Learning Maritime Obstacle Detection from Weak Annotations by Scaffolding

Lojze Žust and Matej Kristan

University of Ljubljana, Faculty of Computer and Information Science

1 Maritime obstacle detection

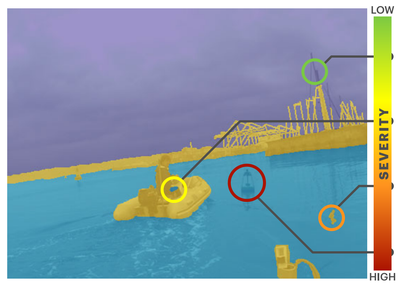
Obstacles:
- Static obstacles (e.g. land, piers)
- Dynamic obstacles (e.g. boats, buoys)

State-of-the-art: **semantic segmentation** [1]

Challenges:
- object reflections, FP detections
- signifficant annotation effort

2 Not all segmentation errors are equal

Similar sized segmentation errors - different severity for detection:



Missing details in the sky
Not important for obstacle detection

"Holes" in large objects
Minor - object is detected and can be avoided

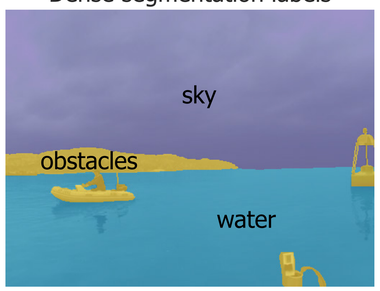
False positive detections in the water
Problematic - can cause frequent slow-downs

Failure to detect objects
Critical - might lead to collisions

3 Detection-oriented annotations

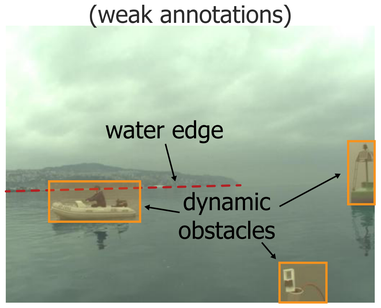
Replace per-pixel labels with detection-oriented annotations [2]:
- Static obstacles → water-obstacle boundary (**water-edge**)
- Dynamic obstacles → **bounding boxes**

Dense segmentation labels



20 minutes / image
segmentation oriented
detection-irrelevant details

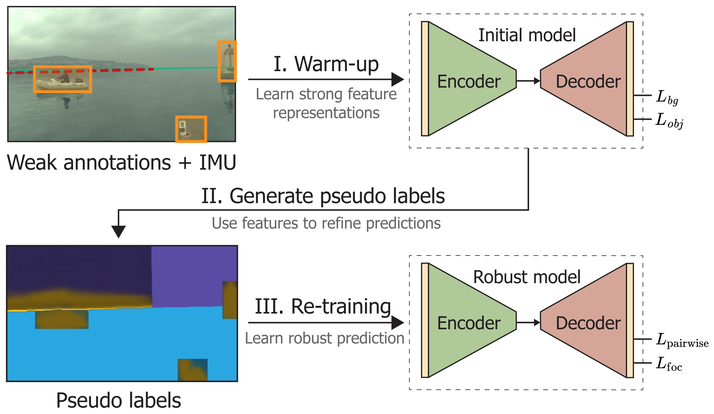
Detection-oriented annotations (weak annotations)



1 minute / image
detection oriented
only relevant information

4 Scaffolded Learning Regime (SLR)

Training **semantic segmentation** models from **weak annotations**:

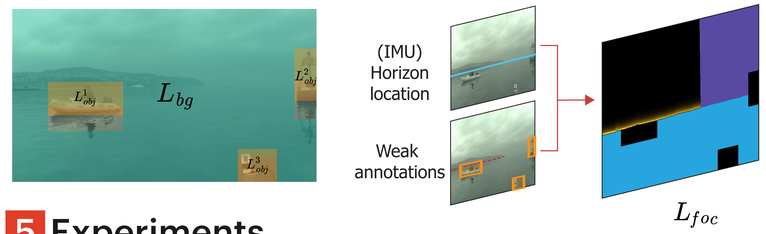


Step I: Feature warm-up

Combination of **object-wise** and **background-wise** objectives

$L_{obj} = L_{proj} + L_{pairwise}$ [3]
Applied to each annotated object (obstacle)

