UC3: Marine Ecosystem Monitoring at EMSO-OBSEA
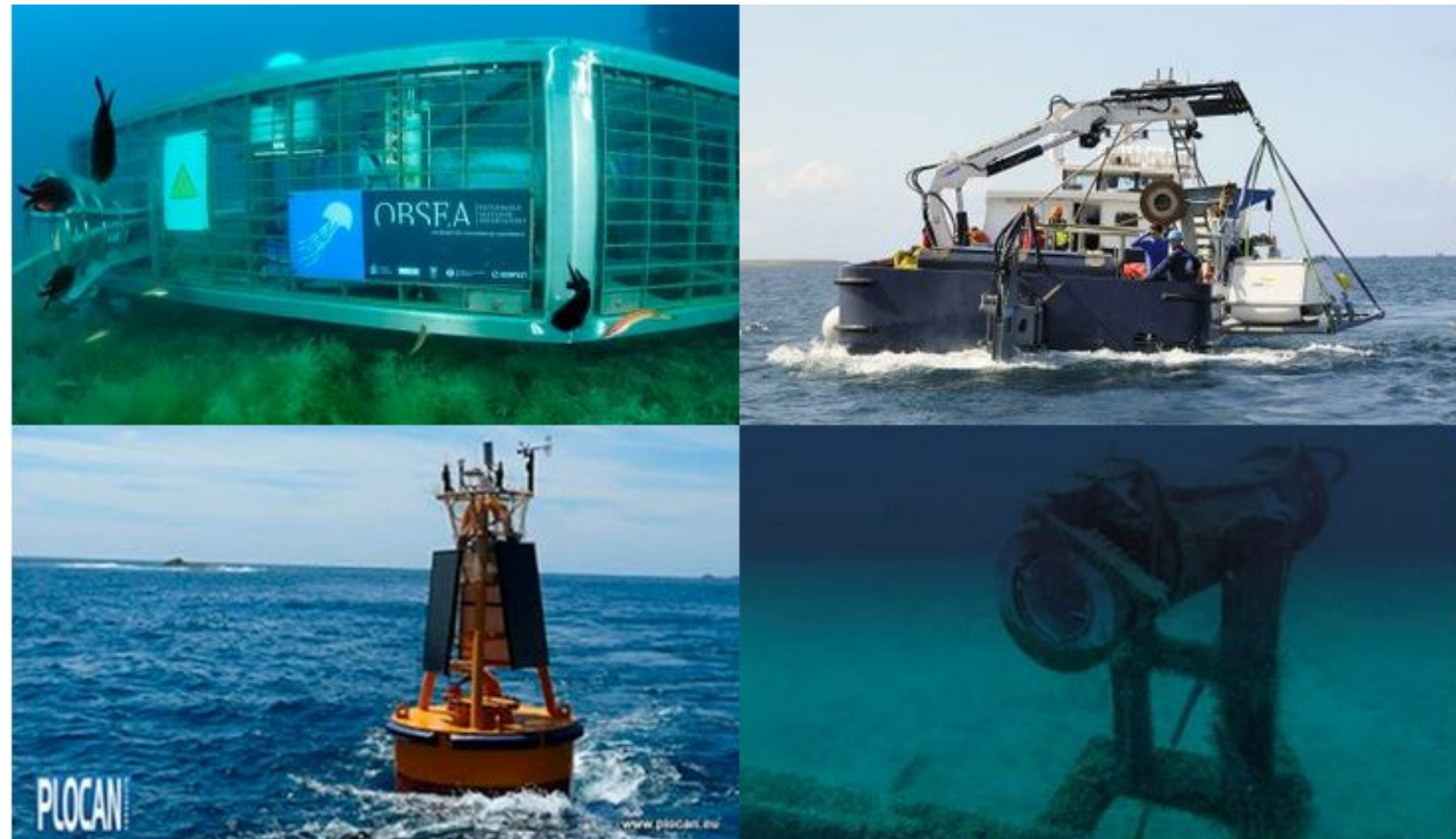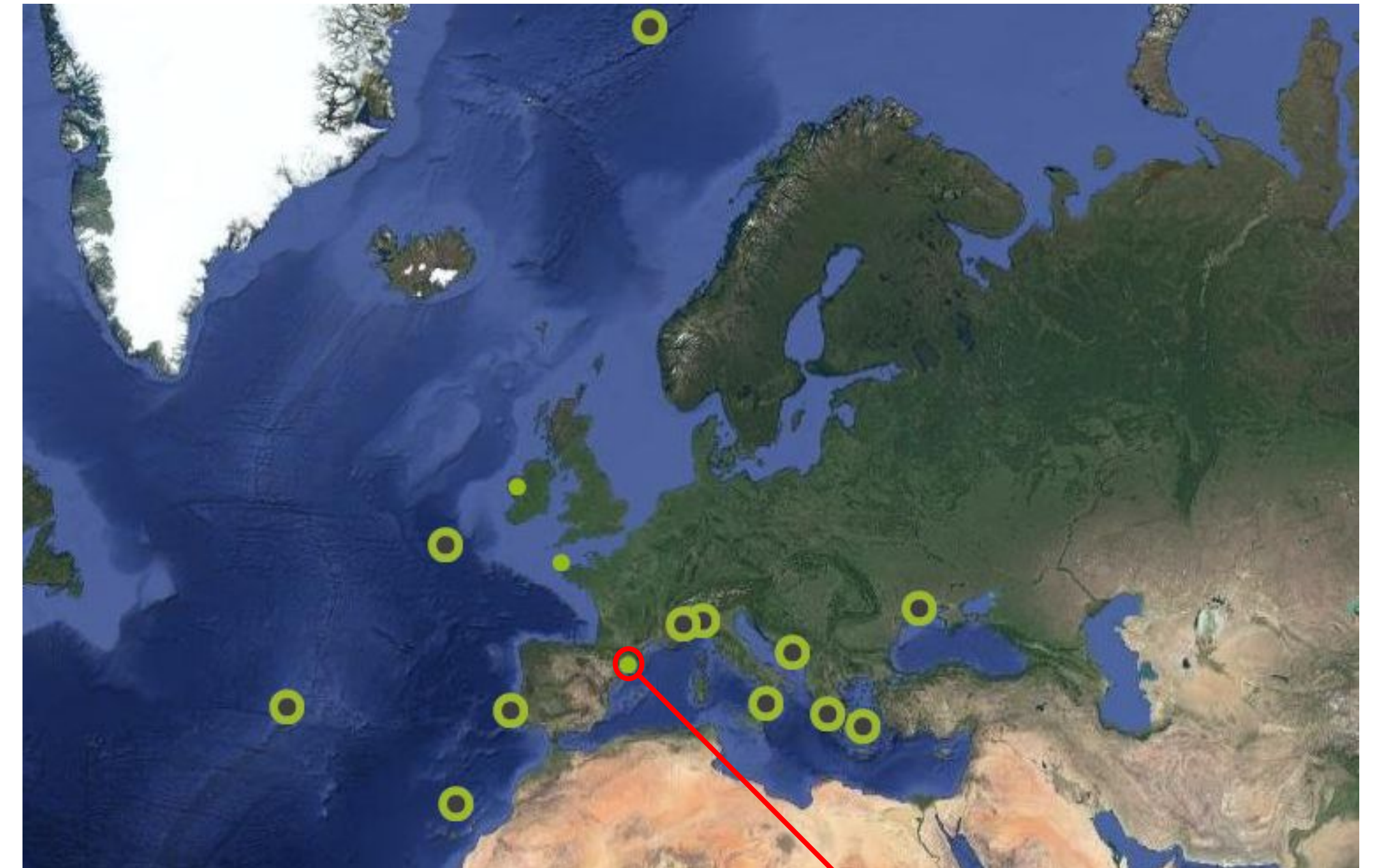
Enoc Martínez
UPC / EMSO ERIC

