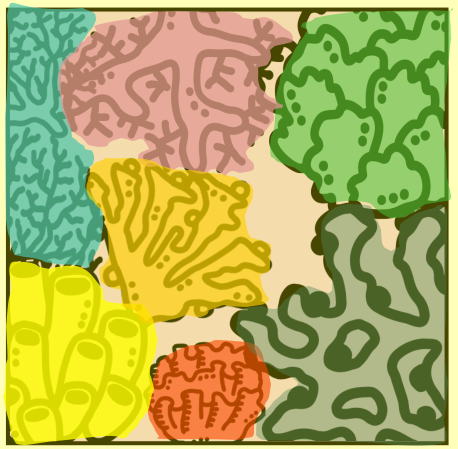
Motivation



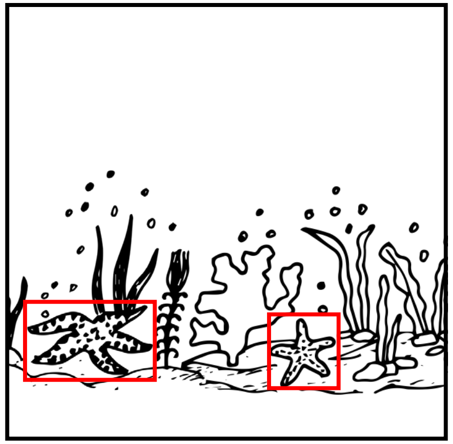
1. Seagrass: Coarse Segmentation



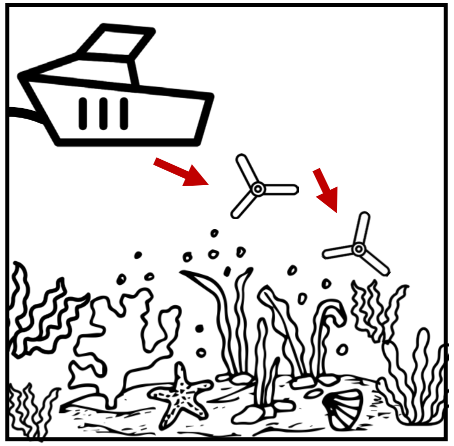
2. Coral: Segmentation



3. Underwater Object Detection



4. Large Scale Reef Restoration (with RRAP)



Future Work

Image taken by G.P. Schmahl/NOAA

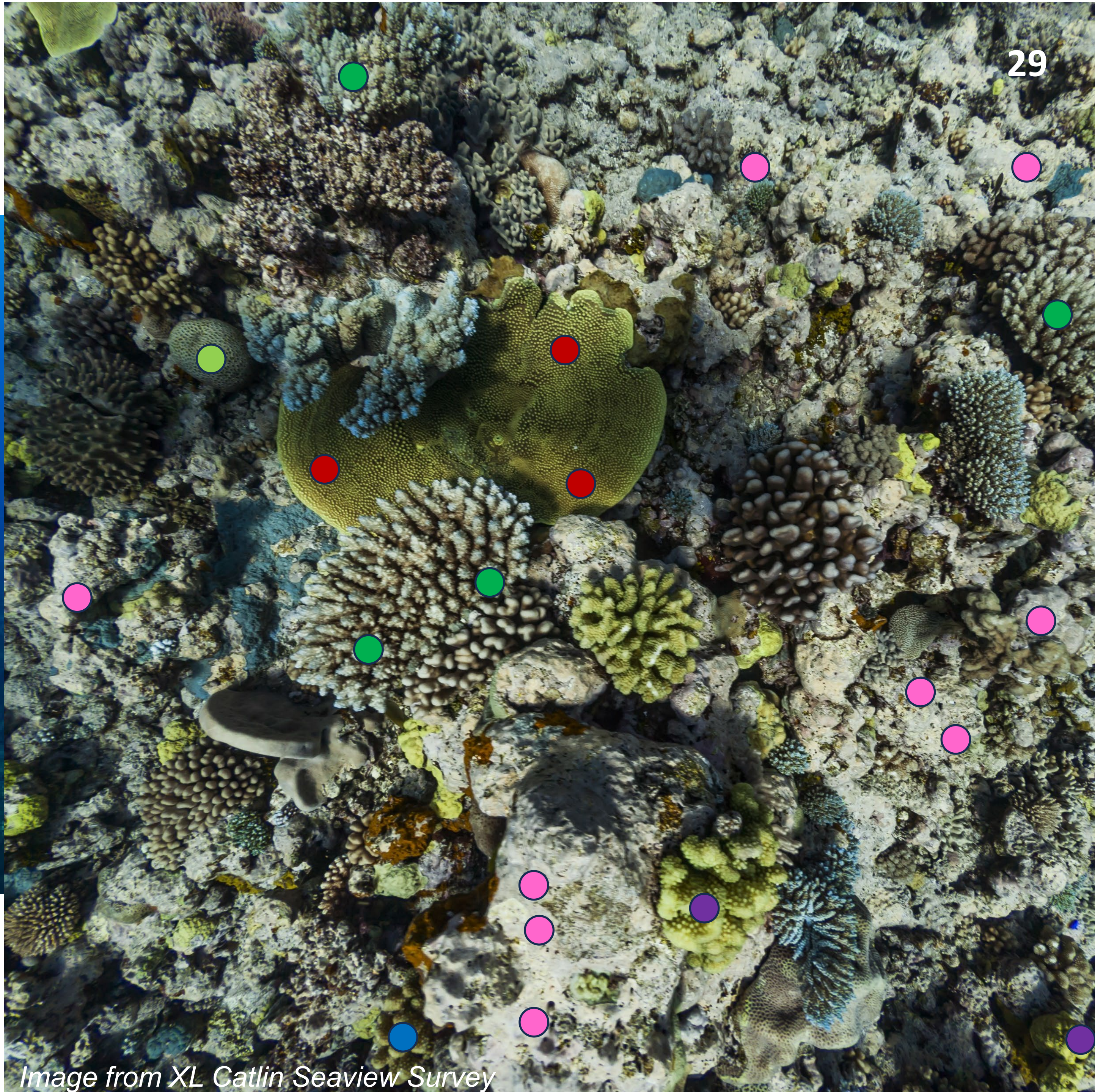
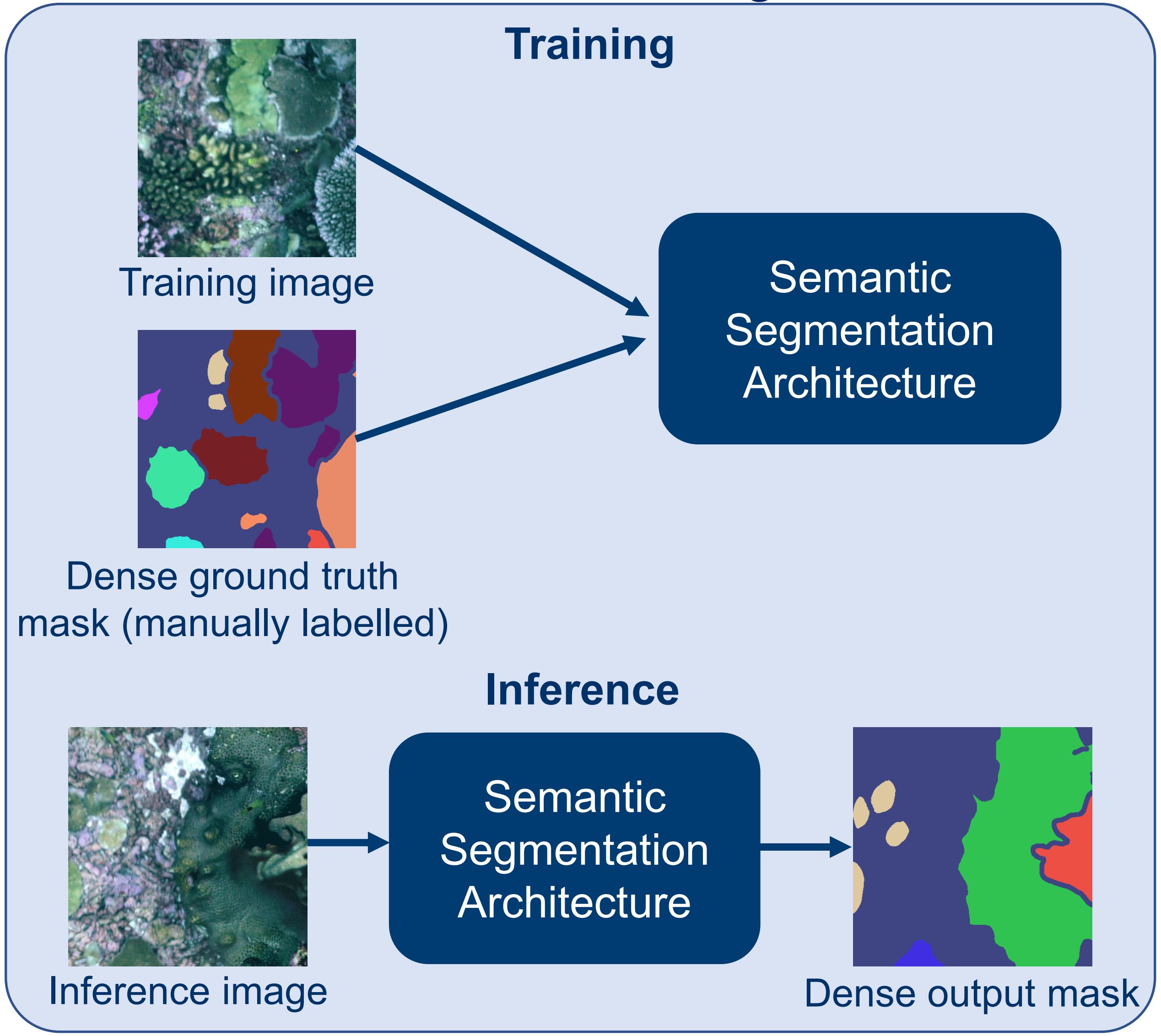
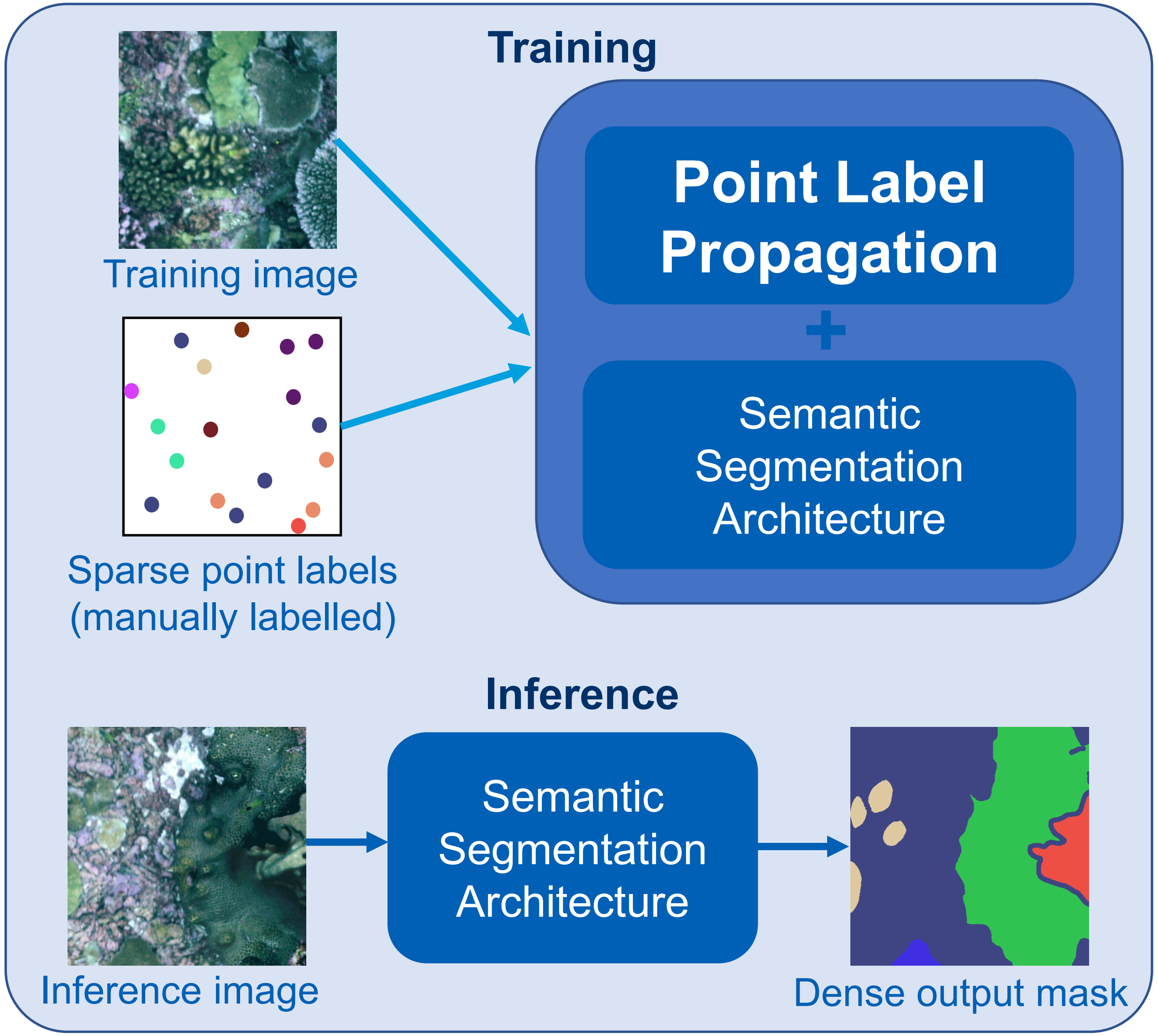


