19th May 2026

Marine Imaging Workshop 2026 - Gibraltar

# Uncovering Anomalous Events for Marine Environmental Monitoring via Visual Anomaly Detection

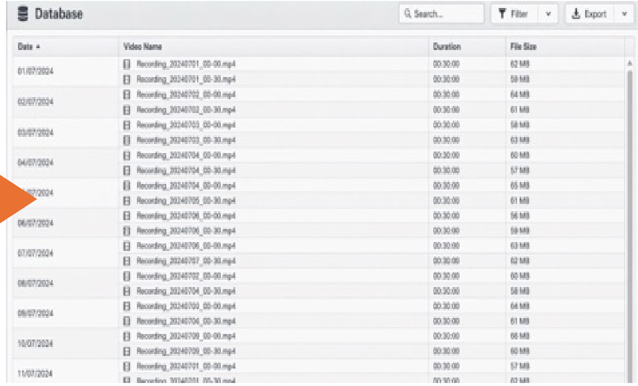

Laura Weihl[1,4], Stefan Hein Bengtson[2,4], Nejc Novak[3], Malte Petersen[2,4]

IT University of Copenhagen[1]
Aalborg University[2]
Anemo Robotics ApS[3]
Pioneer Centre for AI, Denmark[4]

# Environmental Monitoring

- Static cameras for long-term underwater biodiversity monitoring
- 30 secs recording every 30 mins for months
- Problem: get TB of data, but mainly empty/uninteresting videos