iMagine RP1 review

5th December 2023

EMSO ERIC

European Multidisciplinary Seafloor and water column Observatory

- Distributed research infrastructure
- 15 underwater nodes
- Fixed–point ocean monitoring:
  - Physical oceanography
  - Geochemistry
  - **Biological monitoring**

OBSEA Seafloor Observatory

- Underwater cabled observatory

- Located at NW Mediterranean Sea

- Shallow waters (20 m depth)

- Underwater pictures since 2011
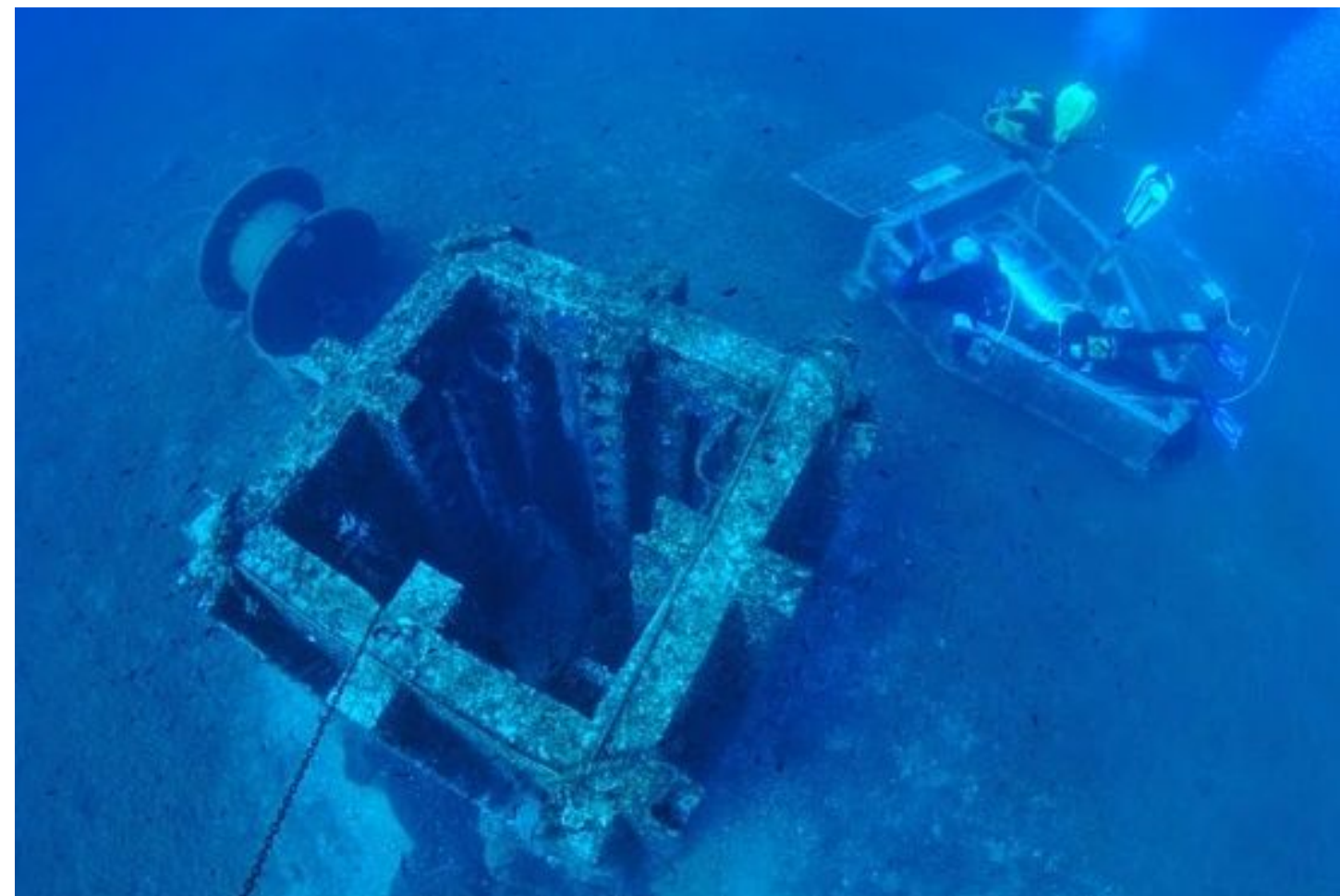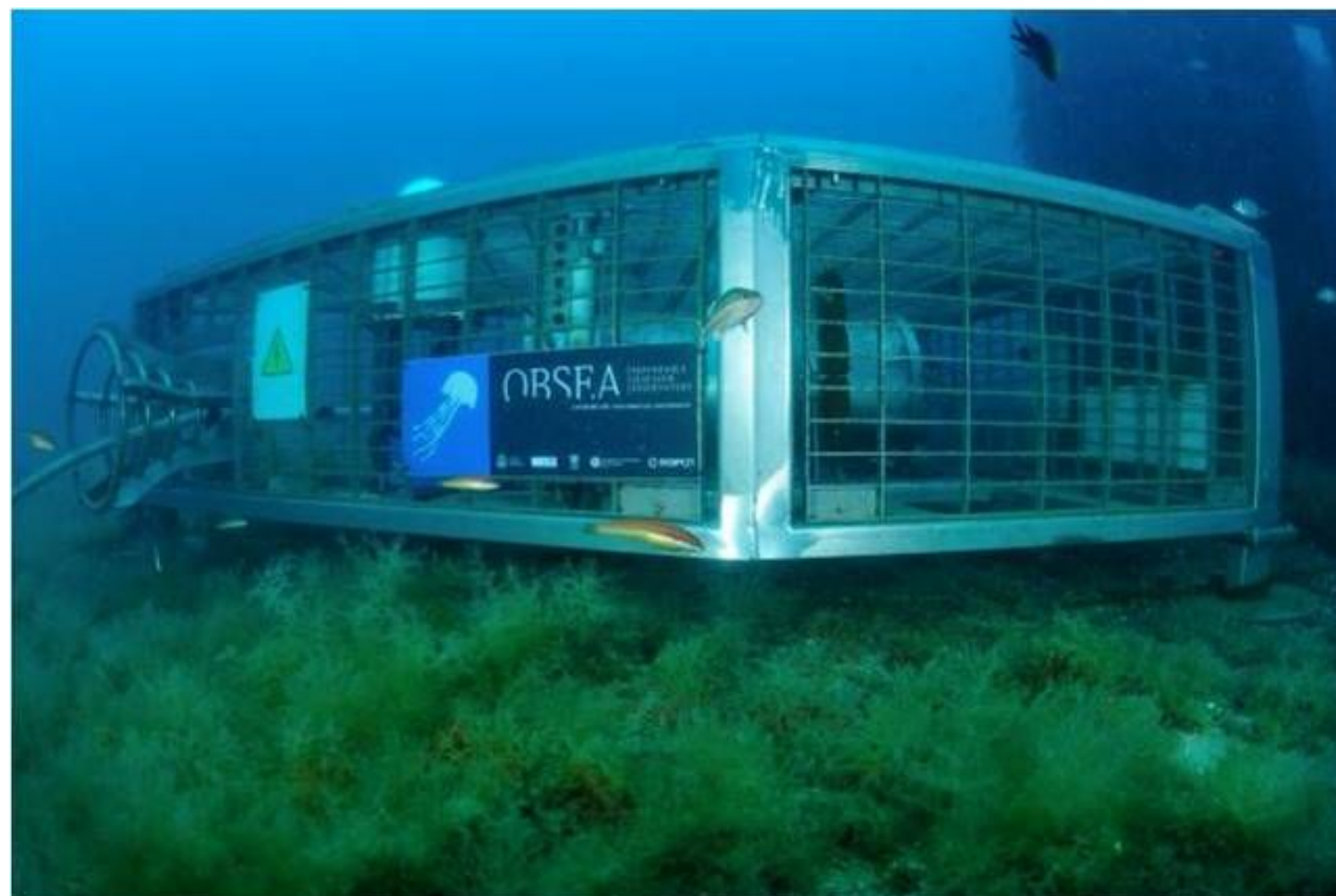
UC3 Objectives