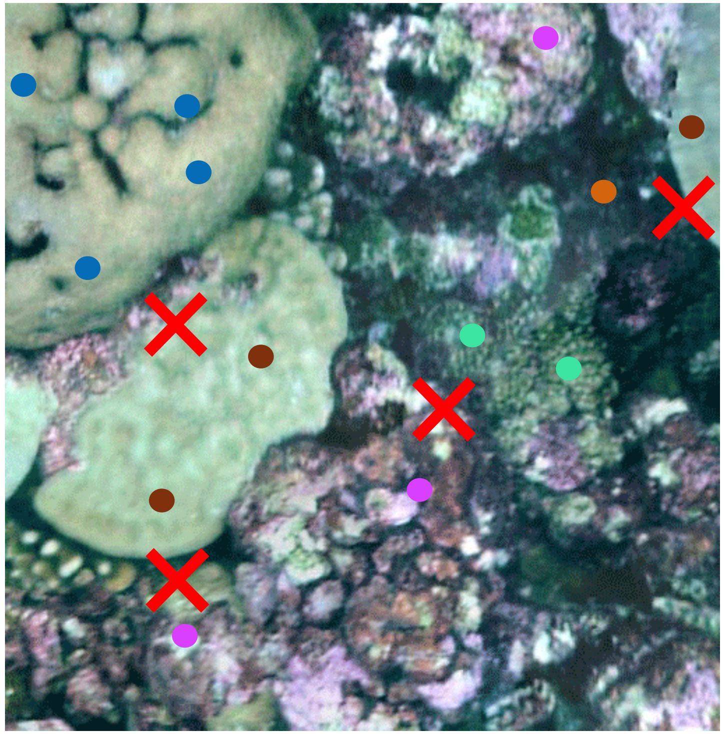
Image from XL Catlin Seaview Survey

Traditional Semantic Segmentation



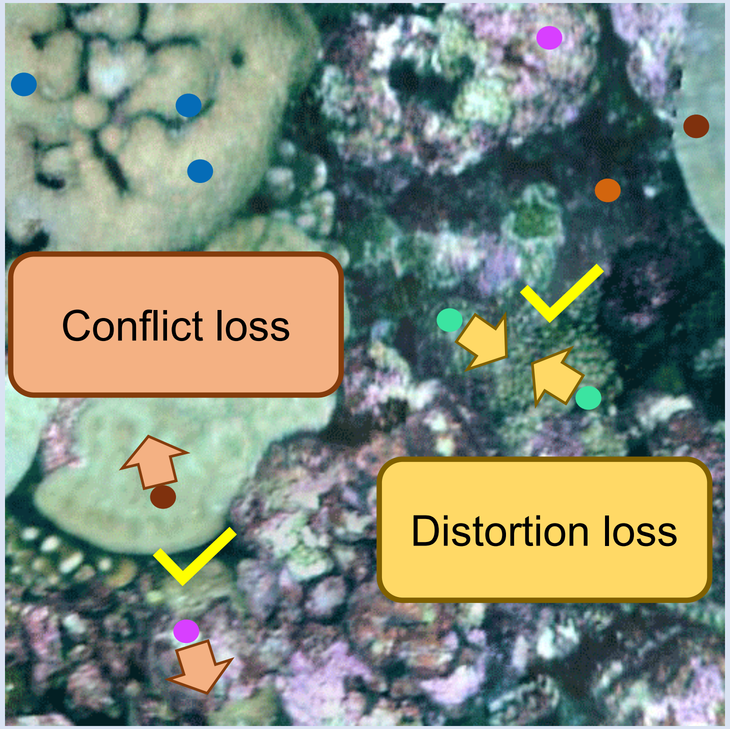
Our Approach





Prior approaches:

- Superpixel regions (shown with) contain conflicting class labels (indicated by ×)
- Superpixels consider only RGB features

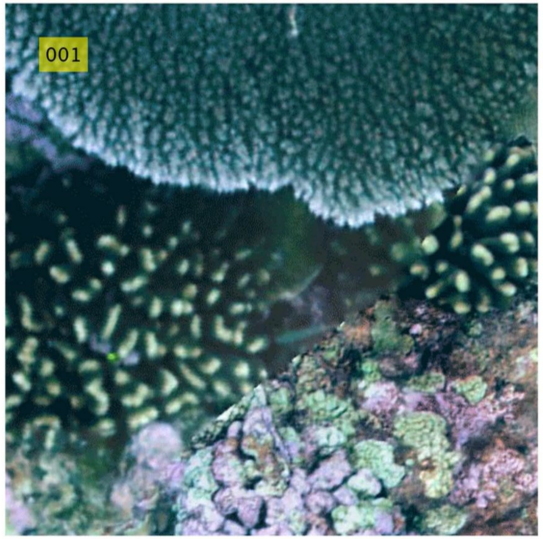


Ours: superpixels informed by points encompass single species regions (✓ denotes consistent inclusion of labels)

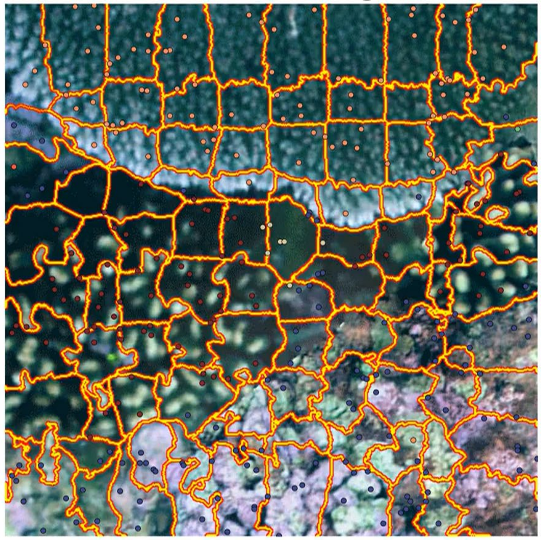
Results – Label Propagation on UCSD Mosaics

- Initially, superpixels are evenly spaced out
- Superpixel centre locations are optimised

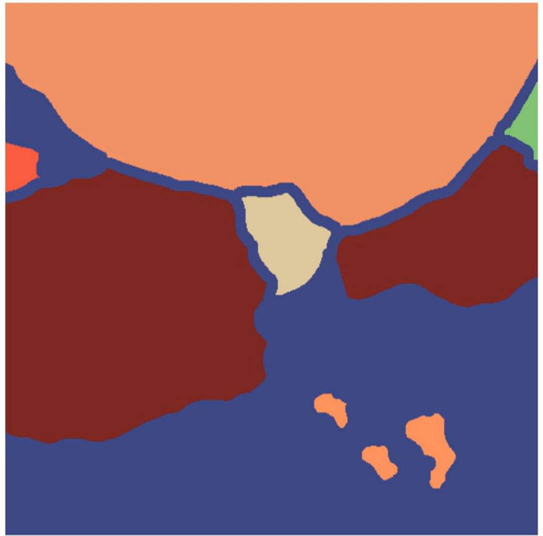
Query Image



Superpixel Regions



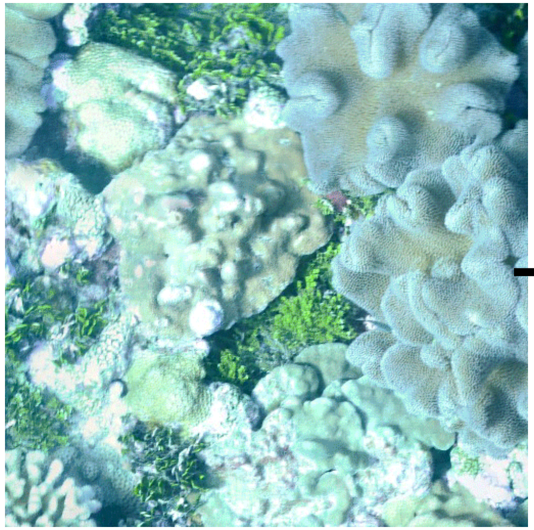
Ground Truth

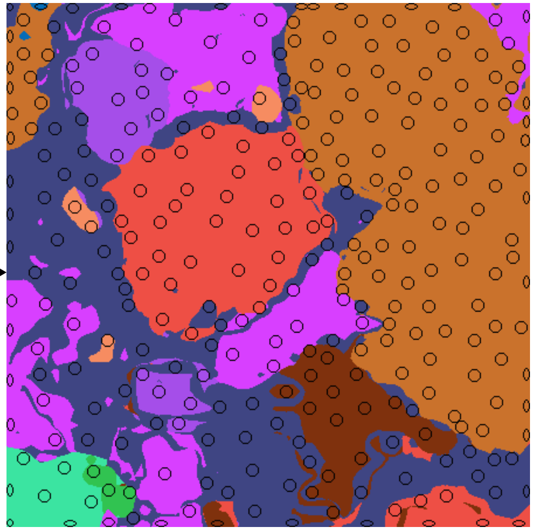


Augmented Ground Truth

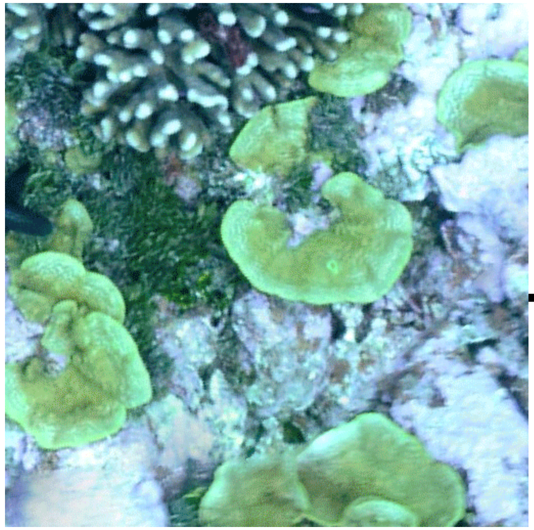
Sparse Label Setting

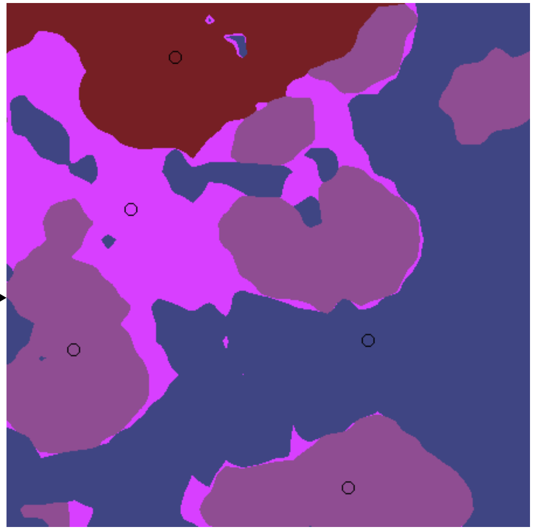
Previously:
100 to 300 point labels per image





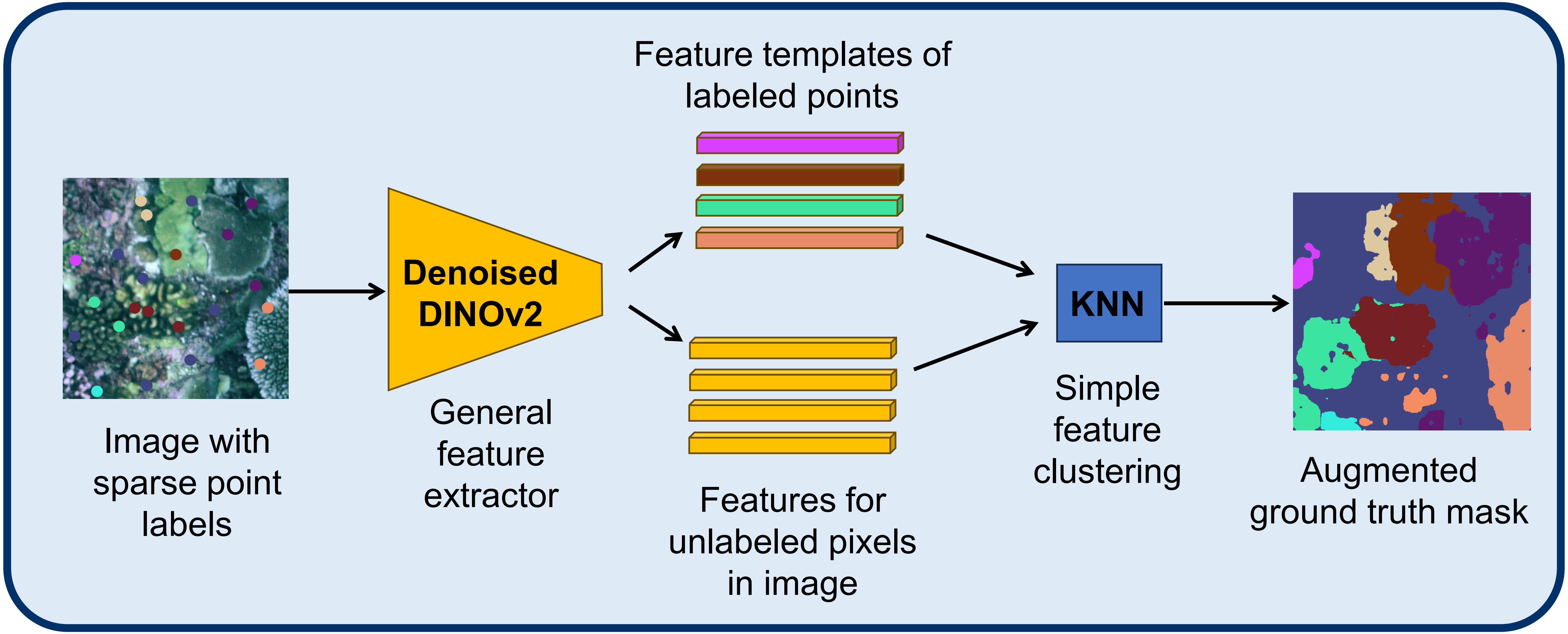
Extremely Sparse Setting:
5 to 25 point labels per image