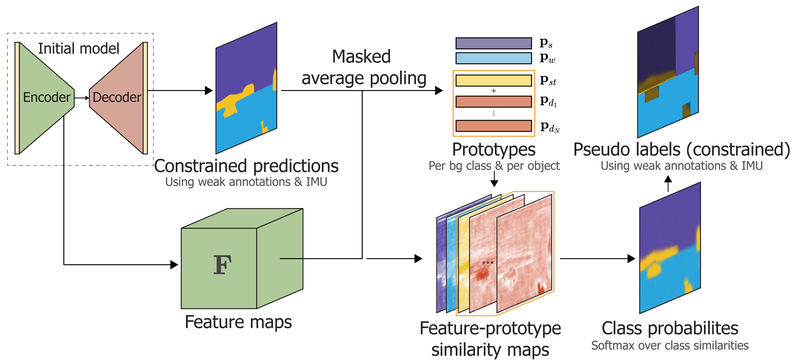
$L_{bg} = L_{foc} + L_{pairwise}$
Applied to background pixels (non-object)



Step II: Pseudo-labels generation

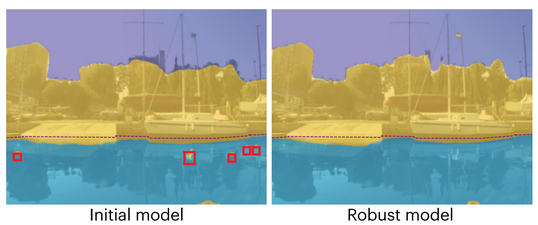
Initial model: strong feature representations, overconfident predictions

Idea: refine predictions using strong features



Step III: Re-training

Re-train the model with the pseudo labels



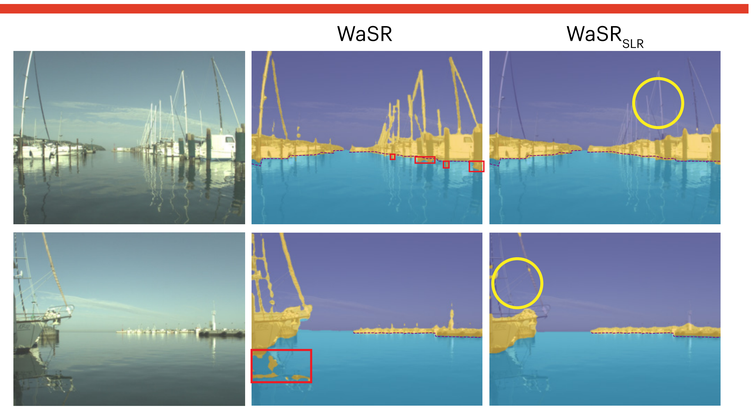
5 Experiments

All models trained on the MaSTr1325 [4] dataset.
- **Traditional**: dense segmentation labels, **SLR**: weak annotations

Evaluation on the MODS [2] benchmark:

		Static obstacles (Segmentation error) A_W	Dynamic obstacles (Detection F1-score) F1 (overall)	F1 (danger zone)
Traditional	RefineNet	11	91.0	61.8
	BiSeNet	**10**	90.2	69.1
	DeepLabV3	**10**	86.0	31.3
	WaSR	11.0	92.6	75.0
Our	$\text{DeepLabV3}_{\text{SLR}}$	11.6 (+1.6)	90.5 (+4.5)	83.3 (+52.0)
	WaSR_{SLR}	11.5 (+0.5)	**93.2** (+0.6)	**85.9** (+10.9)

- **Superior obstacle detection**, especially inside danger zone (<15m)
- Comparable water-edge segmentation performance
- Despite using **signifficantly simpler annotations** (~20x faster)



References

[1] Bovcon, B., & Kristan, M. (2021). WaSR--A Water Segmentation and Refinement Maritime Obstacle Detection Network. IEEE Transactions on Cybernetics, 1–14
[2] Bovcon, B., Muhovič, J., Vranac, D., Mozetič, D., Perš, J., & Kristan, M. (2021). MODS -- A USV-oriented object detection and obstacle segmentation benchmark.
[3] Tian, Z., Shen, C., Wang, X., & Chen, H., BoxInst: High-Performance Instance Segmentation with Box Annotations. CVPR 2021
[4] Bovcon, B., Muhovič, J., Perš, J., & Kristan, M., The MaSTr1325 dataset for training deep USV obstacle detection models. IROS 2019

Code & weights

IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022

https://github.com/lojzezust/SLR
{lojze.zust, matej.kristan}@fri.uni-lj.si