Ecosystem Monitoring at EMSO–OBSEA

Objective 1: Support Data Managers

- Develop an AI model to analyze underwater pictures to increase the impact of existing images archives

Objective 2: Support Biologists

- Use a service for automated analysis of underwater pictures to study fish communities (abundance, behaviour...).

Starting Point

Team Expertise

- Electronic/IT engineering
- Some biology expertise
- No AI experience in-house

Existing dataset

- 30k partially-labelled images
- low image quality (640x480 px)
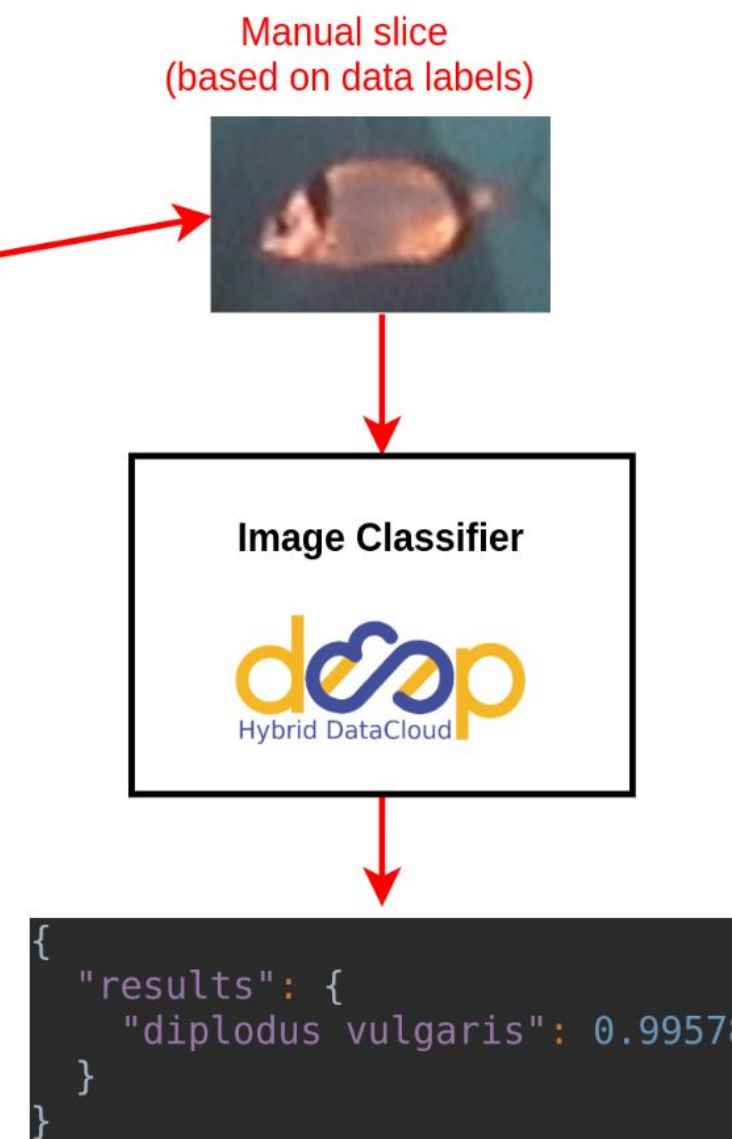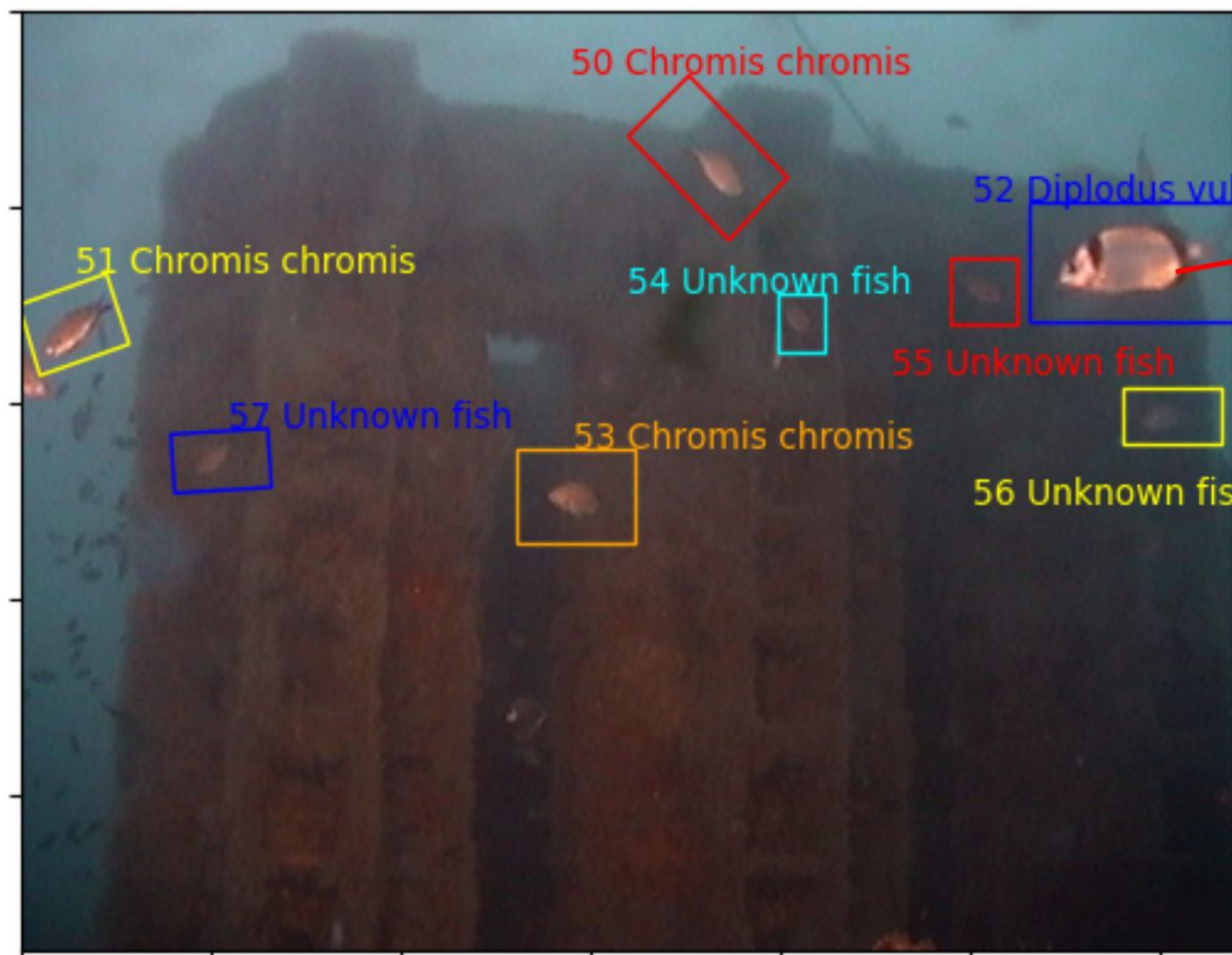- Biofouling issues
- **Unused dataset!**

5

Progress First Year
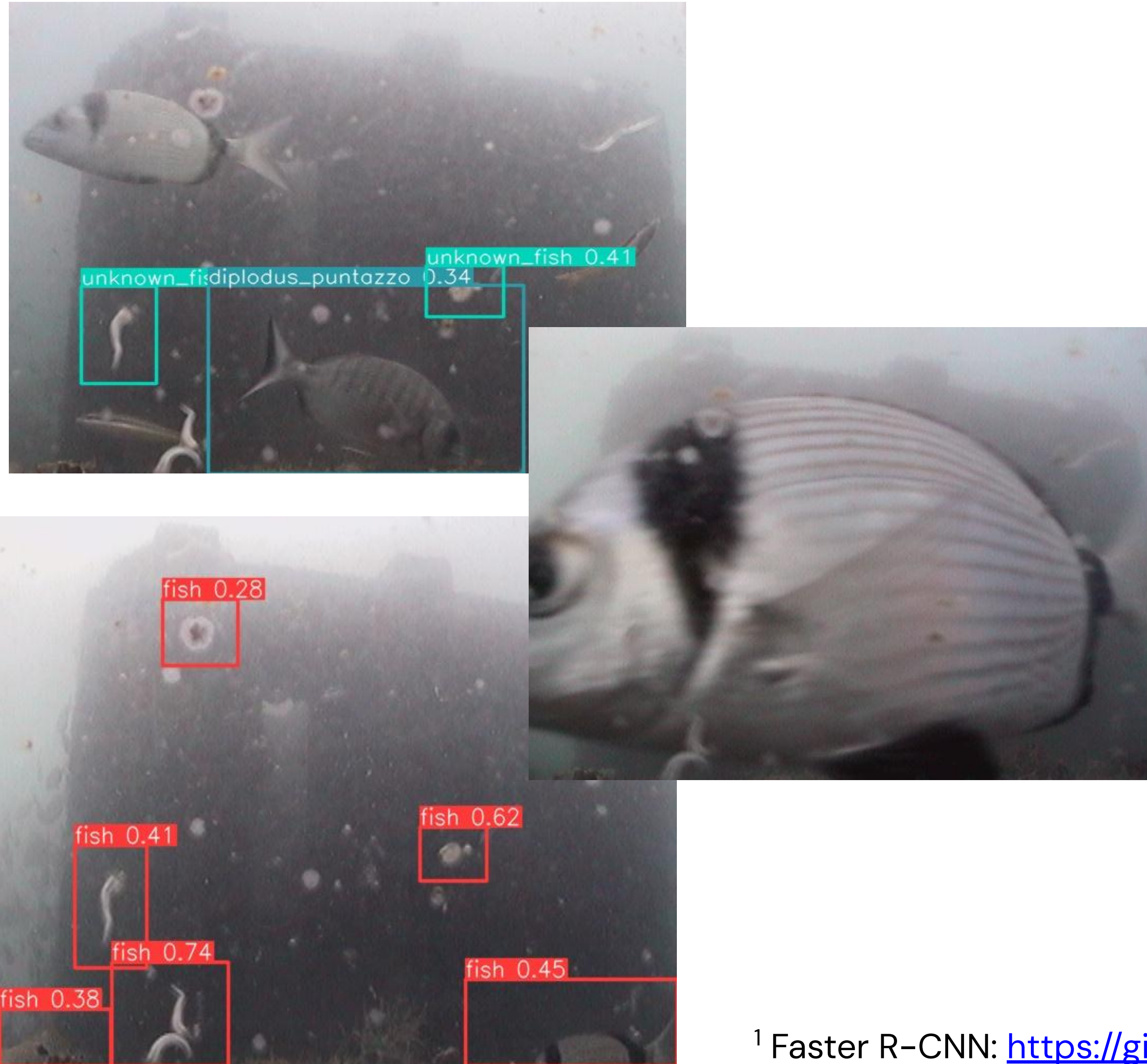
PyTorch Object Classification