Biodiversity Monitoring

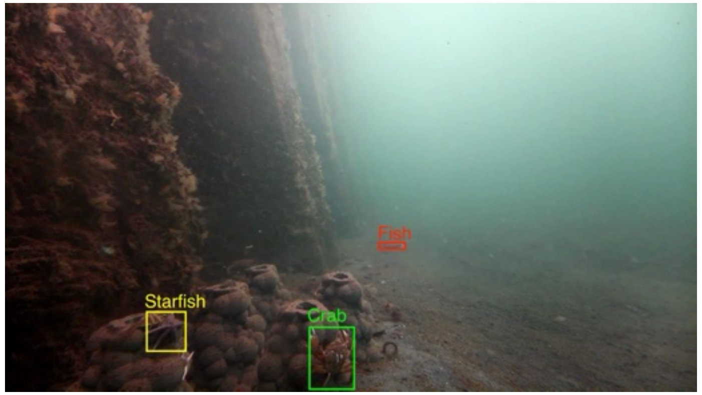



Huge Volumes of Data

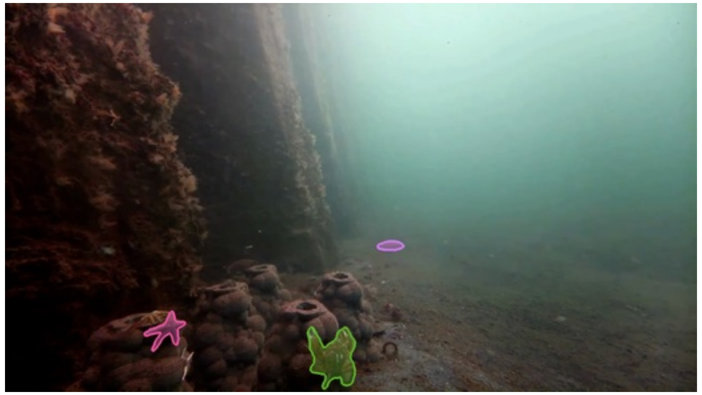
Empty Videos

Anemo Robotics

# Closed-set vs. Open-set



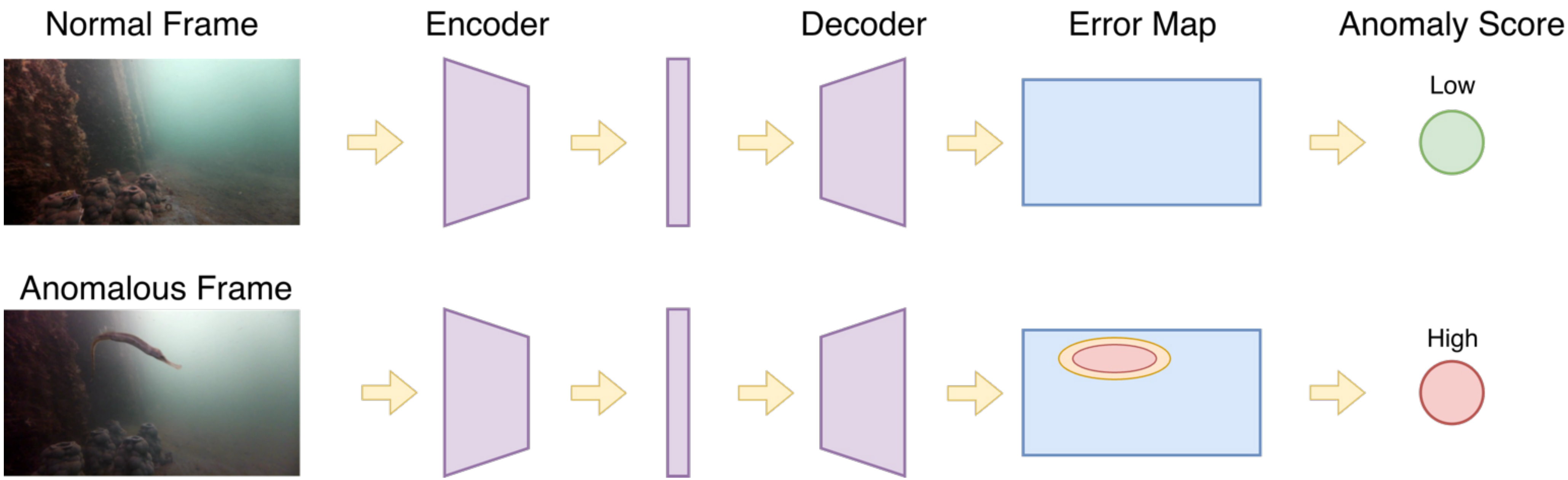

Object Detection



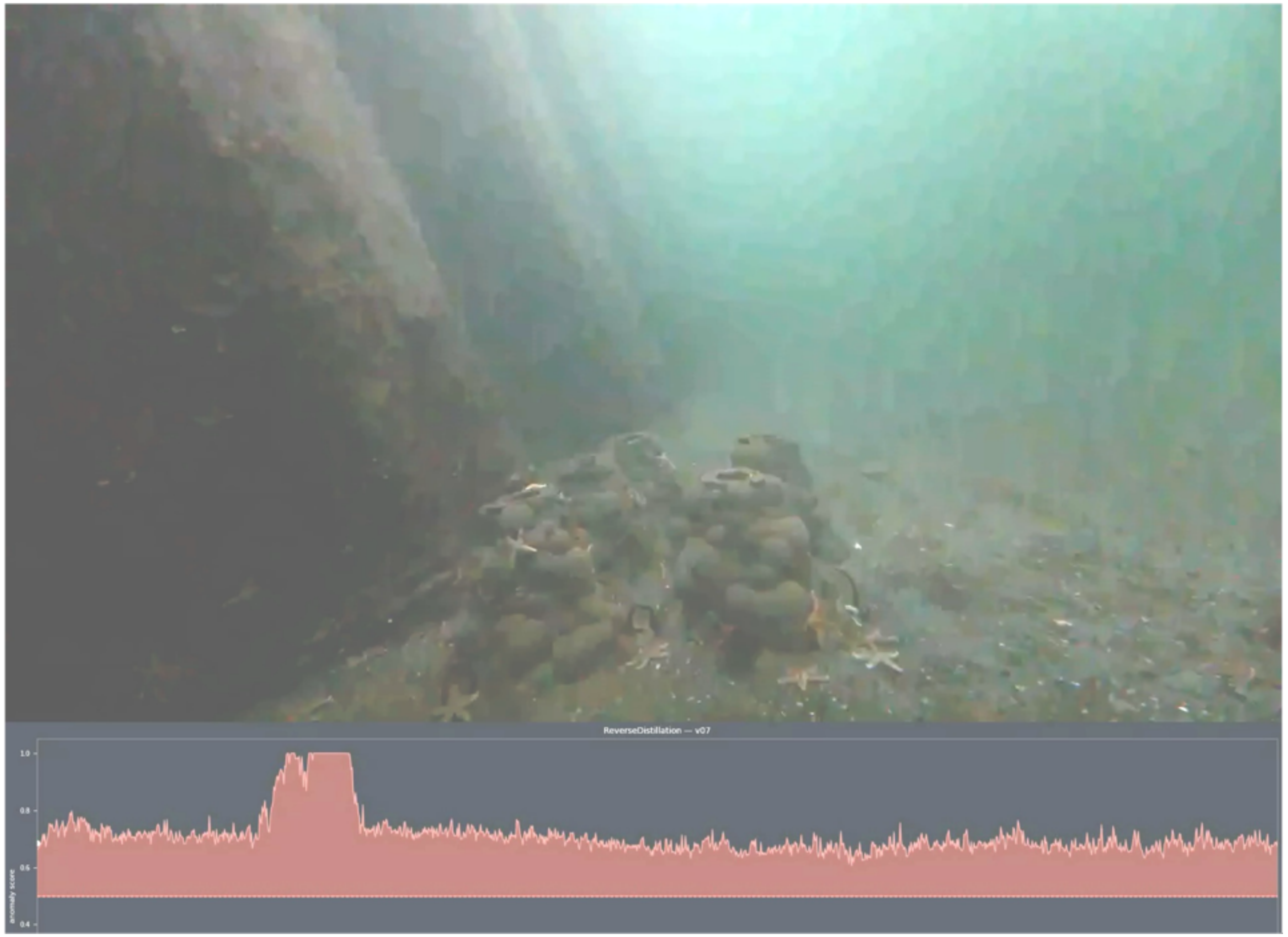
Segmentation