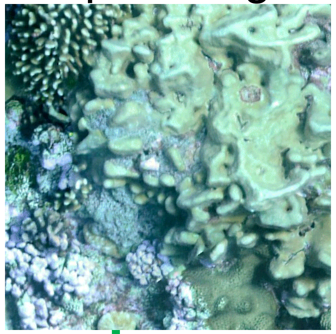
Reduces point labels by up to 98%

Our Approach



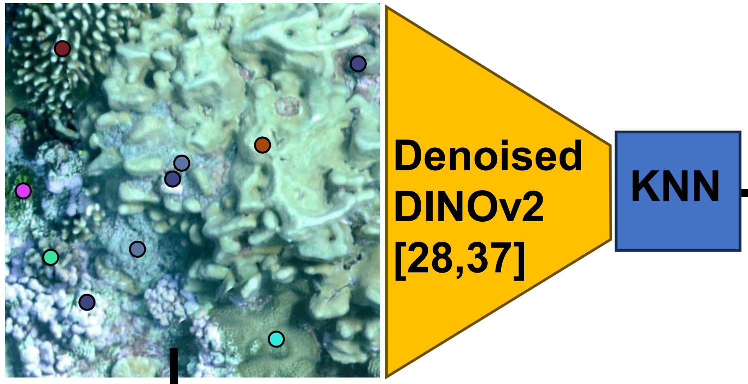
Our Human-in-the-Loop Framework

Input Image

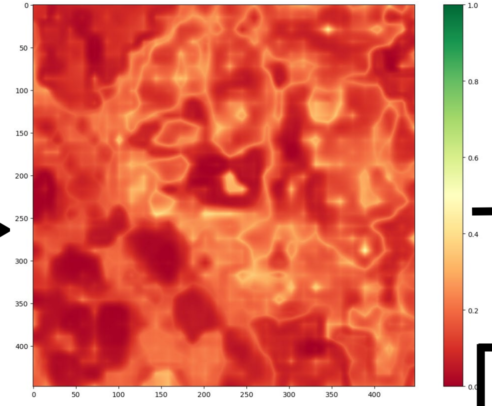


Initially: domain expert labels up to 10 points

Domain Expert Labeller



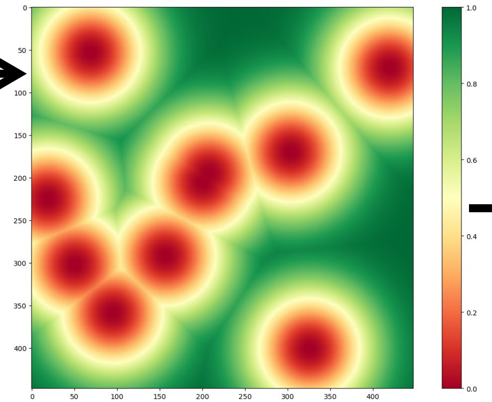
Inverted cosine similarity map to the nearest known labelled pixels:



Create distance map:

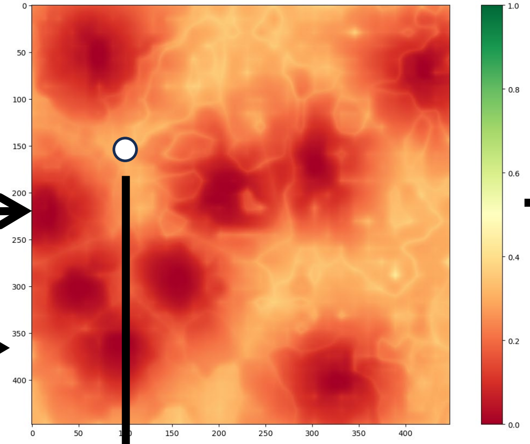
$$\mathrm{D}(x,y) = \min_{(x',y')\in X} \sqrt{(x-x')^2 + (y-y')^2}$$

$$\mathrm{D}_{\mathrm{smooth}}(x,y) = 1 - \exp\left(-\frac{\mathrm{D}(x,y)^2}{2\sigma^2}\right)$$



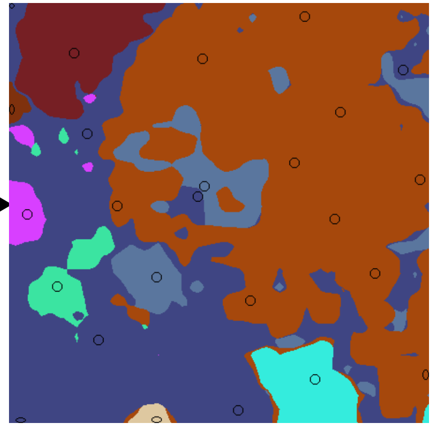
Combine distance and feature similarity:

$$\mathrm{M}(x,y) = \frac{\mathrm{D}_{\mathrm{smooth}}(x,y) + \lambda\,\mathrm{sim}(\mathbf{v}_x, \bar{\mathbf{v}}_l)}{\lambda + 1}$$



Select next point

After max points labelled:



Augmented ground truth mask

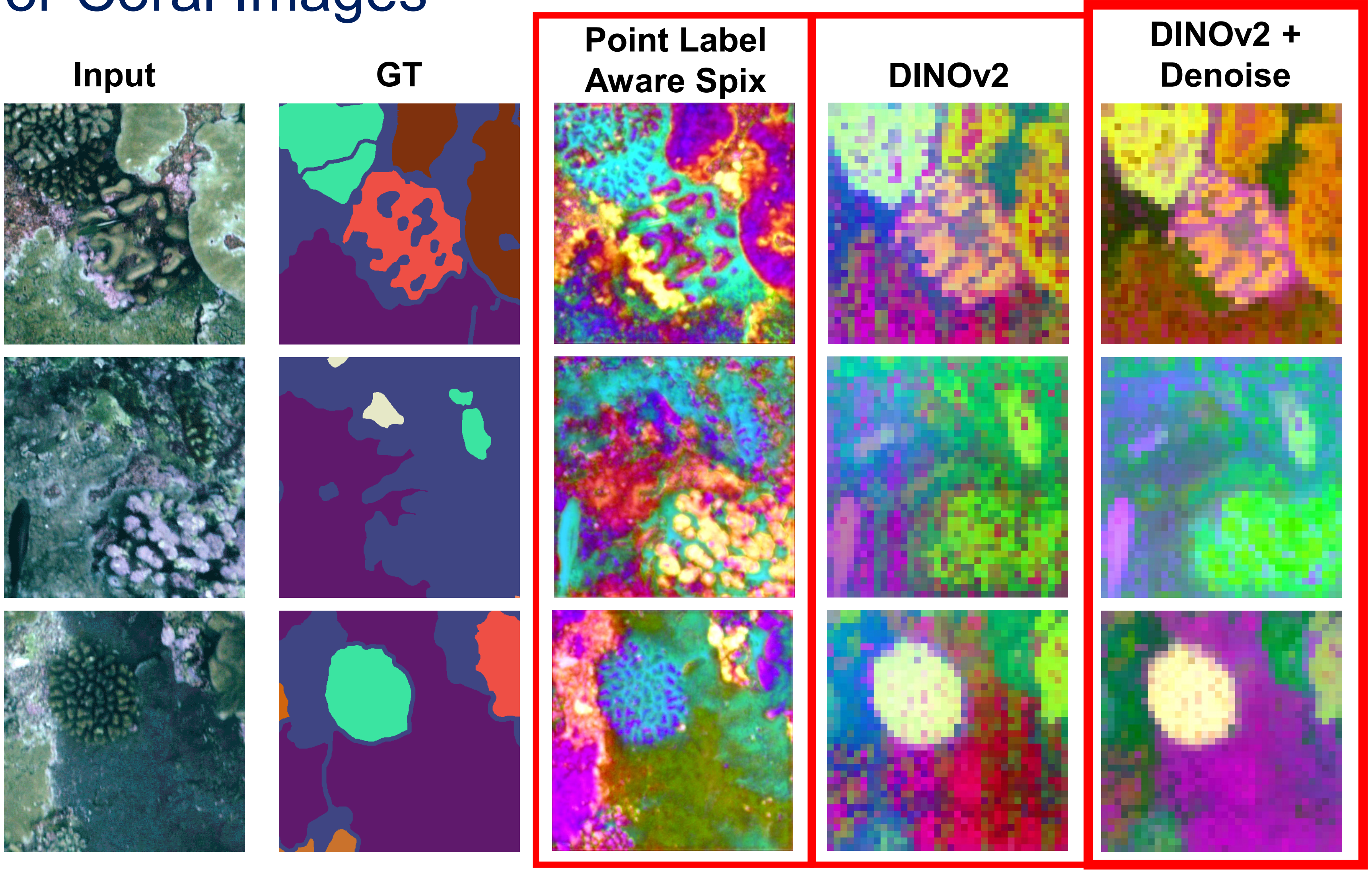
Can be used to train model to perform semantic segmentation

Location of next point is fed back to domain expert

[28] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. DINOv2: Learning robust visual features without supervision. arXiv preprint arXiv:2304.07193, 2023

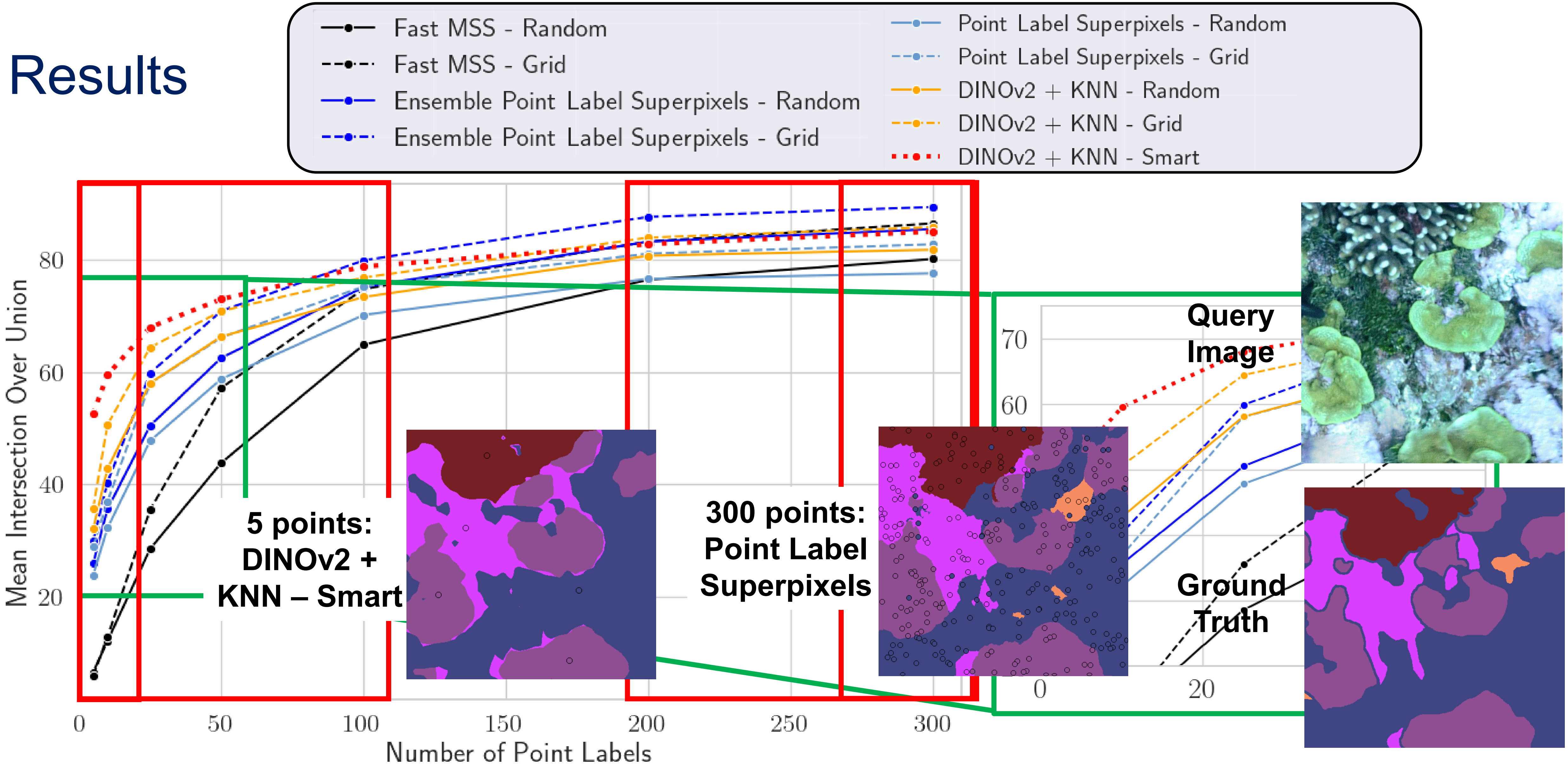
[37] Jiawei Yang, Katie Z Luo, Jiefeng Li, Kilian Q Weinberger, Yonglong Tian, and Yue Wang. Denoising vision transformers. arXiv preprint arXiv:2401.02957, 2024.