- off–the–shelf marketplace module
- trained with "old" dataset
- trained with iMagine platform resources

Results

✓ Reasonably good classification

✗ No detection

✗ Manual selection of regions

Progress First Year

Object Detection/Classification

- Two open-source models used:
  - Faster R-CNN[1]
  - YOLOv8[2]
- Trained with "old" dataset

Results
- ✗ Missing "clear" specimens (top)
- ✗ Big fish not detected (middle)
- ✗ Biofouling/dirt detected as fish (bottom)
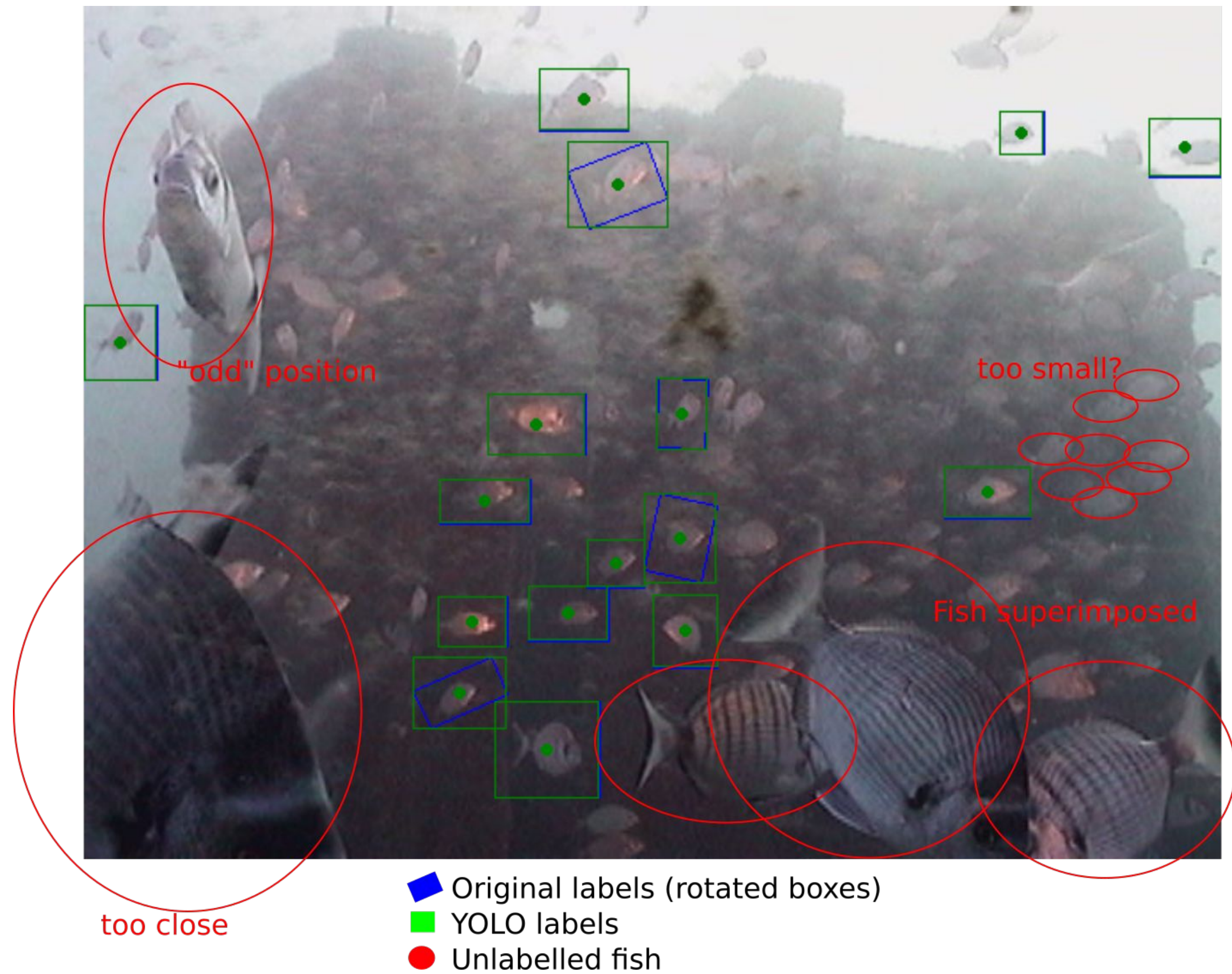
[1] Faster R-CNN: https://github.com/sovit-123/fasterrcnn-pytorch-training-pipeline
[2] YOLOv8: https://docs.ultralytics.com/

Progress First Year

"odd" position

too small?