# Visual Anomaly Detection (VAD)

- Train on “normal” frames of the scene



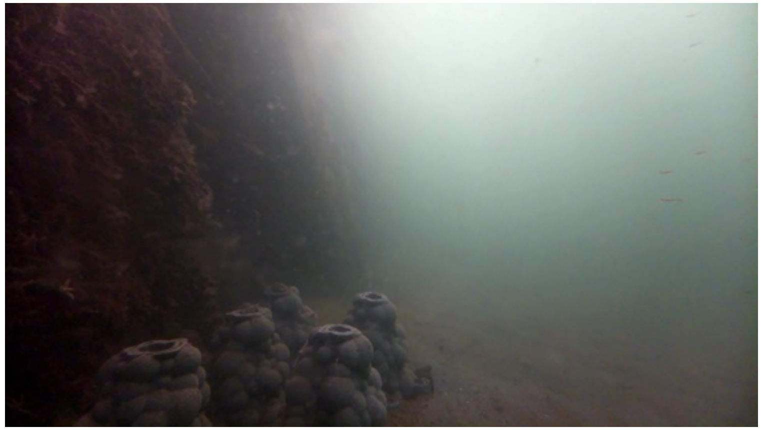
# Visual Anomaly Detection



# Annotation Problem

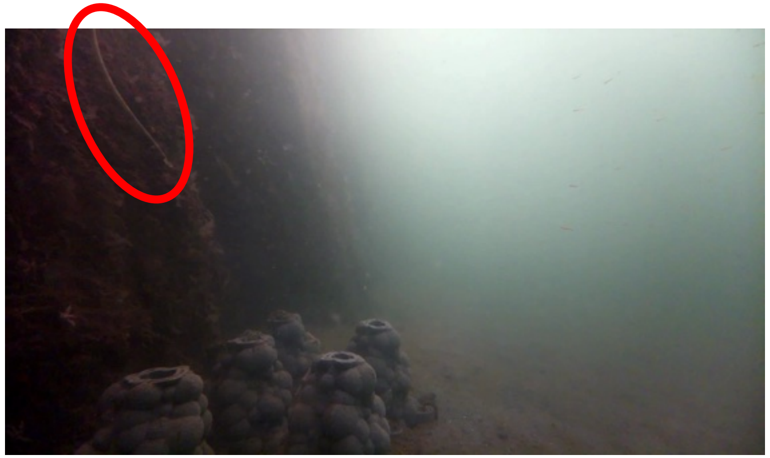
- **Challenge**: when does an anomalous event start/stop?
- We recruit 16 annotators and use soft labels