Deep Features for Coral Images



Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. DINOv2: Learning robust visual features without supervision. arXiv preprint arXiv:2304.07193, 2023
Jiawei Yang, Katie Z Luo, Jiefeng Li, Kilian Q Weinberger, Yonglong Tian, and Yue Wang. Denoising vision transformers. arXiv preprint arXiv:2401.02957, 2024.

Results

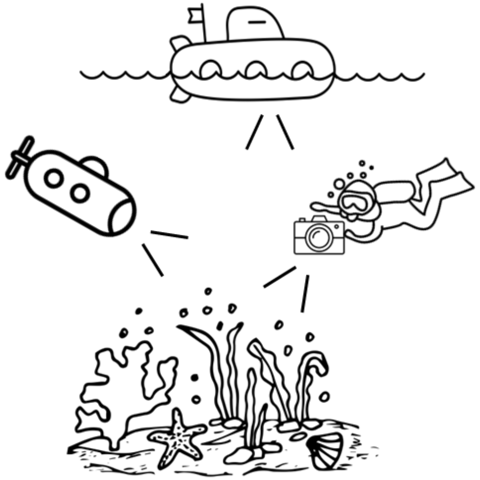

Fast MSS - Random
Fast MSS - Grid
Ensemble Point Label Superpixels - Random
Ensemble Point Label Superpixels - Grid
Point Label Superpixels - Random
Point Label Superpixels - Grid
DINOv2 + KNN - Random
DINOv2 + KNN - Grid
DINOv2 + KNN - Smart
Mean Intersection Over Union
Number of Point Labels
0
50
100
150
200
250
300
20
40
60
80
70
60
0
20
5 points: DINOv2 + KNN – Smart
300 points: Point Label Superpixels
Query Image
Ground Truth

Presentation Outline

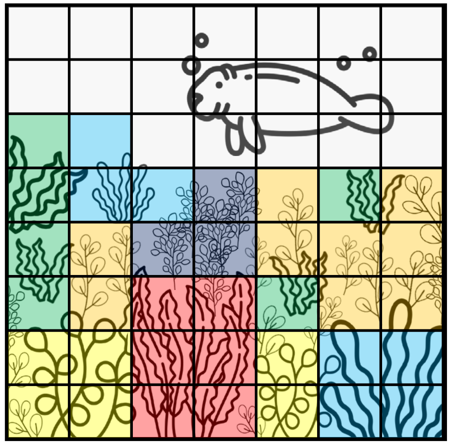
Bodies of Work

Intro

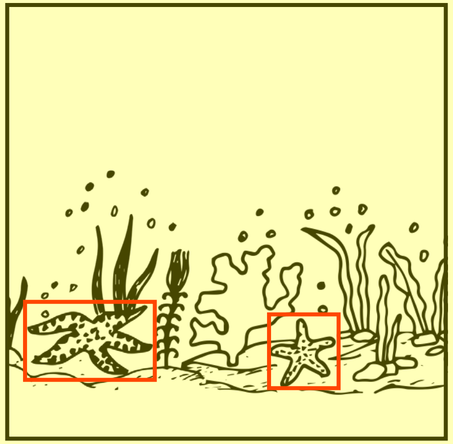
Motivation



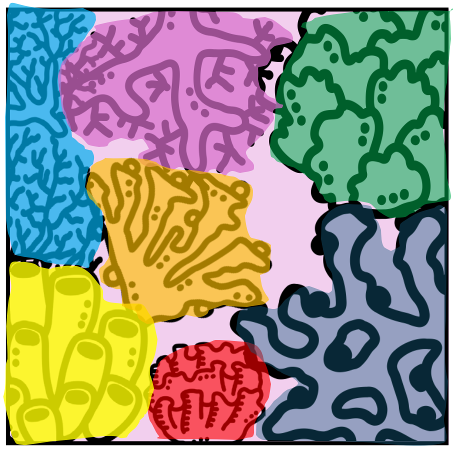
1. Seagrass: Coarse Segmentation



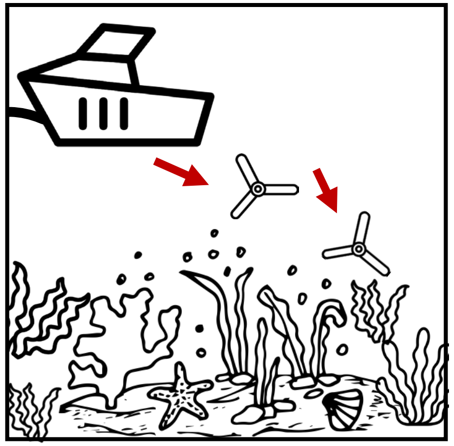
3. Underwater Object Detection



2. Coral: Segmentation



4. Large Scale Reef Restoration (with RRAP)

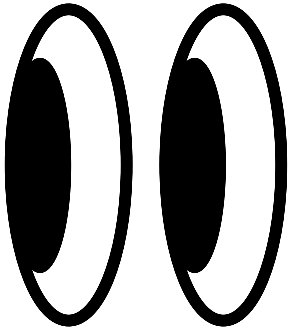


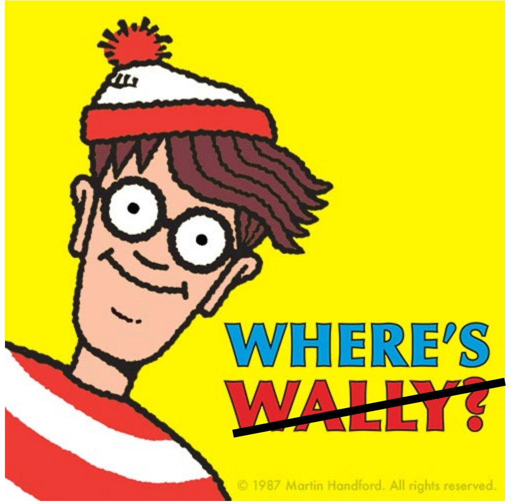
Future Work

Activity: Beat the Machine!

- Crown of Thorns Starfish (COTS) detection: monitoring and mapping outbreaks
- Real imagery from the Great Barrier Reef during a COTS outbreak

Creator: Australian Institute of Marine Science | Credit: LTMP
Copyright: Australian Institute of Marine Science





the Crown-of-Thorns Starfish?

Crown of Thorns Starfish dataset from: Liu, Jiajun, Brano Kusy, Ross Marchant, Brendan Do, Torsten Merz, Joey Crosswell, Andy Steven et al. "The CSIRO Crown-of-Thorns Starfish Detection Dataset." *arXiv preprint arXiv:2111.14311* (2021).

Image 1:

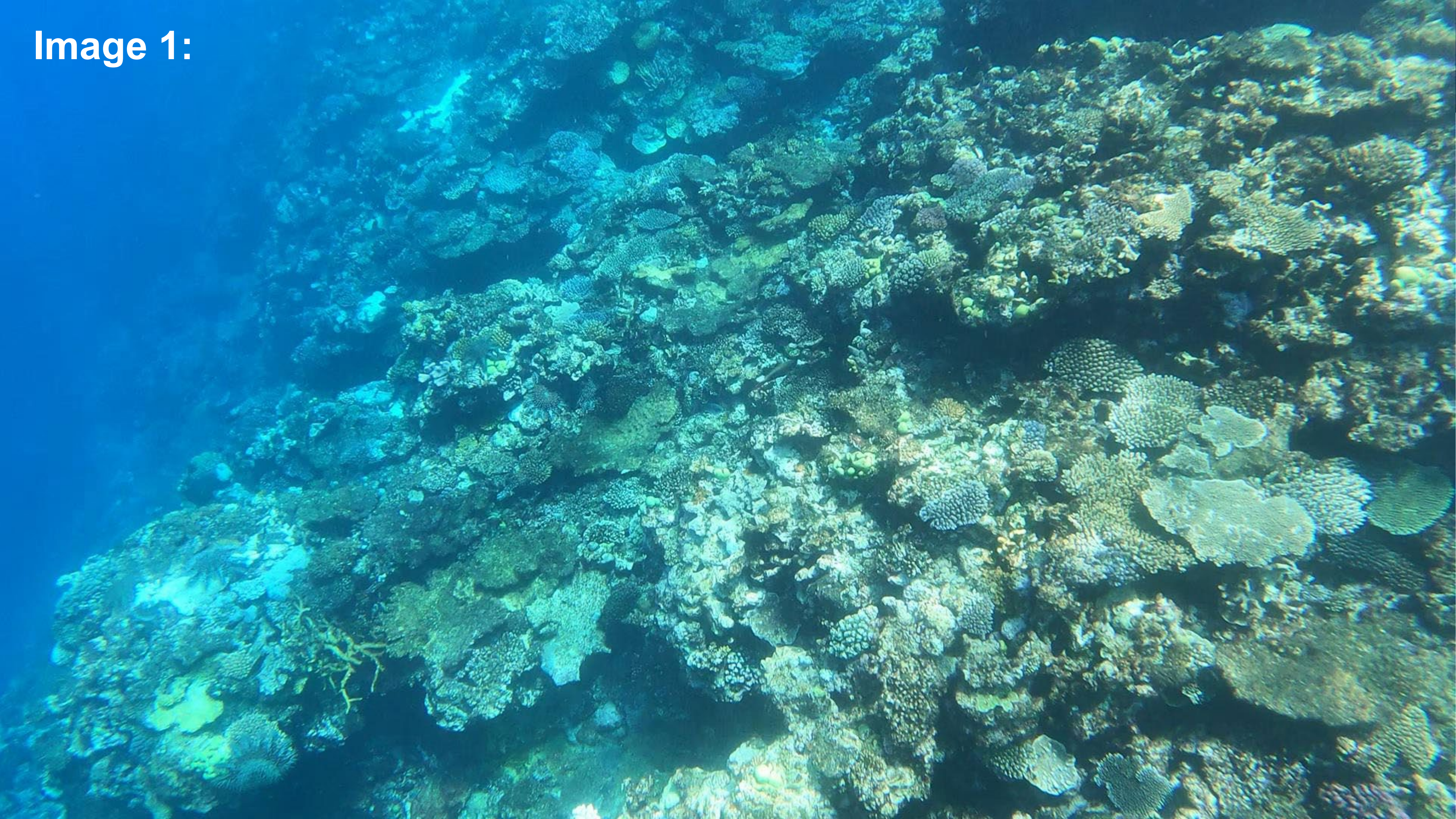
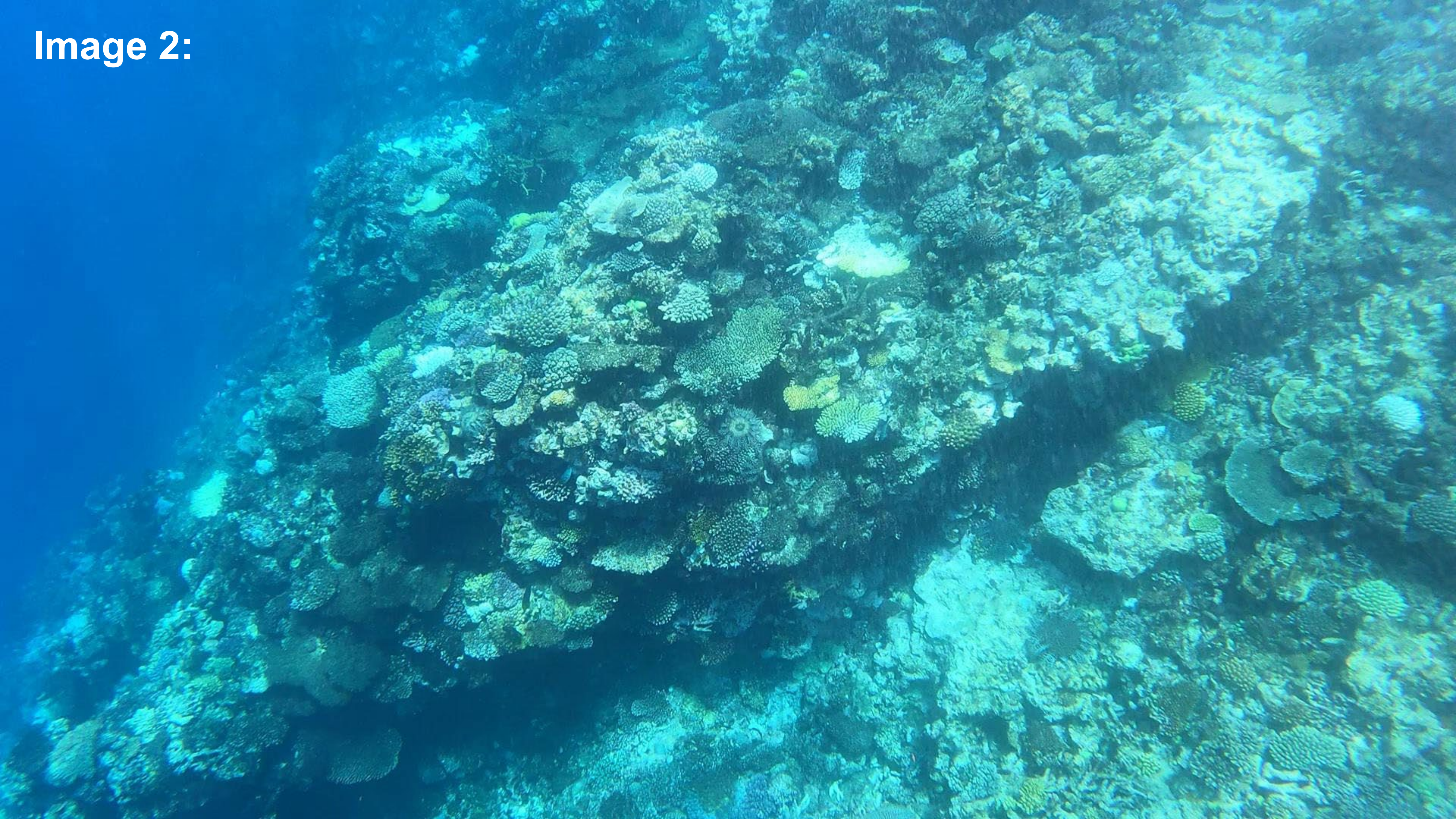