Fish superimposed

too close

■ Original labels (rotated boxes)
■ YOLO labels
● Unlabelled fish
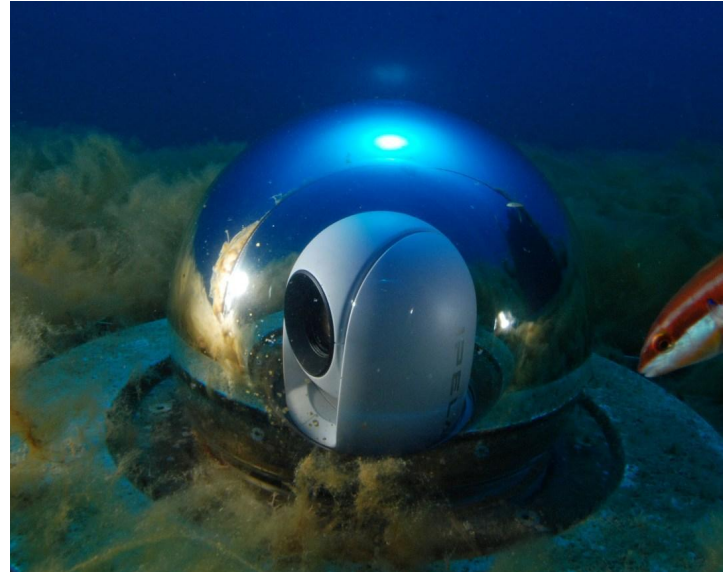
Old Dataset Analysis

Missing labels:
- Overlapped fish
- "Odd" positions
- Close the optics

✗ **Dataset not usable**

Progress First Year

Old Camera
Resolution
Glass sphe

New HD Camera

✓ Reduced Biofouling
✓ Better Resolution
✓ New Dataset with HD pictures

Plans for 2nd Year

New Dataset

Create a new dataset with:

- HD pictures
- Good labels
- Quality over quantity!

Status December 2023

In month M15 we already have:

✓ 1700 labelled images

✓ 15k manual annotations

Plans for 2nd Year

AI-models

Train two YOLOv8[1]-based models:
- *Slow* but precise model: 1 pic/min
  - High-precision scientific timeseries
- *Fast* model for real-time video: 24 FPS
  - Data managers / dissemination
- Publish to the marketplace
- Deploy models in production

Status December 2023

In month M15 we already have:
- ✓ 2 models trained!
- ✓ Published in the marketplace
- ✓ Youtube demo
- ✓ Working with WP4 for video inference

[1] YOLOv8: https://docs.ultralytics.com/
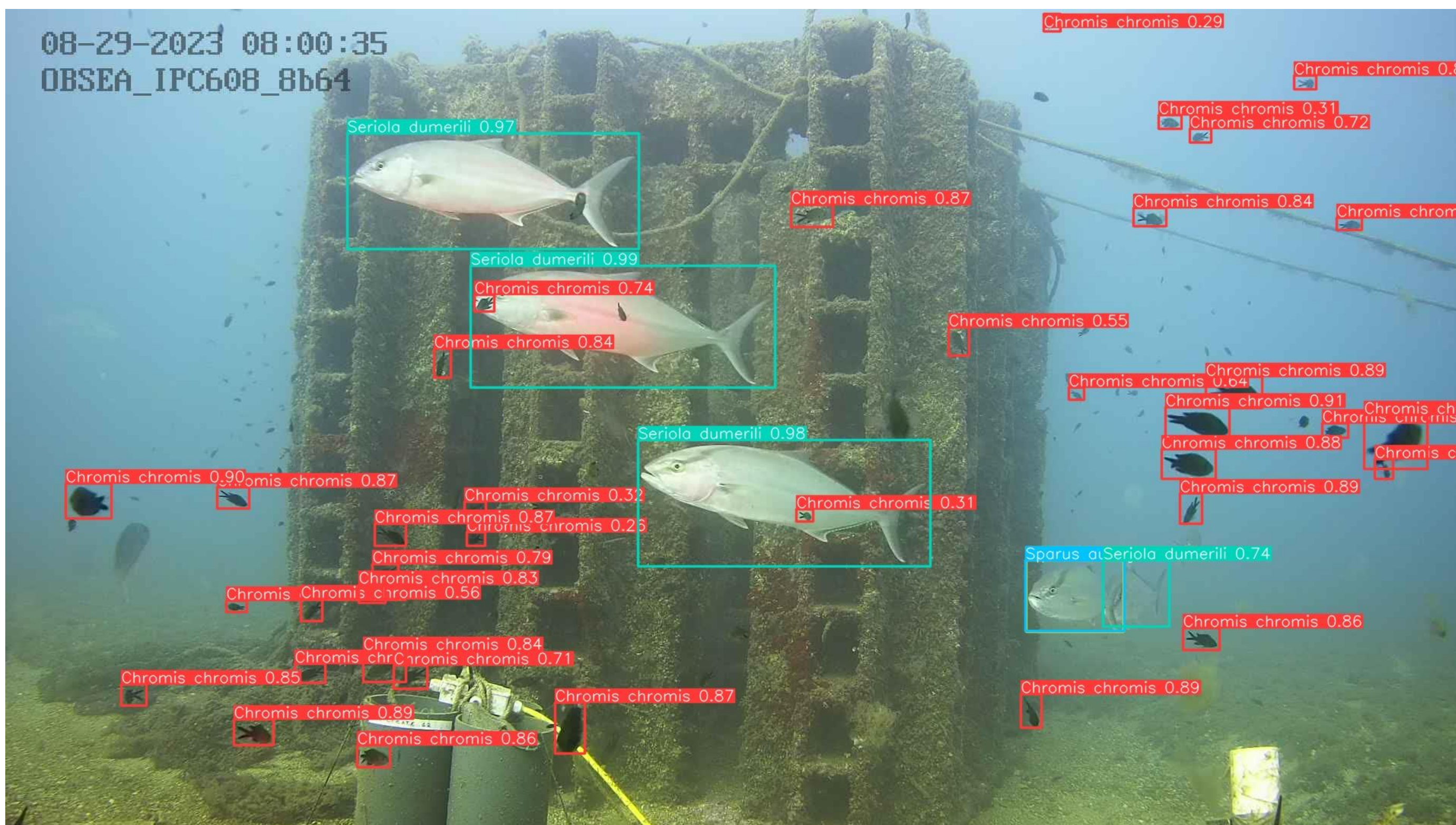
Preliminary Results

