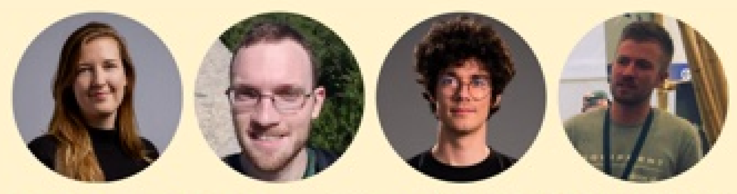
Uncovering Anomalous Events for Marine Environmental Monitoring via Visual Anomaly Detection

Laura Weihl*[1,4], Stefan H. Bengtson[2,4], Nejc Novak[3], Malte Pedersen[2,4]

IT University of Copenhagen[1], Aalborg University[2], Anemo Robotics ApS[3], Pioneer Center for Artificial Intelligence[4] - Denmark



ICCV OCT 19-23, 2025 HONOLULU HAWAII

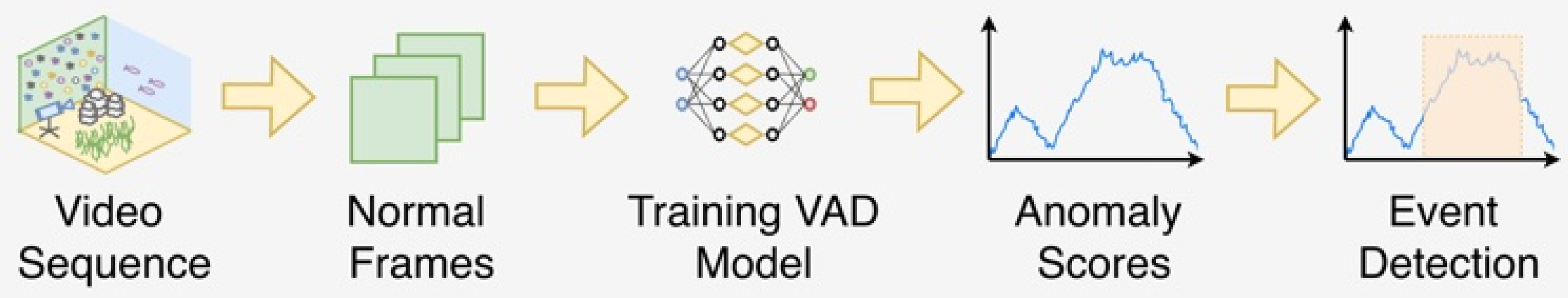
Dataset Full Paper *lawe@itu.dk

Motivation

- For **reliable biodiversity monitoring** we need underwater video cameras but they record vast sections of "nothing", making manual annotation impractical.
- **Difficulty: event annotation is subjective** as visibility varies underwater (turbidity, contrast, lighting), so single-annotator labels are unreliable.
- We train **visual anomaly detection** (VAD) methods on a marine scenes to find "interesting events" e.g. marine animal appearances/movements.
- **Contributions:** AURA multi-annotator dataset, VAD benchmarking, and an evaluation of robust frame-selection for anomalous event extraction.

