

Deployment-Based Hard-Negative Stress Test of Marine Debris Detection Using Water-Dependent Negative Examples for Resource-Constrained ASVs

Advaith Rajagopalan
BASIS Independent Silicon Valley
San Jose, California, USA
rajagopalan.advaith@gmail.com

Abstract—Verifying the presence of floating debris is vital for low-cost autonomous surface vehicles (ASVs) involved in early marine pollution detection. Although traditional image-based analysis is a good starting point for verification tasks, there are specific difficulties associated with the process of operating under real-life water-surface conditions, which include seaweed, surface glare, far-away structures, and low-contrast images. This paper presents a deployment-based evaluation of the YOLOv8n-based RGB debris verifier specifically built for low-cost ASVs. We trained a YOLOv8n model on a publicly available bottle-oriented dataset. Our stress-test analysis is based on a specially reconstructed dataset of 145 images. On that data set, our detector showed average values of precision = 0.947, recall = 0.930, mAP@0.5 = 0.960, and mAP@0.5:0.95 = 0.831. Alongside these overall scores, we provide per-category statistics and confidence-threshold sweep data in order to show strengths and failures of the verifier. We observed 0.977 precision and 0.959 recall on the bottles-only subset; 0.610 and 0.667 false positives per image on the open-ocean and seaweed subsets, respectively; and no false positives on the limited 5-image glare subset. The confidence-threshold sweep allowed us to show how precision and recall can be tuned, which can be used for further deployment decisions regarding the trade-off between false-positive and missed-debris costs. Each per-category finding discussed below suggests a corresponding deployment or architectural mitigation.

Index Terms—marine debris detection, autonomous surface vehicle, hard-negative evaluation, per-category analysis, object detection, marine robotics, YOLOv8.

I. INTRODUCTION

Marine plastic pollution is one of the largest environmental problems [1]–[3]. After being exposed to ultraviolet radiation and other factors, floating plastics can fragment into smaller pieces, making them less visible and harder to clean up [4]. As this process negatively impacts the whole ecosystem of marine organisms, we consider it a strong engineering reason to develop a system capable of detecting floating objects while still visible.

Autonomous surface vehicles (ASVs) operate at a middle ground between satellites or aerial vehicles and fixed or manual collecting equipment. Practically, verifying the presence of floating debris is a tough task. There are many optical and physical aspects of water environments that may confuse a verification device, such as seaweed, surface glare, distant shoreline objects, and low-contrast backgrounds. Ex-

isting research in the area of automatic satellite and marine robotics shows promise for monitoring of marine debris and the necessity to verify debris under marine conditions [5], [6].

While general image-based evaluation is helpful for testing an object detector, aggregate metrics obtained on a typical test split do not necessarily reflect real-life deployment behavior. In the context of autonomous ASV operations, we need to differentiate a situation where a false positive on featureless water leads to maneuvering and a situation when the detector identifies seaweed as a piece of debris. While aggregate measures would lump those situations together, per-category analysis helps to distinguish the two cases.

In this paper, we evaluate a detector in the context of marine operations through stress testing on a hard-negative environment. In particular, we propose:

- Reconstruction of an evaluation dataset for detecting marine debris featuring marine hard-negative samples;
- Per-category analysis which makes aggregate metric evaluation less relevant and allows us to make useful observations on detector behavior;
- Confidence-threshold analysis which enables us to quantify trade-offs during deployment;
- Mapping from identified issues to corresponding solutions in the context of the multi-view verification architecture; and
- Discussion of potential multimodal detector extensions using thermal-delta information in addition to RGB images.

Our key contribution is the evaluation strategy focused on deployment requirements, which allows us to transform abstract accuracy to actionable insights about a detector's strengths and weaknesses.

II. LOW-COST ASV VERIFICATION ARCHITECTURE

A. Verification Process Design Considerations

We are interested in a low-cost debris verification system based on an ASV. Therefore, the verification process design involves consideration of compute, power, and sensing capacity constraints. Rather than running constant inference on all frames, we employ a staging process, in which candidate locations are first identified by the search process, RGB images

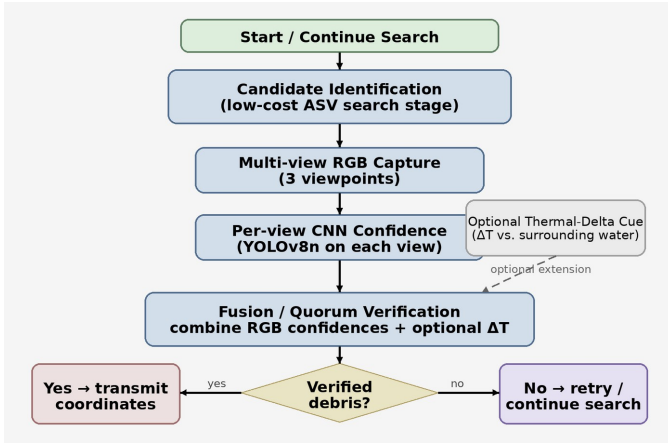


Fig. 1. Workflow of ASV debris verification with an RGB detector. Candidate locations are verified using a 3-view RGB verifier; per-view confidences are aggregated in a quorum-based decision, with the thermal-delta cue shown optionally as a potential supplement to the verification process.

are taken, and detector confidences are then combined in a quorum-based decision step.