Plans for 2nd Year
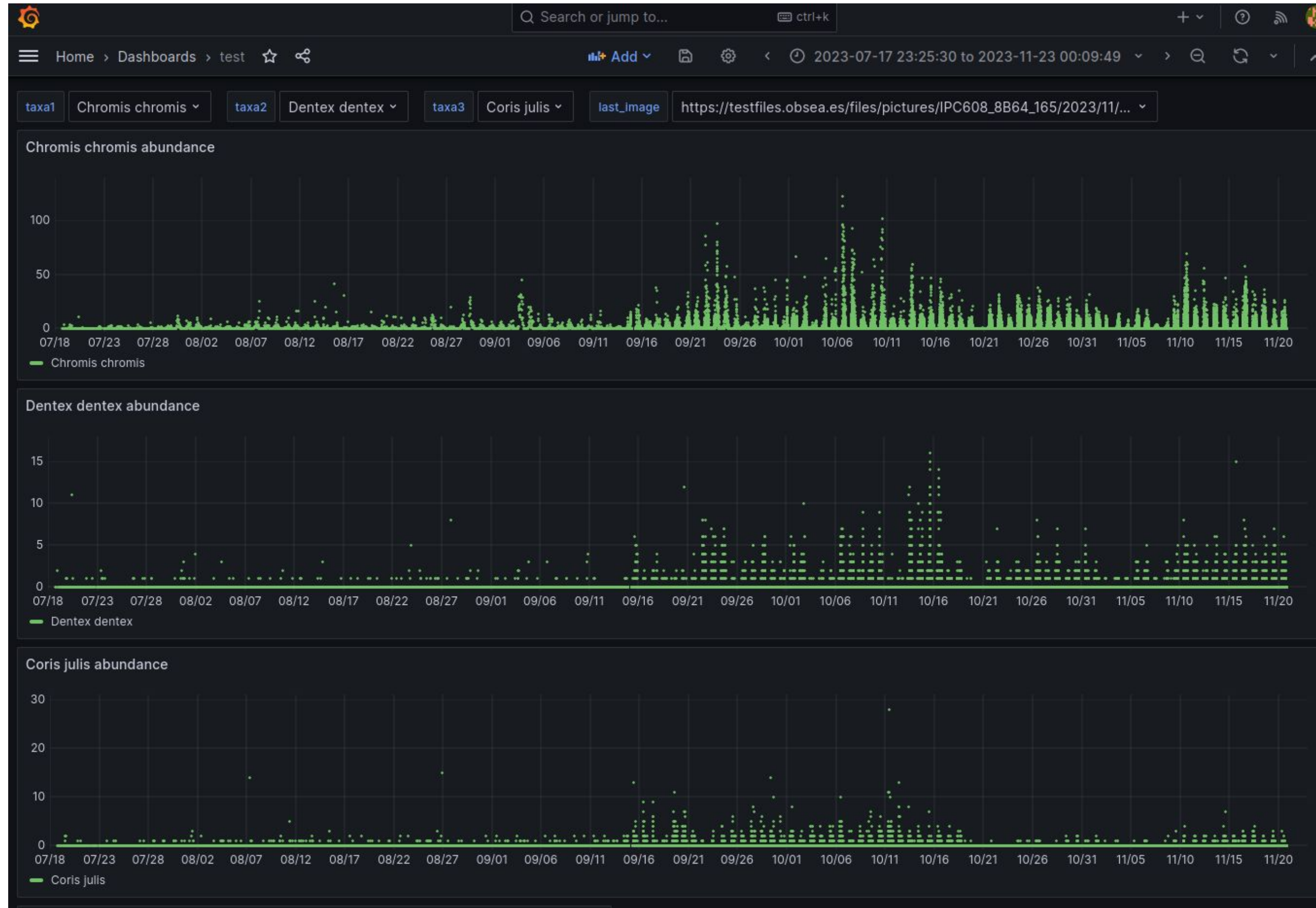
Data Production

- Extract fish abundance information

- Generate abundance time–series

- Publish dataset

Status December 2023
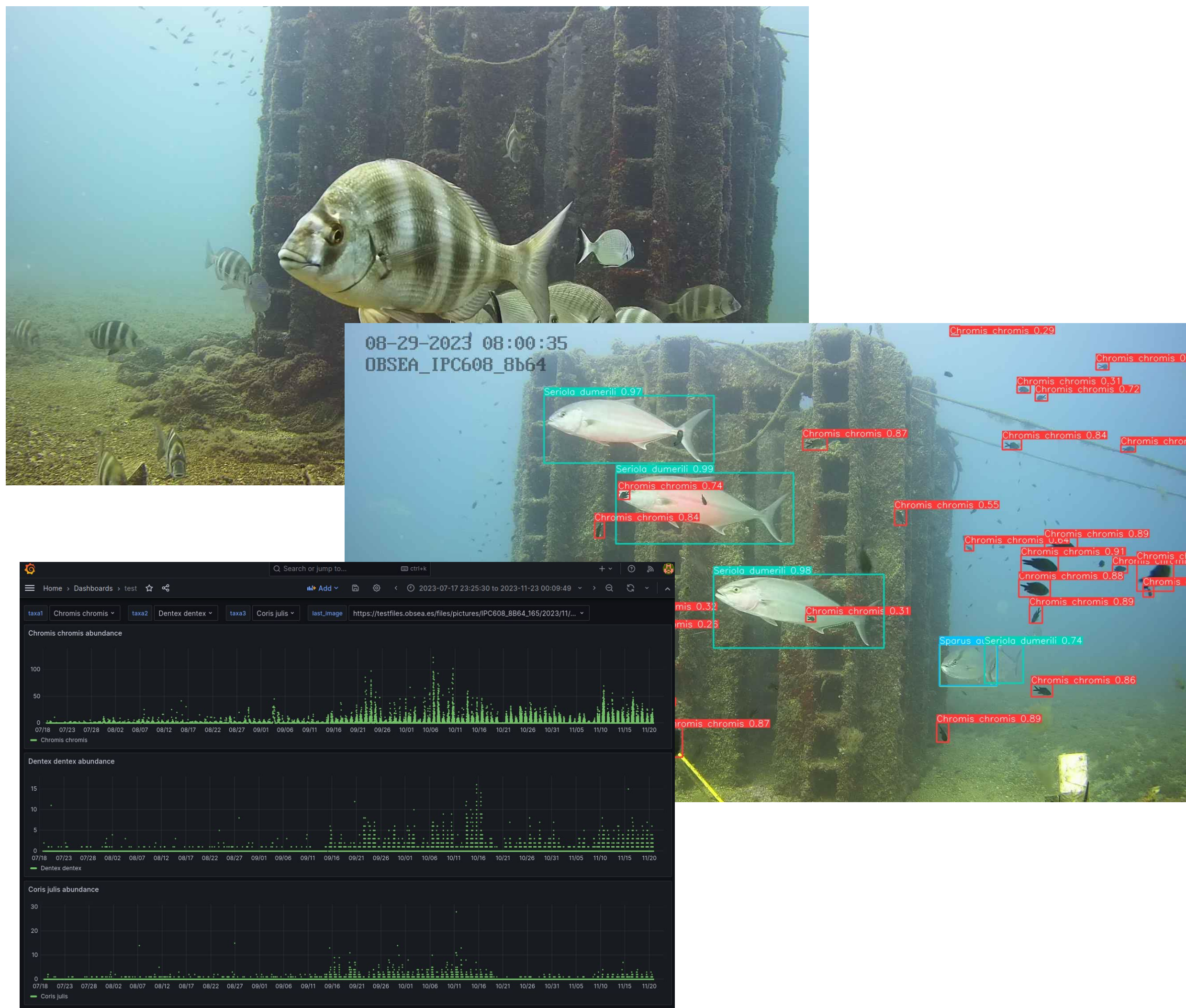
In month M15 we already have:
- ✓ Test time–series
- ✓ 50k pictures analyzed