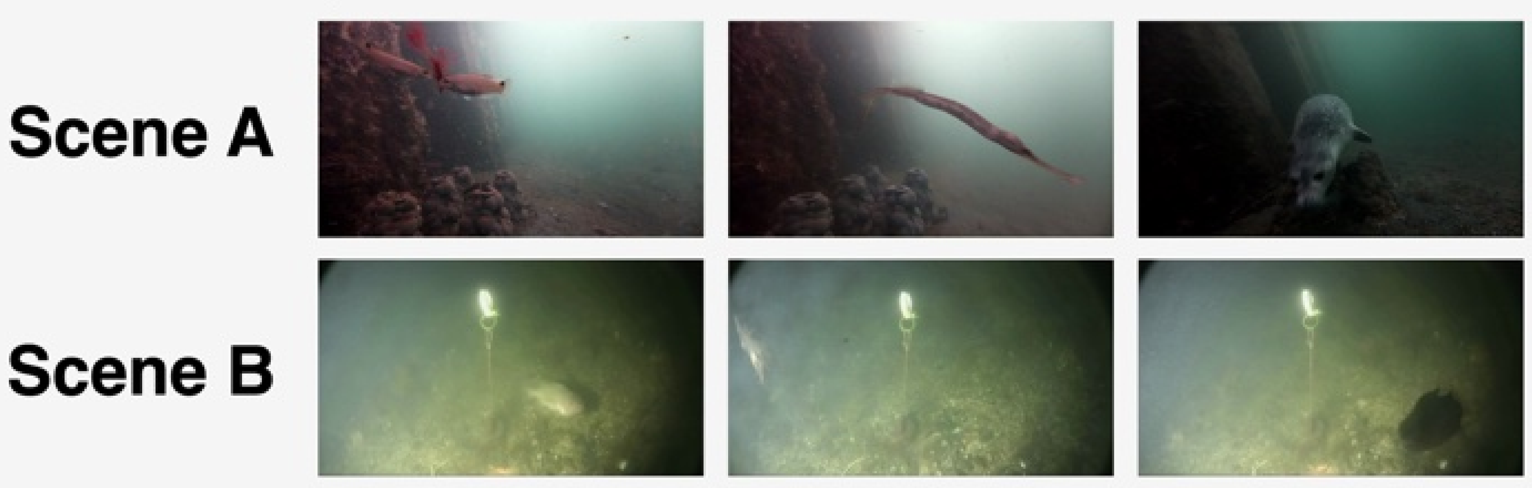
Visual Anomaly Detection (VAD)

- **Objective:** to detect unusual or unexpected visual appearances or motion.
- **Applications:** industrial inspection, security surveillance, medical domain.
- **VAD-based event detection:** train VAD model on normal frames of a scence and find peaks in sequences of anomaly scores in images extracted from video data.



The AURA Dataset: Anomalous UnderwateR Activity

- We introduce **the first multi-annotator underwater VAD** dataset
- **16 annotators, 25 videos, 2 scenes:** Scene **A**nemo and Scene **B**rackish [1], each with a single anomalous event

