Large-Scale Marine Debris Detection with Sentinel-2

Marc Rußwurm, Sushen Jilla Venkatesa, Laura Pasero, Dilge Gül, Devis Tuia

WAGENINGEN UNIVERSITY & RESEARCH · SDSC · TU Delft · EPFL

Motivation

1) marine litter pollutes our environment; Sargassum and Algae blooms are harmful to ecological life

Cuttings Beach, Durban (South Africa)
Image: Lisa Guastella

2) single campaigns collect marine litter at small scale e.g., Ruiz et al., 2022

Bay of Biscay, France
Image: Oihane Basurko

3) lack of large-scale satellite-bad detection methods limits collection efforts

even though

an abundance of satellite data is freely available:

Sentinel-2, PlanetScope

Research question:

How to use modern machine learning for automatic detection of generic marine debris with publicly available satellite imagery?

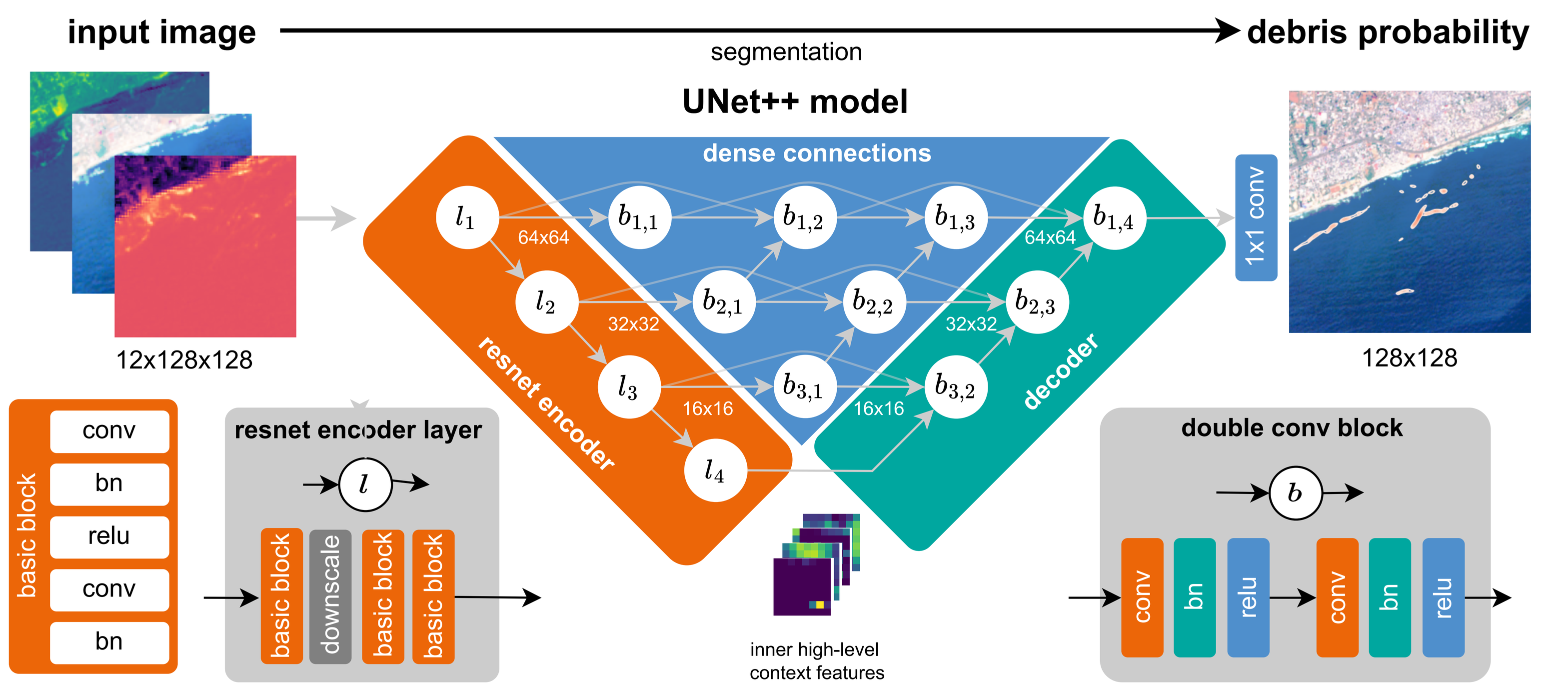
Approach:

Explore Data-Centric Machine Learning (DCAI) principles for large-scale marine debris detection

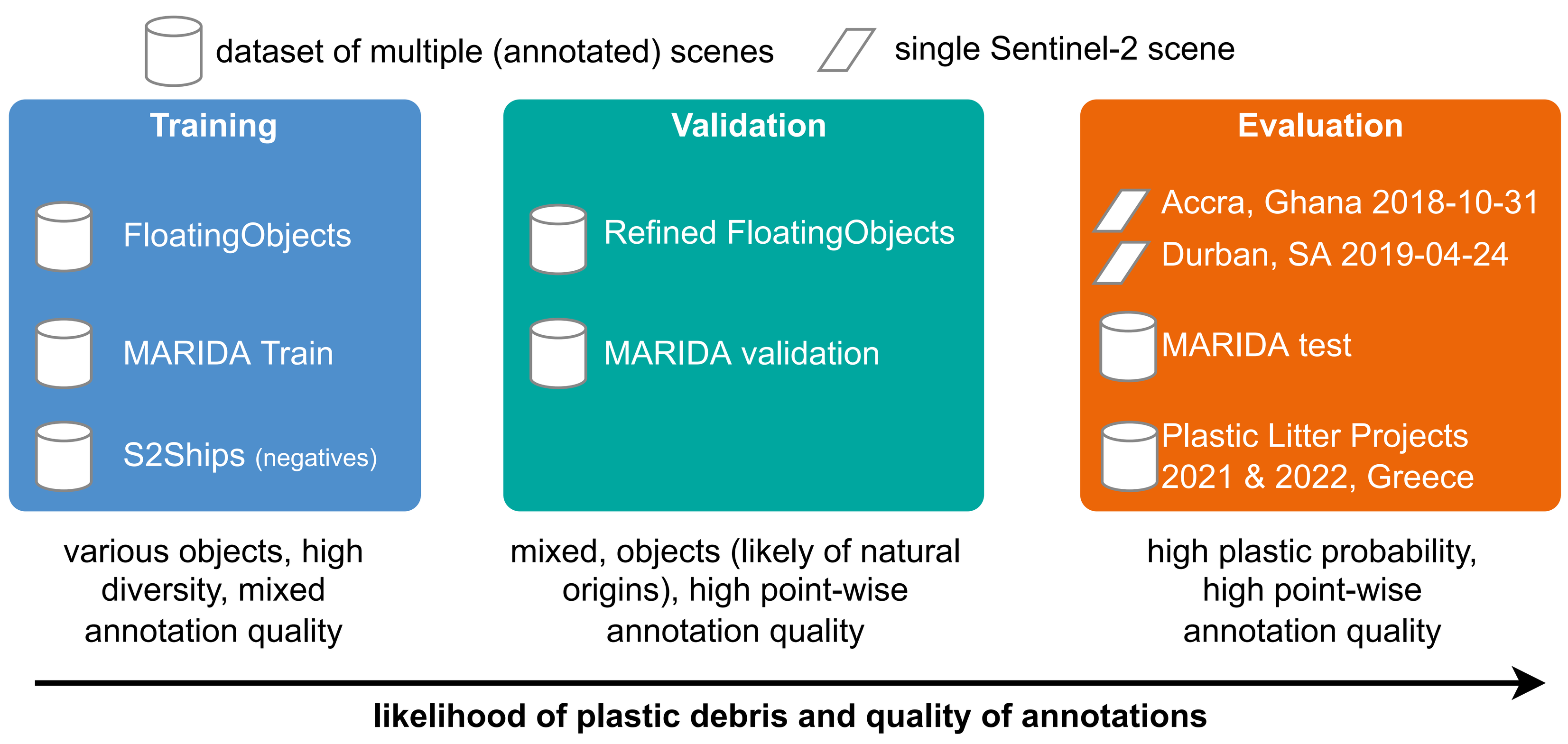
Method & Data

UNet and UNet++ segmentation models

Estimate a probability for "marine debris" for each image pixel



trained on a quantity-; evaluated on quality-focused datasets



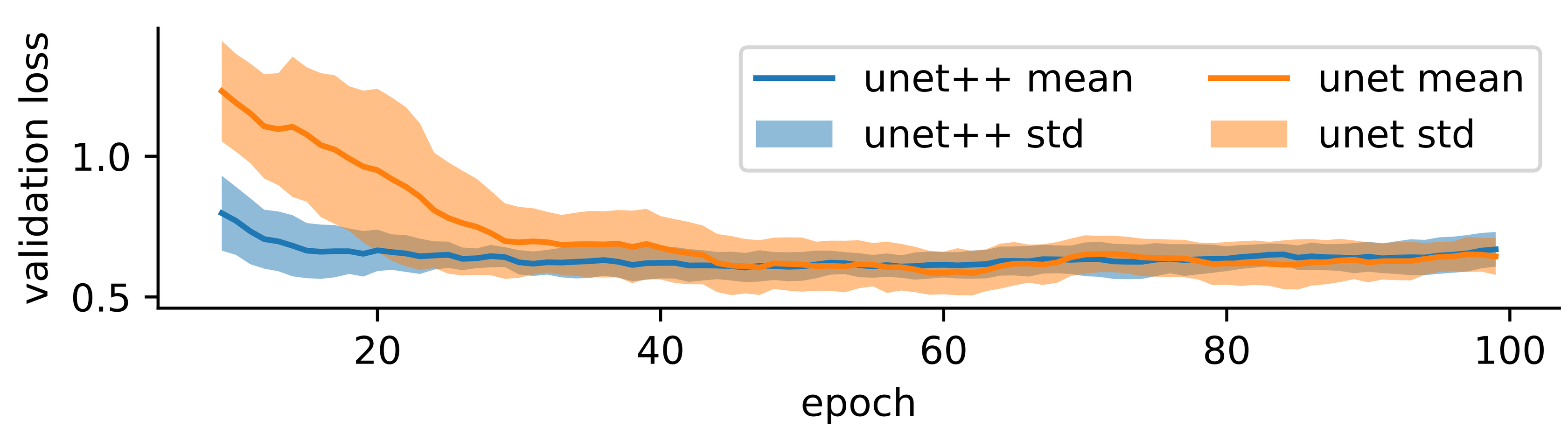
Takeaways

In accordance with data-centric machine learning

six principles according to Jarrahi et al., 2023:

1) data-fit 2) data consistency 3) iterative process 4) reflect local context
5) meet needs of local stakeholders 6) exchange between ML and domain experts

both UNet++ and UNet deep learning models result in similar accuracy:

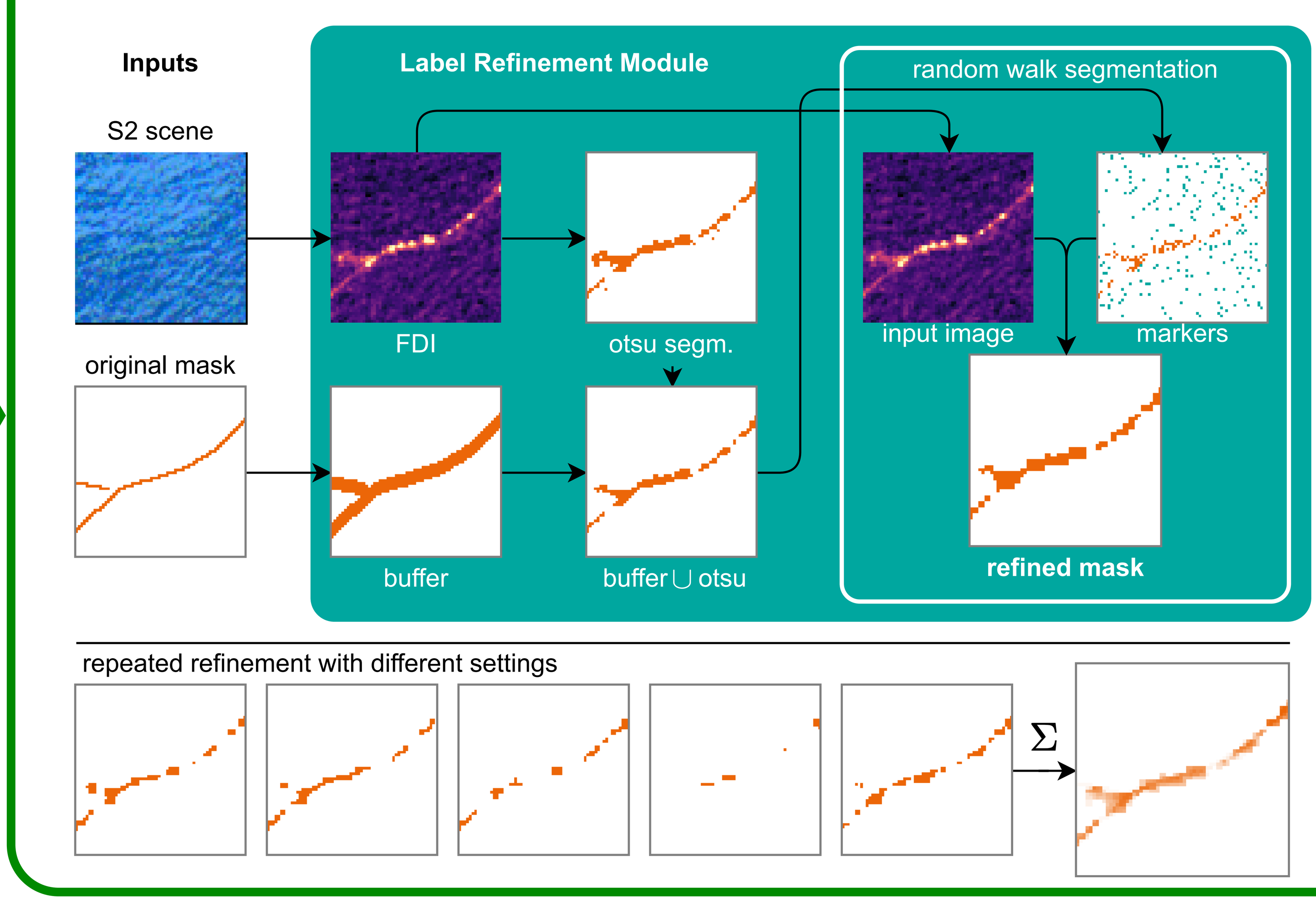


automated label refinement had greater impact on accuracy (see no-ref)

Marida-test set

trained on	original data		our train set			
	RF	UNET	RF	UNET	UNET++	UNET++ no-ref
ACCURACY	0.697	0.838	0.811	**0.865** ± 0.006	**0.867** ± 0.005	0.851 ± 0.006
F-SCORE	0.288	0.701	0.708	**0.741** ± 0.012	**0.749** ± 0.009	0.710 ± 0.015
AUROC	0.488	0.764	0.862	**0.738** ± 0.012	**0.746** ± 0.021	0.733 ± 0.006
JACCARD	0.168	0.539	0.548	**0.589** ± 0.015	**0.598** ± 0.012	0.551 ± 0.018
KAPPA	0.197	0.593	0.569	**0.654** ± 0.016	**0.661** ± 0.012	0.615 ± 0.017

Example: Improving data-fit (principle 1) with label refinement



Contact

Marc Rußwurm Wageningen University
Assistant Professor ML & RS

marc.russwurm@wur.nl

Model & Sources

github.com/marccoru/marinedebrisdetector

References

- Rußwurm M, Venkatesa J. S., Tuia D. (2023) Large-scale Detection of Marine Debris in Coastal Areas with Sentinel-2. To appear in Cell iScience.
- Mifdal, J., Longépé, N., & Rußwurm, M. (2021). Towards detecting floating objects on a global scale with learned spatial features using sentinel 2. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 285-293.
- Ruiz, I., Basurko, O. C., & Rubio, A. (2022). Modelling the distribution of fishing-related floating marine litter within the Bay of Biscay and its marine protected areas. Environmental Pollution, 292, 118216.
- Jarrahi, M. H., Memariani, A., & Guha, S. (2023). The Principles of Data-Centric AI. Communications of the ACM