Manually analysis of 30k images took **1 year**
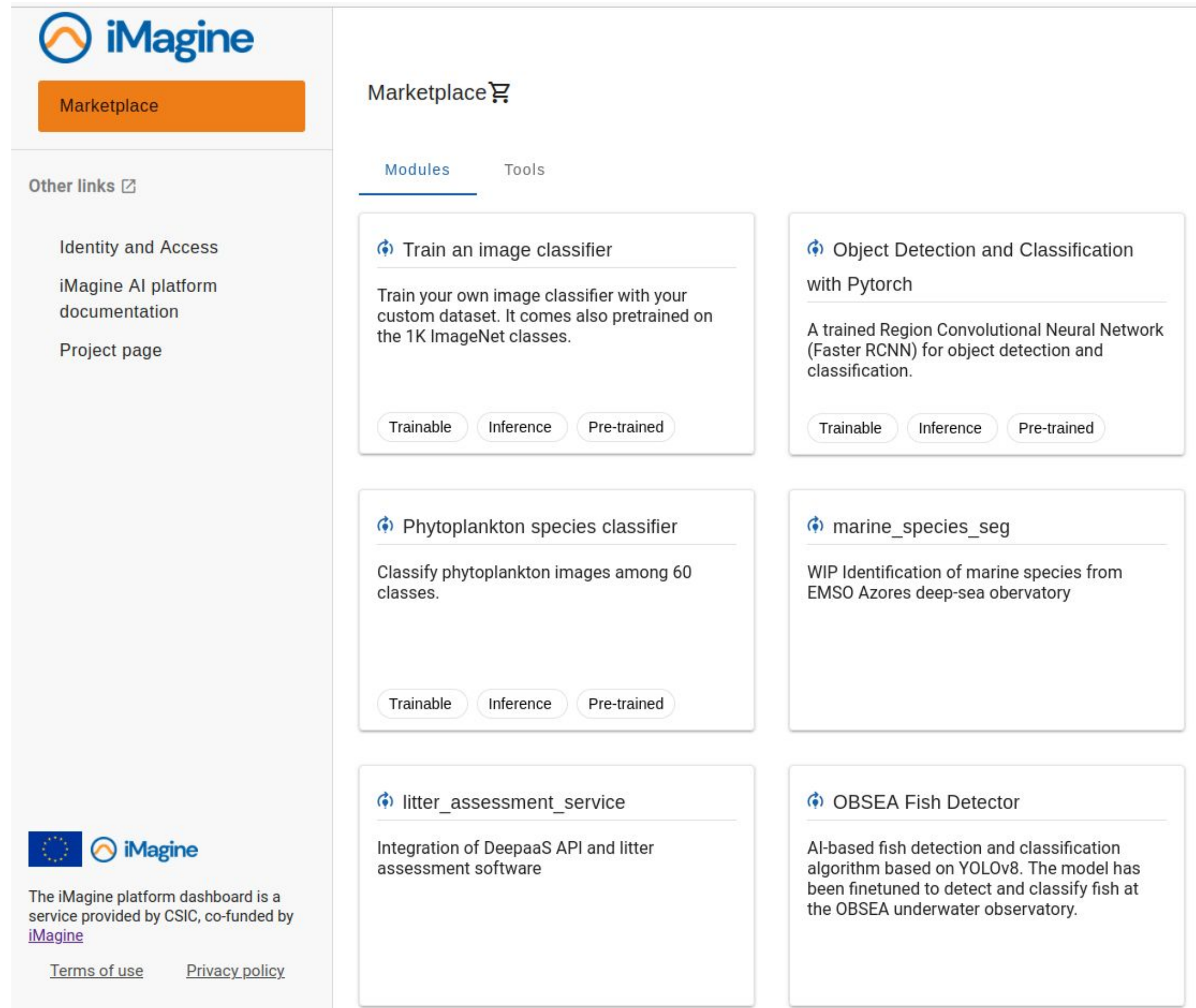50k images analysis on a laptop took **3 hours**

Plans for 3rd Year

Biological Information

- Publish scientific results

- Apply drift detection techniques

- MLOPs

- Promotion with other infrastructures:

  ○ EMSO SmartBay

  ○ EMSO Azores

  ○ PAP

  ○ JERICO

  ○ LifeWatch

  ○ EMBRC

iMagine Platform

Experiences with the Platform

✓ User friendly

✓ Easy to deploy modules

✗ Some issues with virtualization

    ✓ WP4 team fixed them

✗ Modules with broken dependencies

    ✓ Good docs –› easy to recreate

    ✓ Sandbox module

UC3 Marine Ecosystem Monitoring at OBSEA

Thank you for your attention