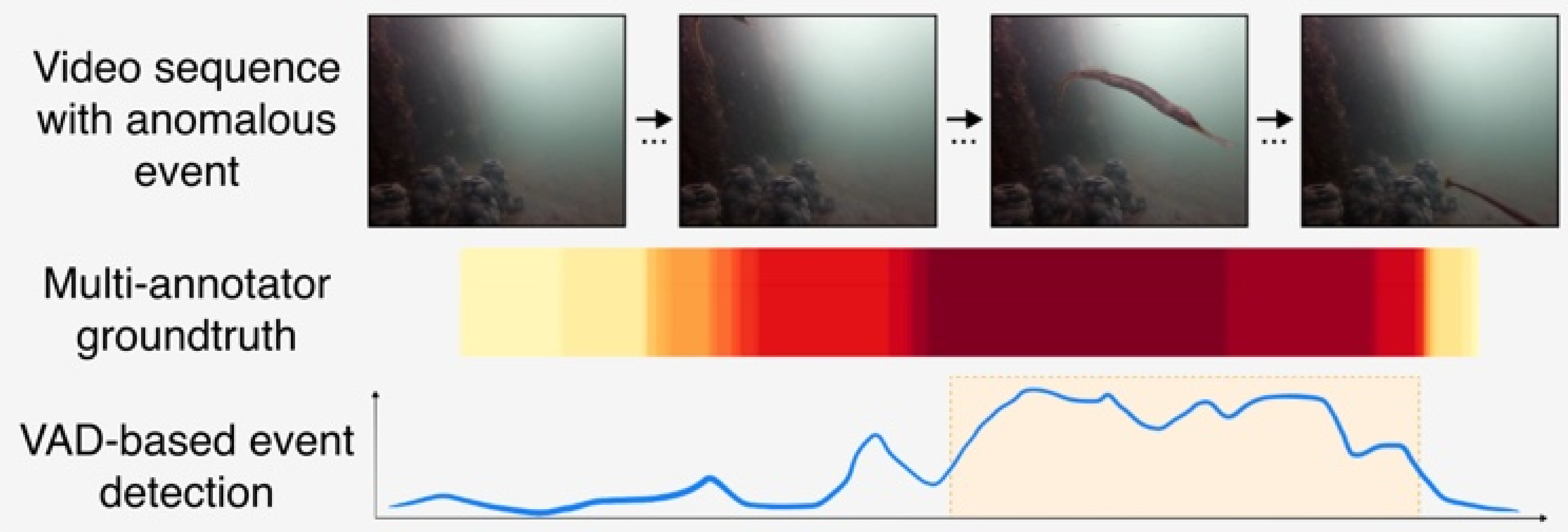


Anomalous Events Underwater

Learn the "normal" scene. Uncover the moments that matter.

Evaluation

- **Setup**: Evaluate four VAD models on two scenes with two train splits
- **Annotation Alignment**: mean absolute error vs. soft labels and temporal intersection over union (t-IoU) vs. consensus labels
- **Event Detection**: comparing peak-based event detection vs simple thresholding



Results

- **Annotation Alignment**: Mean absolute error vs soft labels on large training set (↓)

	Scene A	Scene B
EfficientAD [2]	0.393 ± 0.101	0.379 ± 0.048
GANomaly [3]	0.425 ± 0.127	0.435 ± 0.099
STFPM [4]	0.465 ± 0.107	0.320 ± 0.058
Reverse Distillation [5]	**0.271** ± 0.151	**0.259** ± 0.041

- **Event Detection**: Average t-IoU per frame selection method (↑)

	Scene A		Scene B	
	Find-peak	Threshold	Find-peak	Threshold
EfficientAD [2]	0.347	0.374	0.745	0.709
GANomaly [3]	0.242	0.083	0.471	0.217
STFPM [4]	0.301	0.225	**0.863**	0.748
Reverse Distillation [5]	**0.673**	**0.461**	0.861	**0.806**

→ VAD scores from **Reverse Distillation** align best with human annotation.

Conclusion

- **Event Detection:** Peak-based method shows higher t-IoU than thresholding
- **Reverse Distillation as strong baseline:** more data and consistent scenes help
- **AURA Dataset:** Scene A more challenging for VAD models than Scene B
- **Need for multi-annotated data:**
  - VAD model performance varies depending on ground-truth
  - single annotation is highly subjective
- **Future Work:** expand sites, longer sequences with multiple anomalous events

References

[1] Pedersen, M., Bruslund Haurum, J., Gade, R. and Moeslund, T.B., 2019. Detection of marine animals in a new underwater dataset with varying visibility. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 18-26).

[2] Batzner, K., Heckler, L. and König, R., 2024. EfficientAD: Accurate visual anomaly detection at millisecond-level latencies. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 128-138).

[3] Akcay, S., Atapour-Abarghouei, A. and Breckon, T.P., 2018, December. GANomaly: Semi-supervised anomaly detection via adversarial training. In Asian Conference on Computer Vision (pp. 622-637). Cham: Springer International Publishing.

[4] Wang, G., Han, S., Ding, E. and Huang, D., 2021. Student-teacher feature pyramid matching for anomaly detection. arXiv preprint arXiv:2103.04257.

[5] Deng, H. and Li, X., 2022. Anomaly detection via reverse distillation from one-class embedding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9737-9746).

Anemo Robotics | IT UNIVERSITY OF COPENHAGEN | REMARO | European Commission | PIONEER CENTRE FOR ARTIFICIAL INTELLIGENCE | AALBORG UNIVERSITY