Frame 230



Soft label 0.0

Frame 370



Soft label 0.5

Frame 510

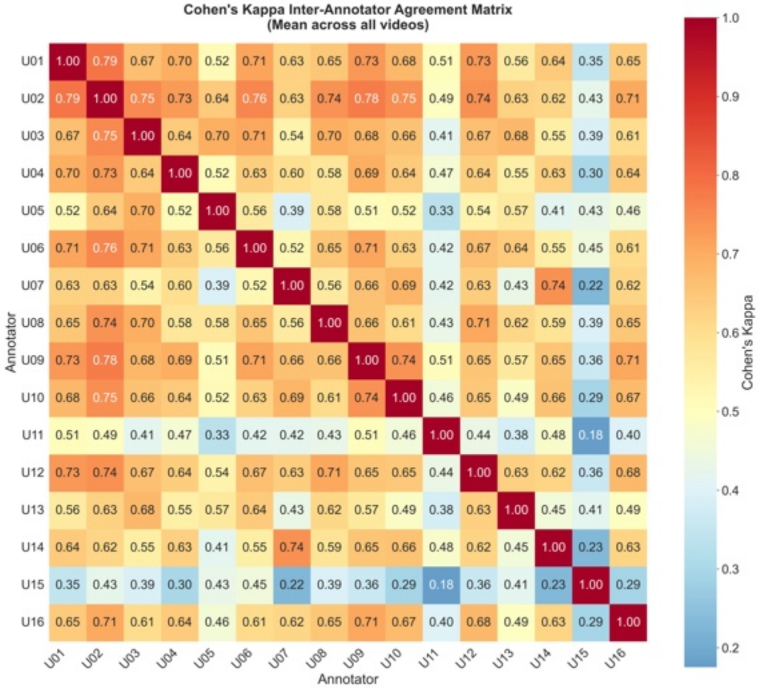
Soft label 1.0

# The AURA Dataset

- **New Dataset: A**nomalous **U**nderwate**R A**cticity
- 2 scenes, 25 videos, 16 annotators
- Moderate to substantial agreement