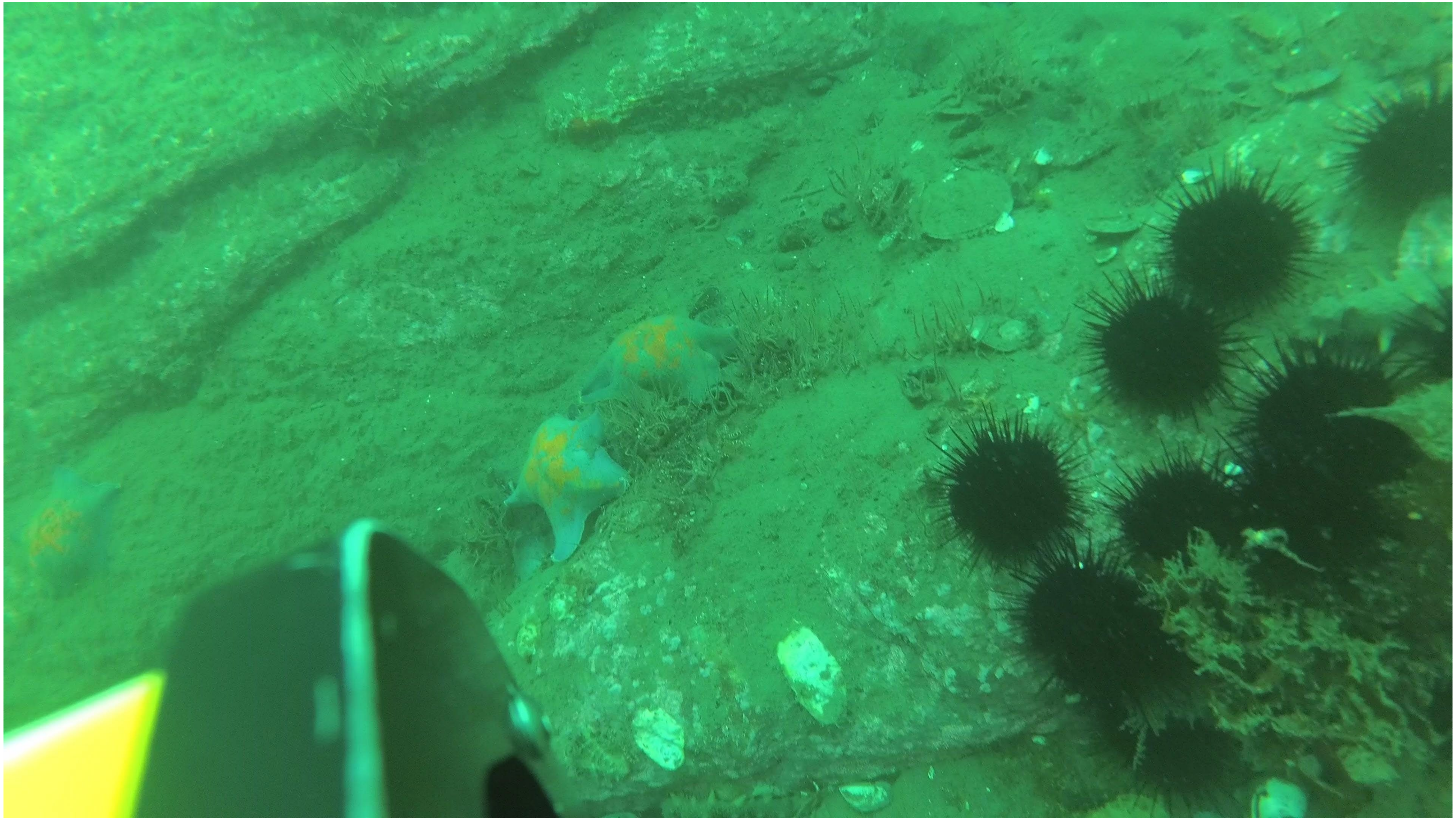
Image 2:



What about in real-time?

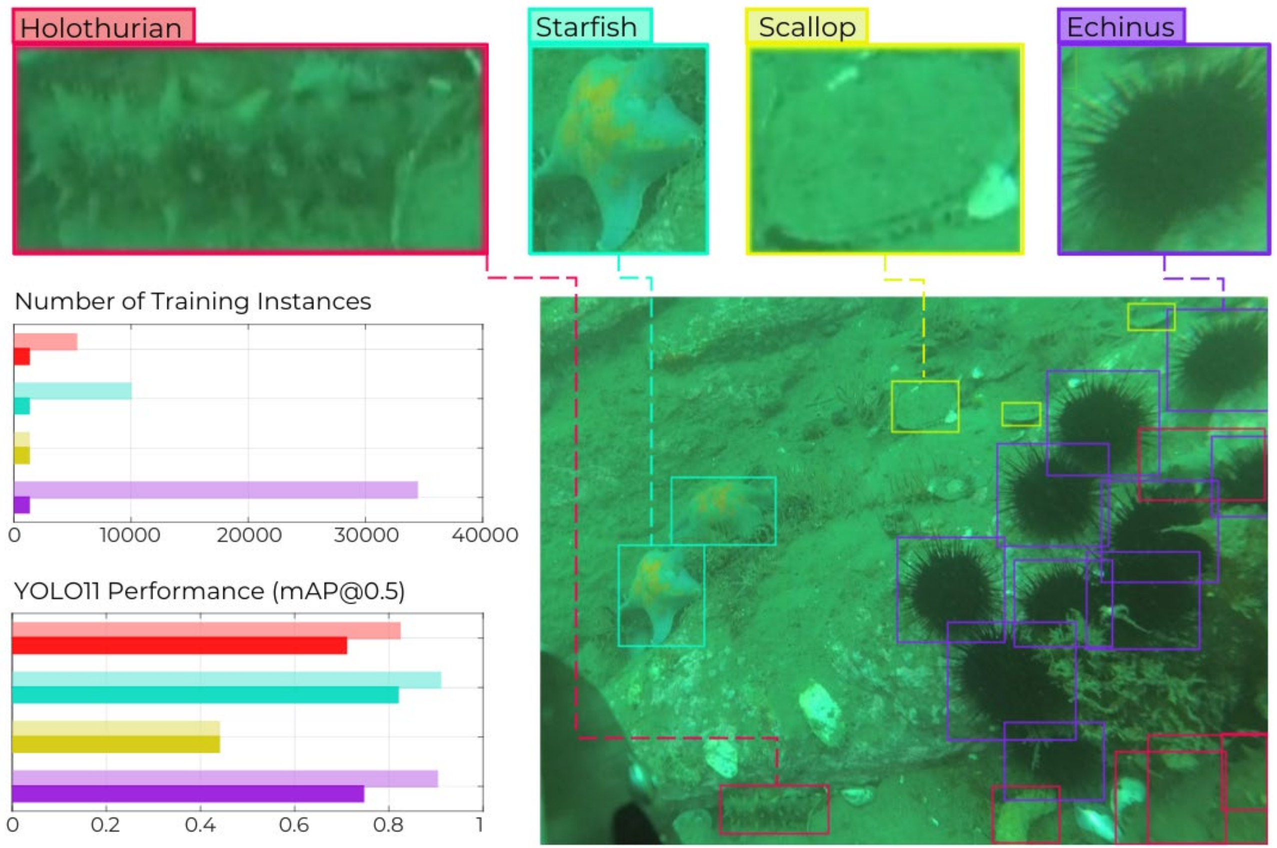
Melanie Wille, Tobias Fischer, ***Scarlett Raine***





Are All Marine Species Created Equal? Performance Disparities in Underwater Object Detection

- Investigating why certain marine species are easier or harder to detect
- Systematic analysis of object detector components and the impact of dataset characteristics

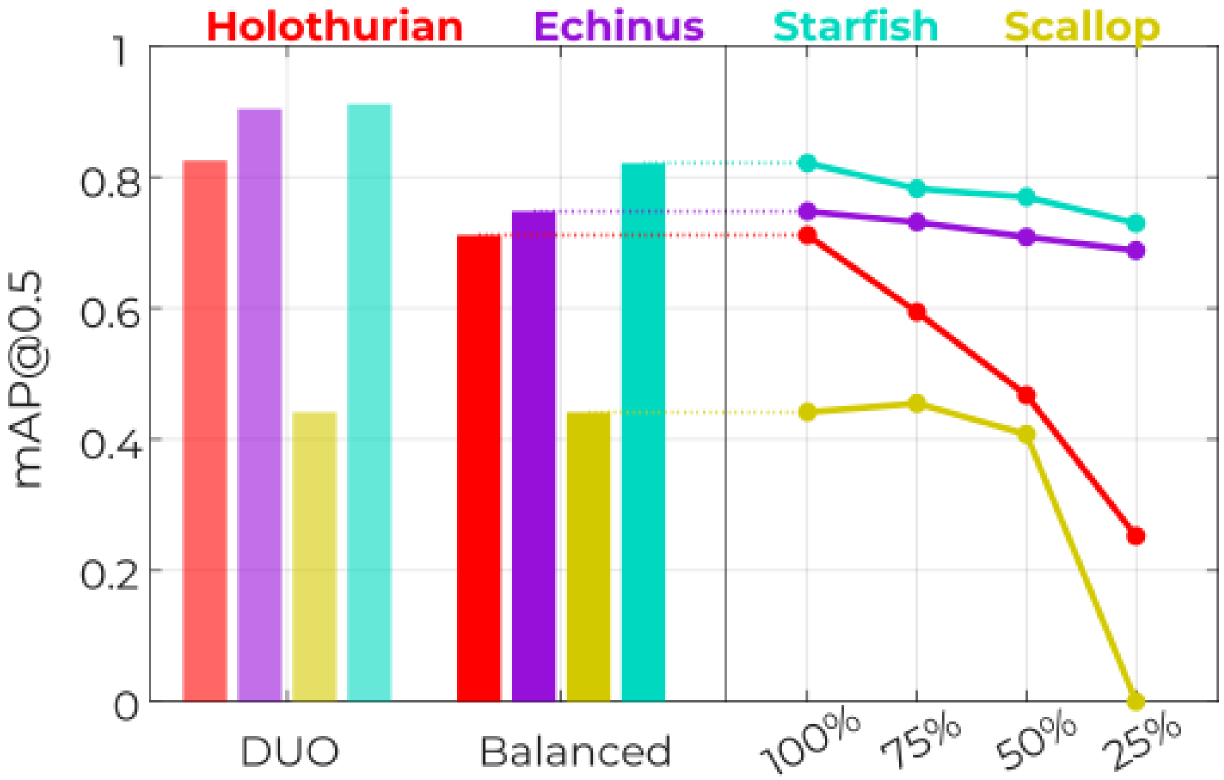


Melanie Wille, Tobias Fischer, ***Scarlett Raine***

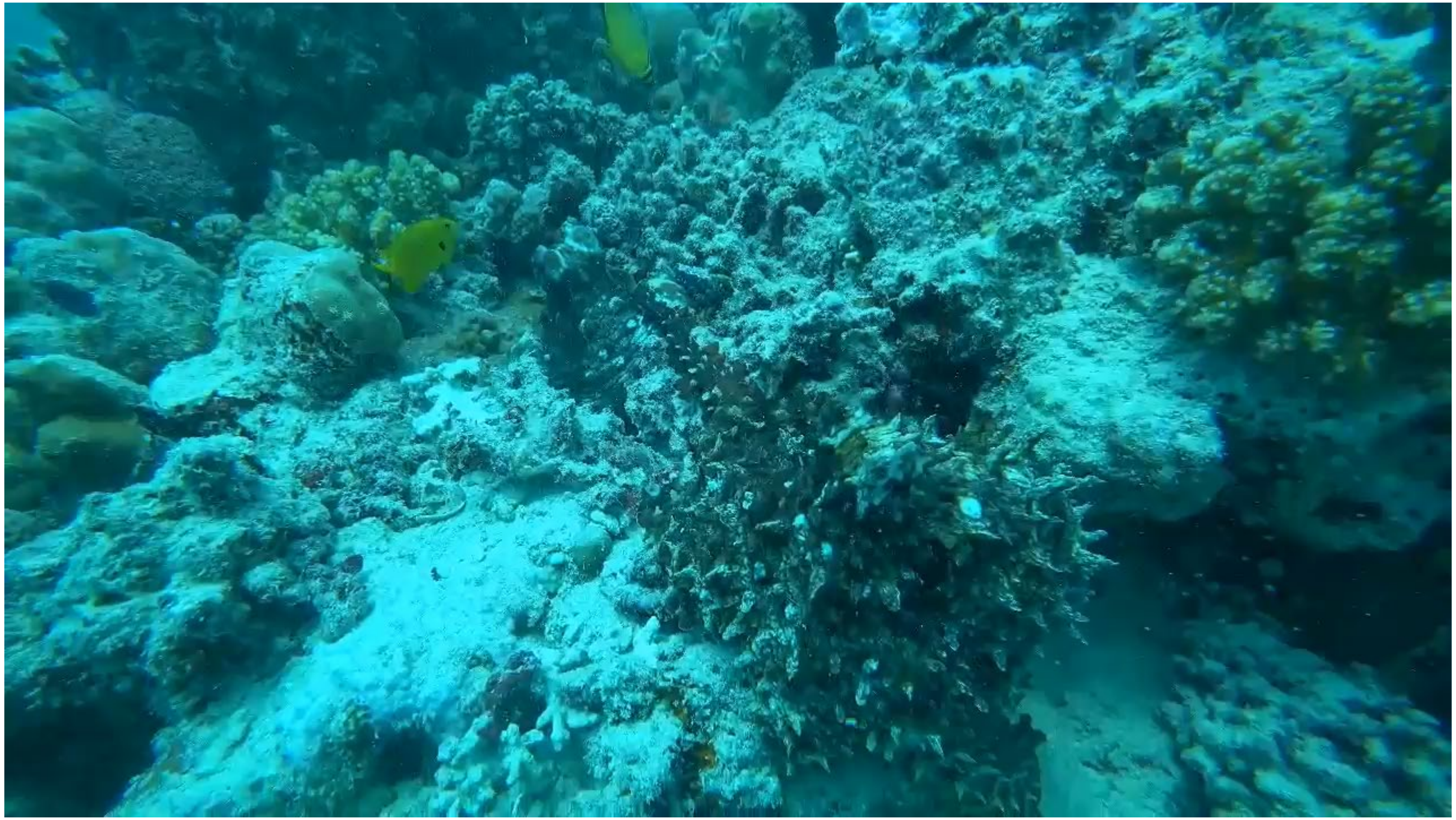


Are All Marine Species Created Equal? Performance Disparities in Underwater Object Detection

- Localization analysis (YOLO11 and TIDE) found foreground-background discrimination is the most problematic, regardless of data quantity
- Classification experiments reveal gaps even with balanced data, suggesting intrinsic feature-based challenges



Camouflaged Underwater Object Detection



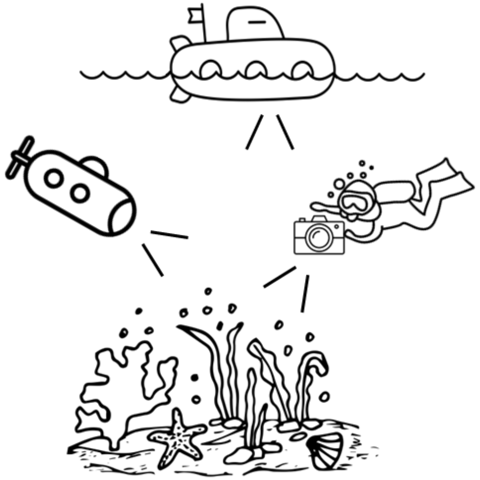
Video by Melanie Wille

Presentation Outline

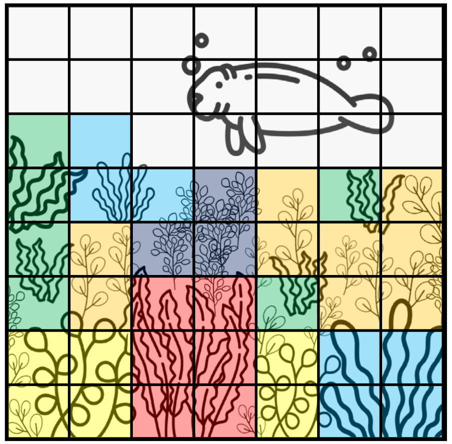
Bodies of Work

Intro

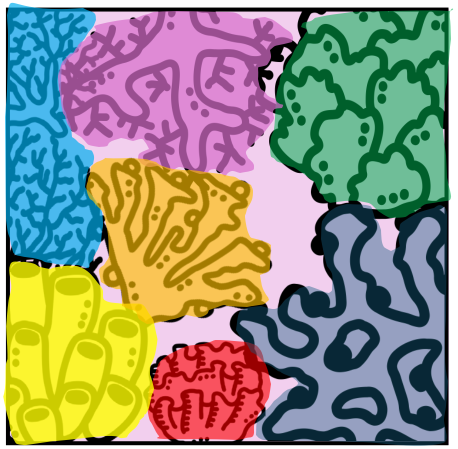
Motivation



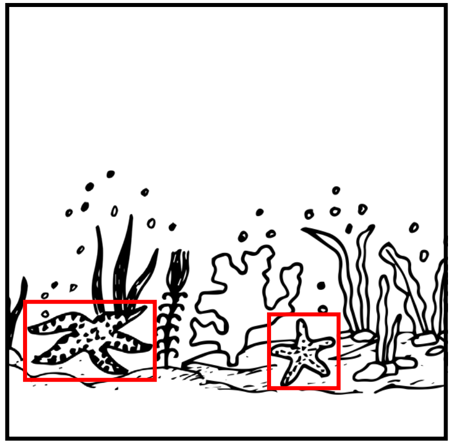
1. Seagrass: Coarse Segmentation



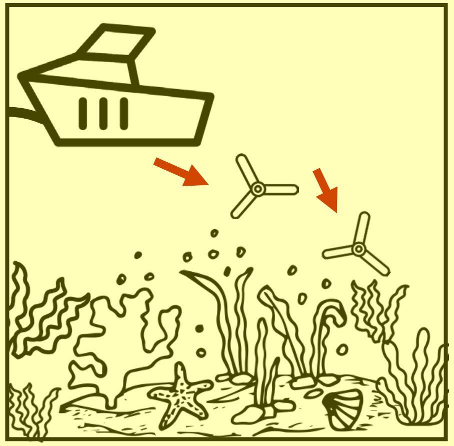
2. Coral: Segmentation



3. Underwater Object Detection



4. Large Scale Reef Restoration (with RRAP)



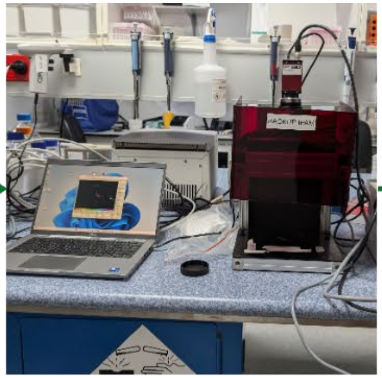
Future Work

Dr Dorian Tsai and many collaborators

Coral Aquaculture Pipeline

REEF RESTORATION & ADAPTATION PROGRAM

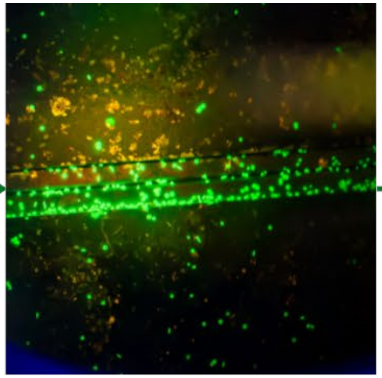
Coral Re/Collection



Broodstock Selection & Holding

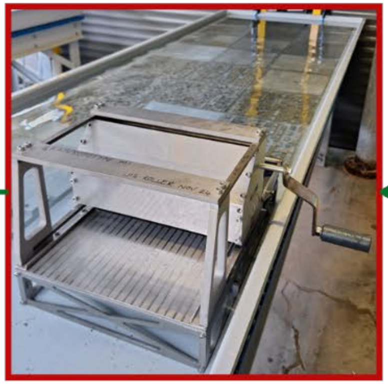
Mass Spawning & Selective Breeding
CSLICS Desktop

High-density Larval Rearing
CSLICS



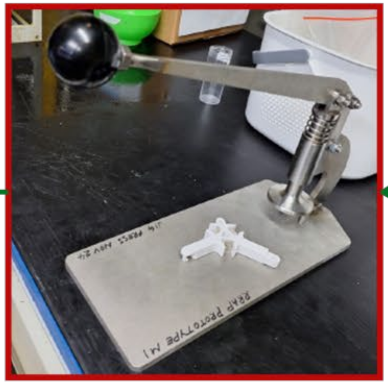
Mass Settlement

Coral Growout
CGRAS



Coral Tile Breakage
Jigs (cutter)

Tab Selection
OCTAV



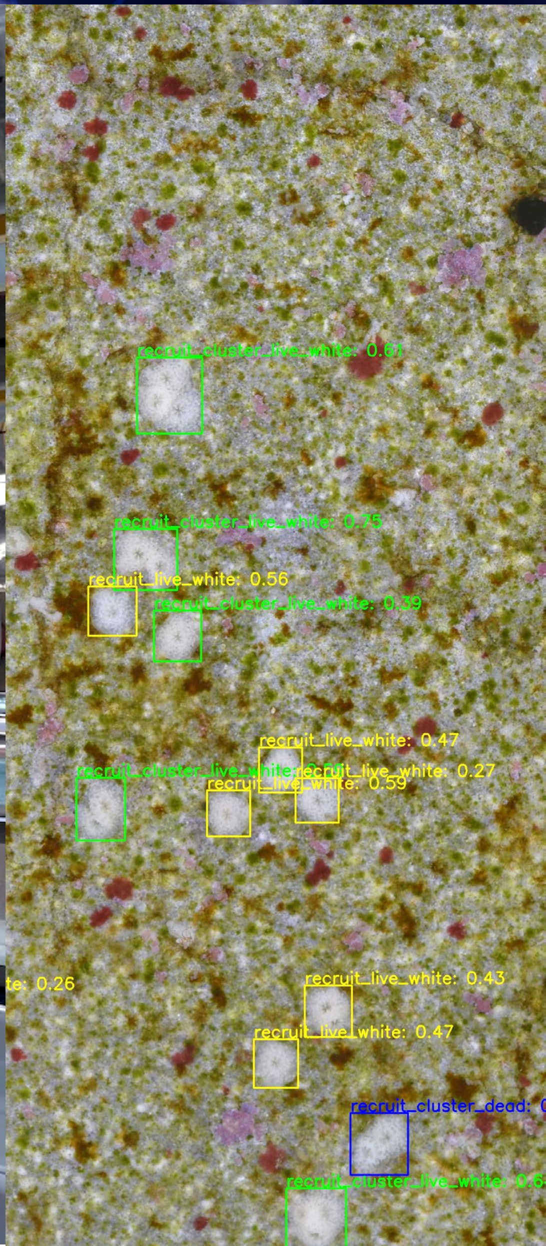
Device Assembly
Jigs (press) + Device Assembly Machine