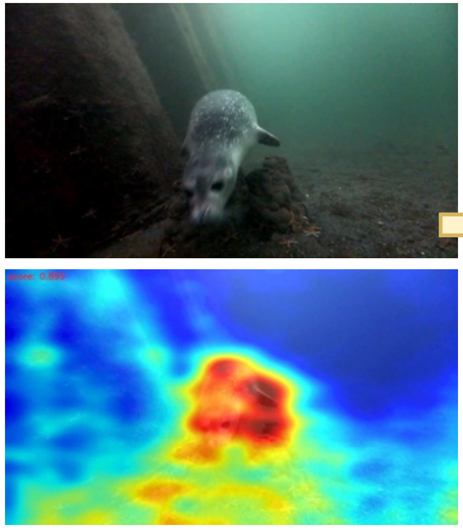
Scene A

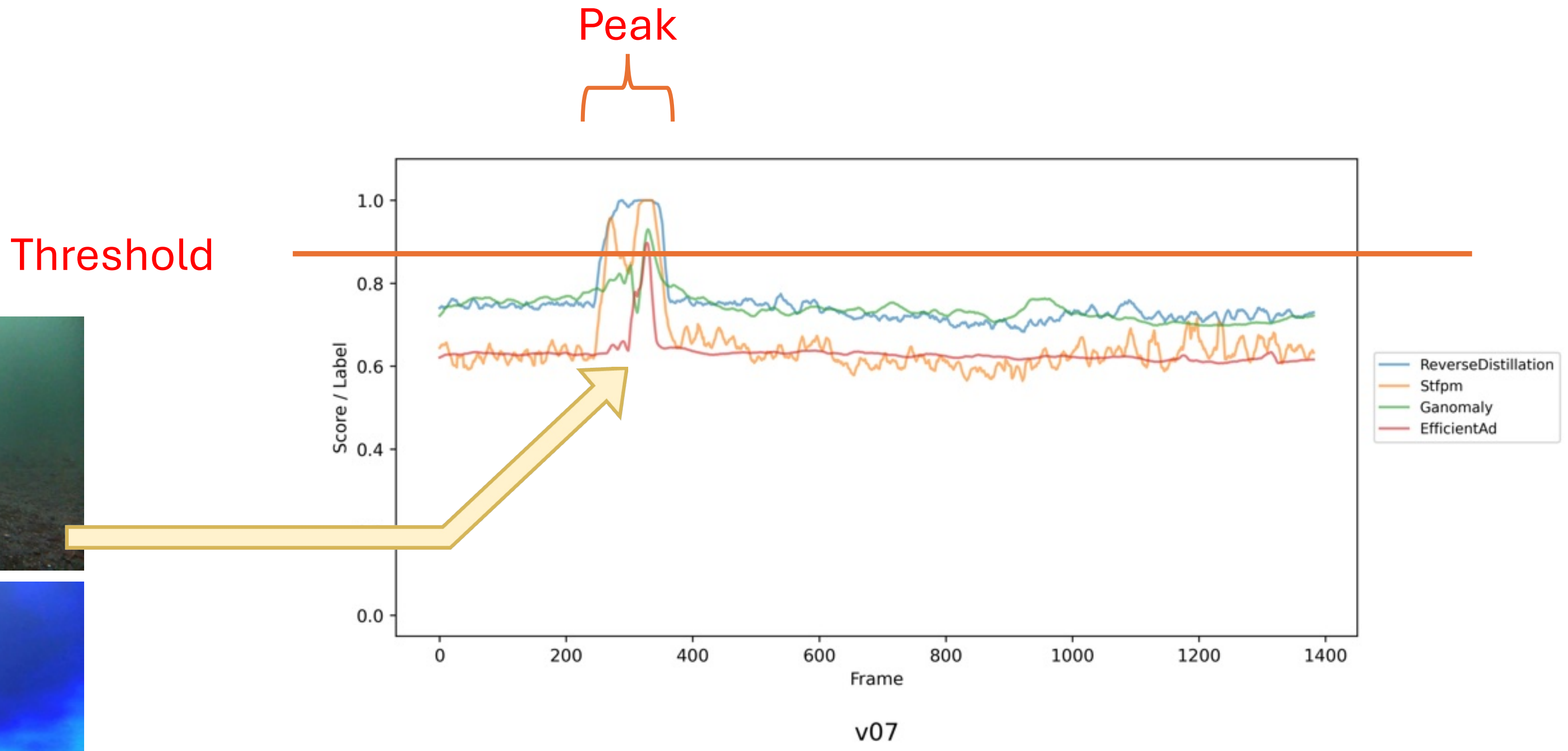
Scene B



# Experiments

- Evaluate 4 visual anomaly detection methods in anomalib [1]
- Event extraction: threshold-based vs. peak-based






[1] https://anomalib.readthedocs.io/en/lib-v2.3.3/

# Anomaly scores vs. soft labels

- Report mean average error (↓)

| | MAE (Mean ± Std) | |
|---|---|---|
| | **Scene A** | **Scene B** |
| EfficientAD | 0.393 ± 0.101 | 0.379 ± 0.048 |
| GANomaly | 0.425 ± 0.127 | 0.435 ± 0.099 |
| Stfpm | 0.465 ± 0.107 | 0.320 ± 0.058 |
| Reverse Distillation | **0.271** ± 0.151 | **0.259** ± 0.041 |

- Key takeaway:
  - Reverse Distillation aligns best with human annotations in both scenes
  - Scene B is consistently easier across models

# Results - tIoU

- t-IoU: temporal intersection over union

| | | Scene A | Scene B |
|---|---|---|---|
| EfficientAD | Find Peaks | 0.347 | 0.745 |
| GANomaly | Find Peaks | 0.242 | 0.471 |
| Reverse Distillation | Find Peaks | **0.673** | 0.861 |
| Stfpm | Find Peaks | 0.301 | **0.863** |
| EfficientAD | Threshold | 0.374 | 0.709 |
| GANomaly | Threshold | 0.083 | 0.217 |
| Reverse Distillation | Threshold | **0.461** | **0.806** |
| Stfpm | Threshold | 0.225 | 0.748 |

- Key takeaway:
  - Reverse Distillation performs best at finding anomalous events
  - Peak-based event extraction works better than threshold-based

# Conclusions

## Underwater VAD

- Consistent/artificial lighting (scene B) makes the scene easier to learn

## Event Detection

- Peak-based method performs better than threshold-based extraction

## Annotation Guidance

- Performance remains dependent on dataset size, variability, and model architecture




# Uncovering Anomalous Events for Marine Environmental Monitoring via Visual Anomaly Detection

**Laura Weihl[1,4], Stefan Hein Bengtson[2,4], Nejc Novak[3], Malte Petersen[2,4]**

IT University of Copenhagen[1]
Aalborg University[2]
Anemo Robotics ApS[3]
Pioneer Centre for AI, Denmark[4]

**AURA Dataset**

**Full Paper**

**Published at**
ICCV Joint Workshop on Marine Vision, 2025

lw112.github.io

lawe@itu.dk