Transport

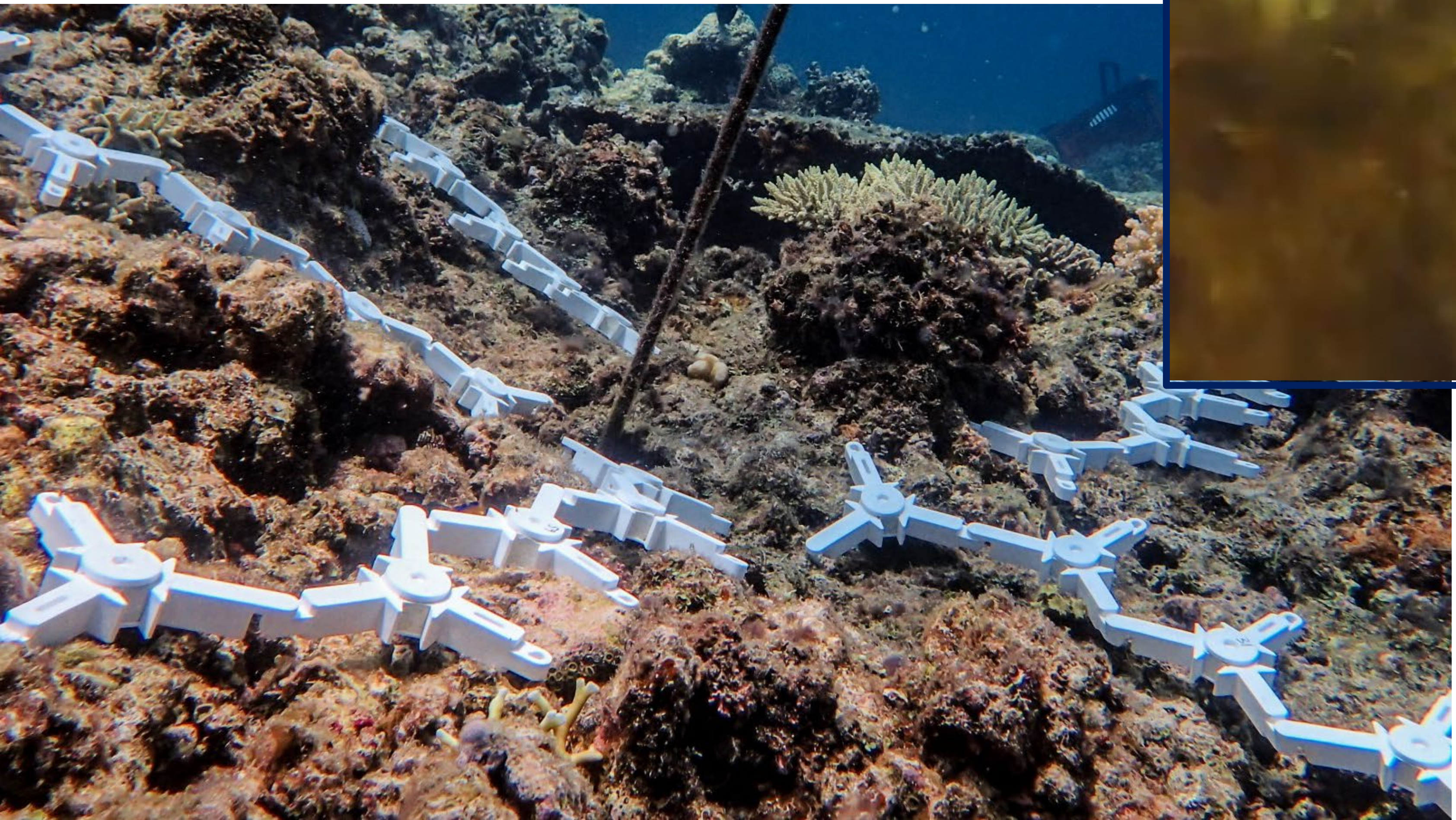
Device Deployment
Deployment Guidance System

Monitoring

AUSTRALIAN INSTITUTE OF MARINE SCIENCE



Coral Reseeding



REEF RESTORATION & ADAPTATION PROGRAM

Australian Government

Australian Institute of Marine Science

AI-driven Dispensing of Coral Reseeding Devices for Broad-scale Restoration of the Great Barrier Reef (under review)

Scarlett Raine, *Benjamin Moshirian and Tobias Fischer*

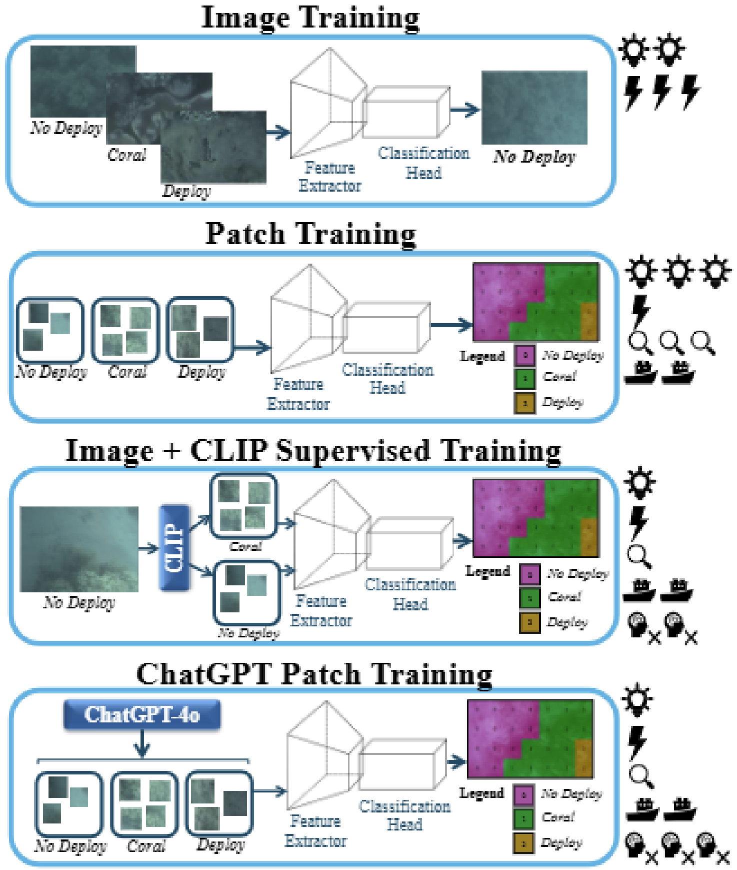
REEF RESTORATION & ADAPTATION PROGRAM

Australian Government | Australian Institute of Marine Science

Reef Guidance System

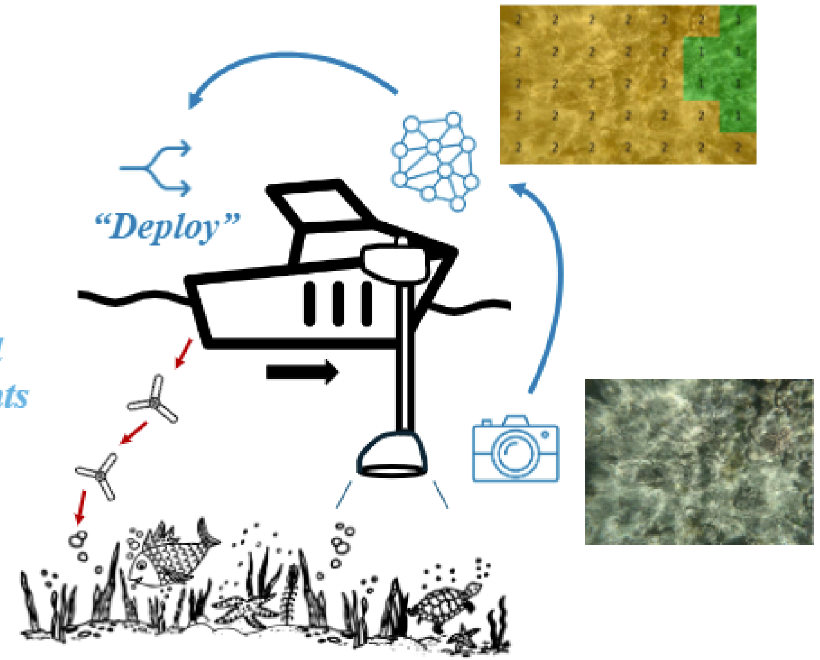
Operational Considerations

- **Model Accuracy**
- **Inference Speed**
- **Desired Level of Model Interpretability**
- **Flexible Boat Configuration i.e. single vs dual device deployment**
- **Domain Expert Unavailable for Data Labelling**



Select model to suit operational requirements

Reef Guidance System at Inference Time



Scarlett Raine, *Benjamin Moshirian and Tobias Fischer*

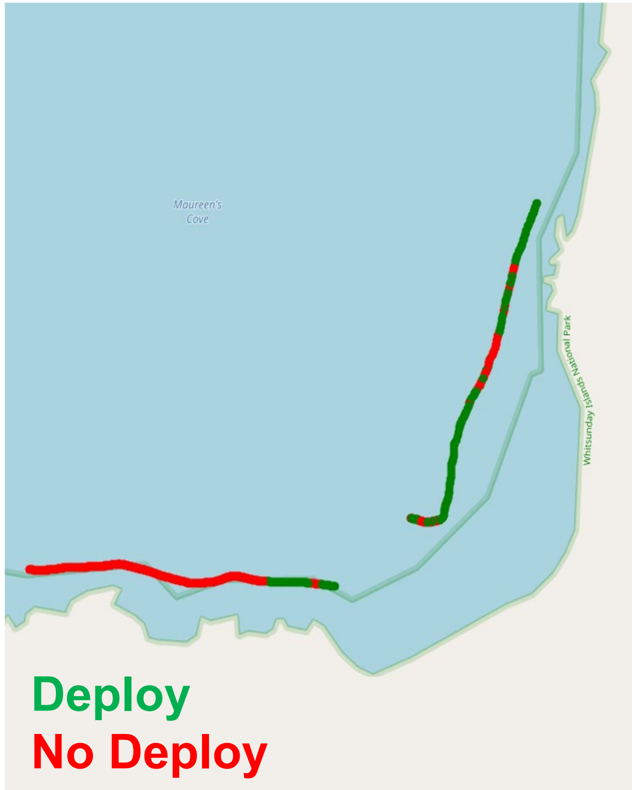


REEF RESTORATION & ADAPTATION PROGRAM

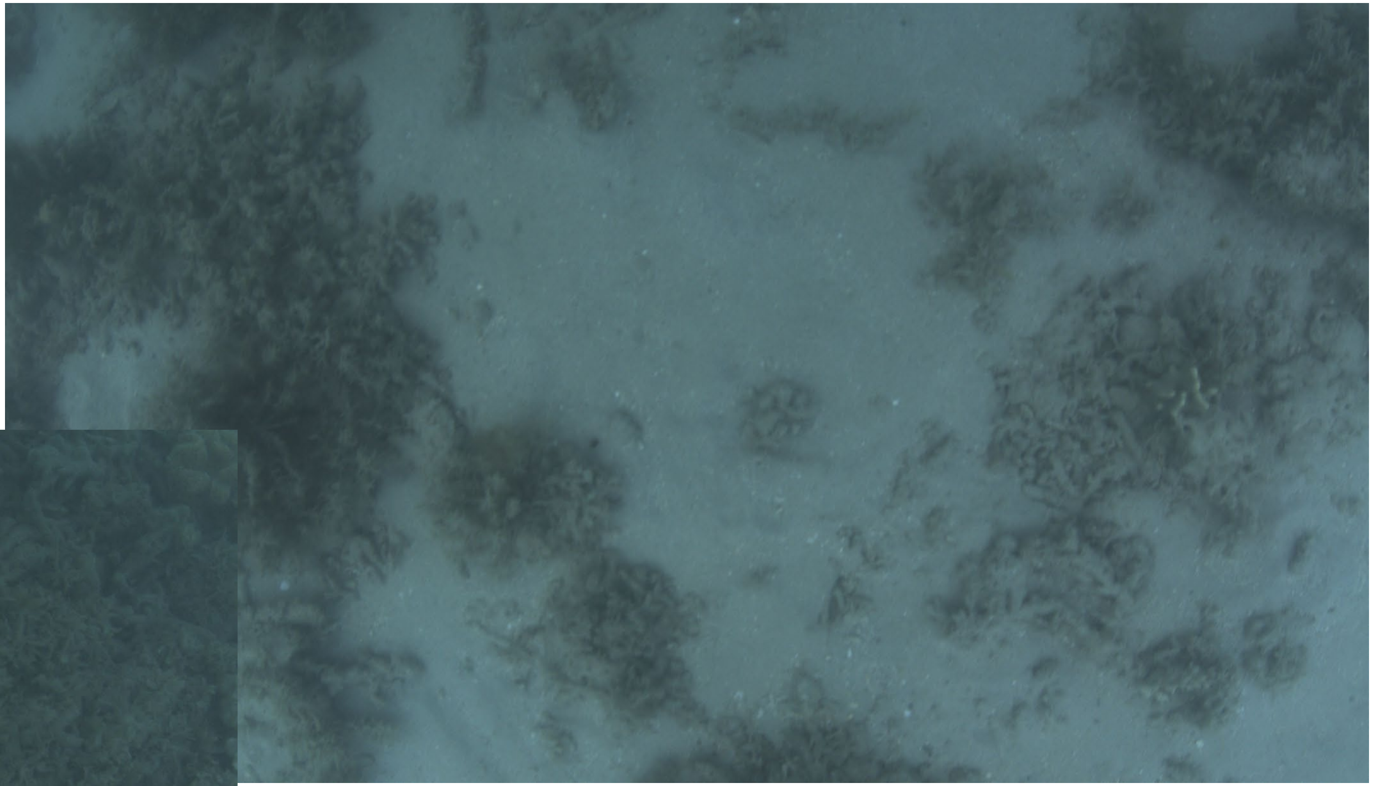
Australian Government

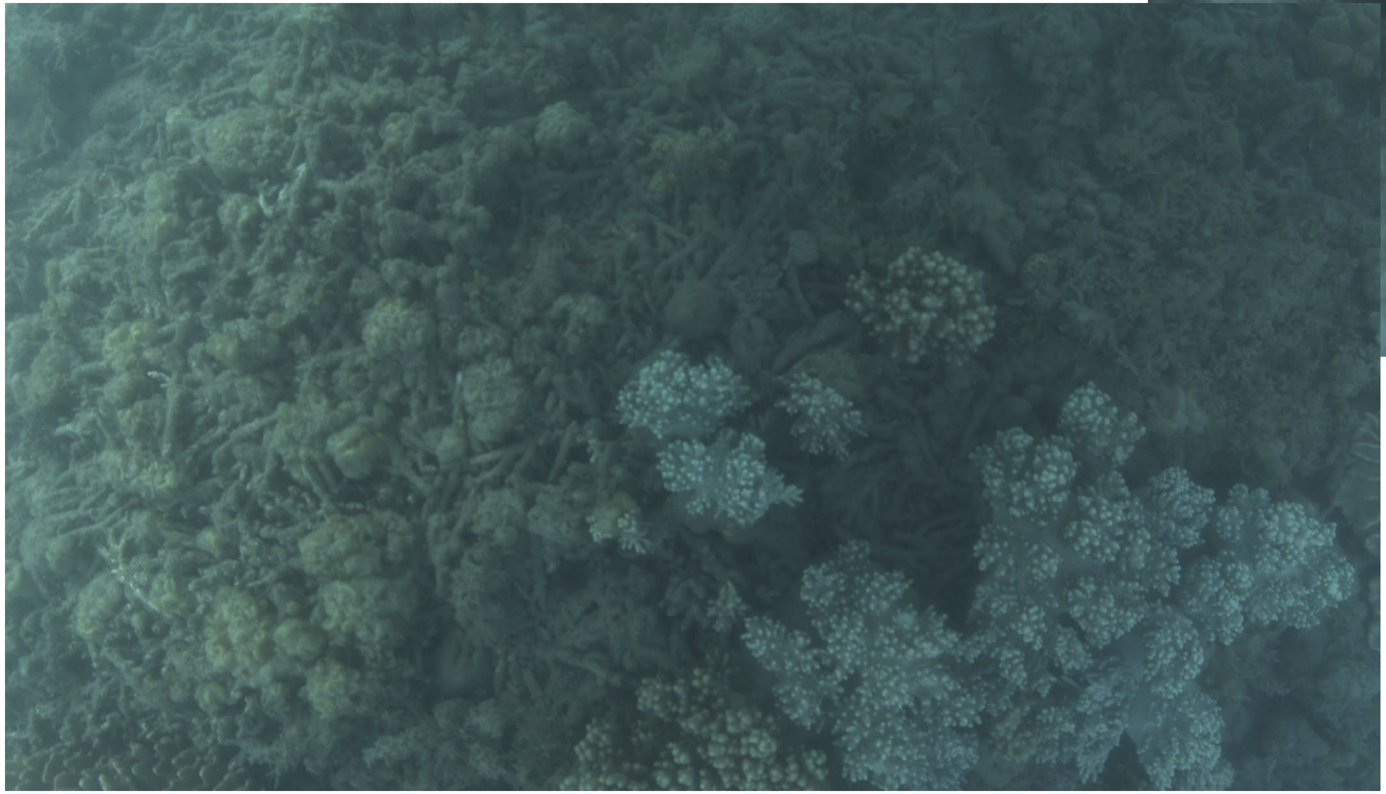
AUSTRALIAN INSTITUTE OF MARINE SCIENCE

Reef Guidance System



Raine, S., Moshirian, B., & Fischer, T. (2025). AI-driven dispensing of coral reseeding devices for broad-scale restoration of the Great Barrier Reef. *arXiv preprint arXiv:2509.01019*. [under review]





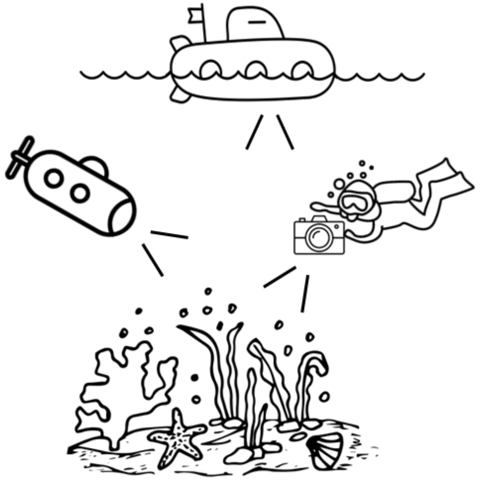
In underwater images, which colour is attenuated (reduced) the **most**?

Red, resulting in blue-green tinted imagery

Presentation Outline

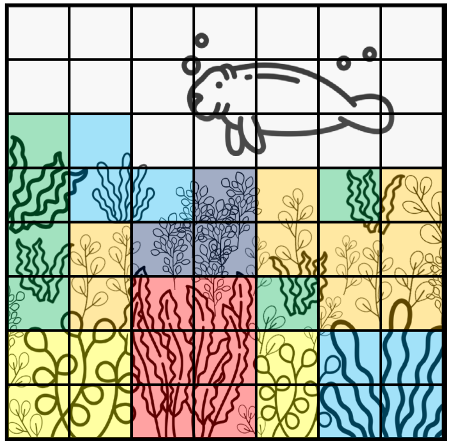
Intro

Motivation

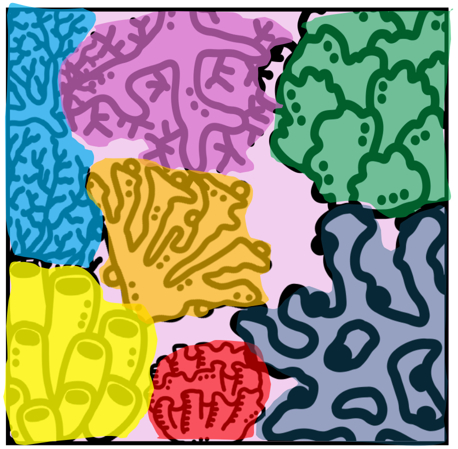


Bodies of Work

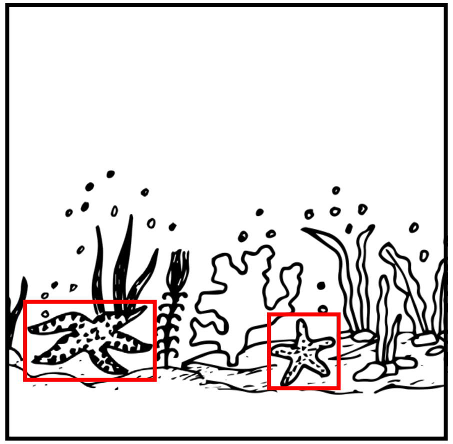
1. Seagrass: Coarse Segmentation



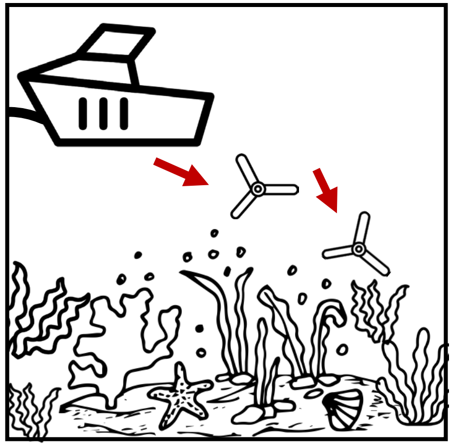
2. Coral: Segmentation



3. Underwater Object Detection



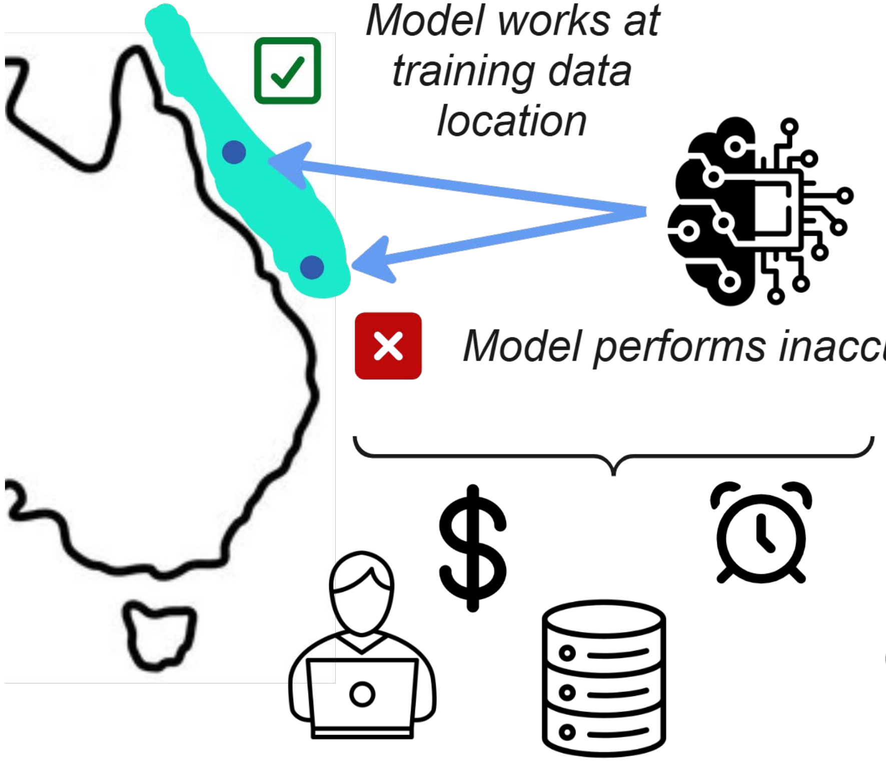
4. Large Scale Reef Restoration (with RRAP)



Future Work

Upcoming Research

- Open-set recognition for novel marine species
- Domain-transfer between training and deployment sites
- Inter-observer variation and uncertainty-aware approaches



AI models do not generalise to new locations, conditions

Model retraining is costly and time-consuming, requires training data and experts for labelling

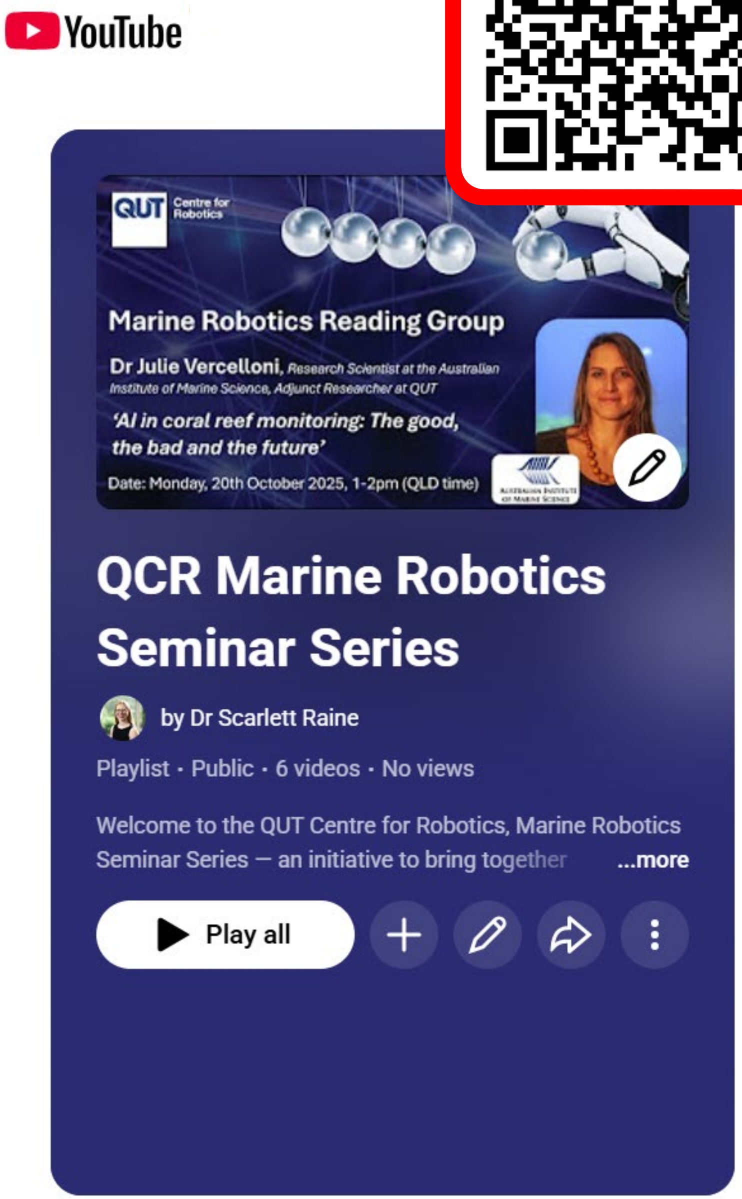



QUT Centre for Robotics

Marine Robotics Seminar Series

- Once a month
- International speakers
- Whole range of topics: from marine science, computer vision, robotics, reef restoration, sensing etc
- Let me know if you'd like to join!

https://sgraine.github.io/marine-robotics-seminars/



Centre for Robotics

Underwater Robotic Vision for Ecosystem Monitoring and Reef Restoration

Dr Scarlett Raine

Chief Investigator, QUT Centre for Robotics

sg.raine@qut.edu.au