Such a low-cost ASV can run verification tasks with the help of a Raspberry Pi-based computer, which imposes limitations on the type of detector. As opposed to continuous full-frame inference, we are concerned about verifying debris with a lightweight detector on a few selected images. This paper focuses not only on the detector itself, but on the feasibility of performing such a verification task using the YOLOv8n-based model.

B. High-Level Verification Architecture

In Fig. 1, the architecture of our verification process is demonstrated. Starting with a candidate debris location selected previously, RGB views of this location are taken from different angles and analyzed by an RGB detector, whose confidences are combined in a quorum-based decision block. Thermal-delta is included as an additional cue to be used in multimodal extension.

Notably, this multi-view process is non-trivial since, as shown in Section VI, precisely viewpoint-dependent error cases such as silhouette views of seaweed are the ones the quorum decision should reject.

III. METHODS

A. RGB Verifier Backbone and Training Setup

As for the RGB verifier, YOLOv8n, the nano version of the YOLOv8 family, was chosen due to its lightweight design, making it applicable for robotic applications [7]. YOLOv8n was used as a lightweight baseline rather than as a claim of using a state-of-the-art detector, which is out of scope here. Its choice is based on the light architecture, widespread use, and the availability of advanced training and inference frameworks, which make this detector a good candidate for deployment-oriented experiments on ASVs. The comparison with recent lightweight detectors and later YOLO variants

TABLE I
CLASS COMPOSITION OF THE RECONSTRUCTED ON-WATER TEST DISTRIBUTION, TOTALING 145 IMAGES.

Category	Images	Inst.	Role
Bottles (original test split)	81	88	Positive cases
Bottles (new, diversity)	9	9	Positive, env. variety
Open ocean, no debris	41	0	Hard negatives
Seaweed	9	0	Known FP mode
Glare / reflection	5	0	Reflection FP mode
Total	145	97	—

is left as future work. As for the training data, a publicly available dataset of bottles was chosen. To train the model, pre-trained YOLOv8n weights were used together with 1,343 images for training and 258 for validation taken from the `bottles-svzvd-w6q4v` dataset in Roboflow Universe [8] and formatted into a YOLO-compatible dataset. Training was done using 50 epochs with default image size 640 and Ultralytics augmentations. Configurations included Ultralytics 8.3.165, Python 3.11, PyTorch 2.6.0+cu124, and an NVIDIA Tesla T4 GPU. Evaluation was done using Ultralytics 8.4.36 without any fine-tuning to account for the hard negatives found among the test images.

Model selection reflects deployment expectations. Whereas a heavier detector would likely perform better, the current experiment emphasizes deployment in compute-constrained ASV pipelines. In the illustrative workflow, the candidate inspection process is multi-view rather than frame-based: three RGB views are analyzed, and the aggregate confidence is computed through a quorum mechanism.

B. Reconstructed On-Water Test Distribution

Unlike in standard machine-learning experiments, testing is not limited to the exported test distribution, which contains only marine debris imagery from a particular perspective. Rather, the test distribution is recreated to include various marine scenarios that may occur in ASV operations, but retaining the original positive cases. Extra images are drawn from the cited Roboflow dataset, openly licensed Wikimedia/iNaturalist/Unsplash image sources, and an existing USV obstacle-detection dataset [6], [8]–[11]. Per-class metrics are computed alongside aggregate metrics in order to attribute failure modes to specific classes (marine conditions). Table I summarizes the class composition of the reconstructed on-water test distribution.

C. Evaluation Protocol

Both aggregate and per-bottles metrics are computed using Ultralytics’ `model.val()` interface on ground-truth annotations, obtaining precision, recall, `mAP@0.5`, and `mAP@0.5:0.95`. Because the open-ocean, seaweed, and glare subsets have no examples of debris, ground-truth annotations are unavailable, and therefore cannot be scored using `model.val()`. For these negative-only subsets, evaluation

TABLE II
AGGREGATE PERFORMANCE ON THE RECONSTRUCTED ON-WATER TEST SPLIT.

Split	P	R	mAP@.5	mAP@.5:.95
Test (145 images)	0.947	0.930	0.960	0.831

was carried out using `model.predict()`, where the number of predicted boxes counts as false positives. This allows us to obtain the easily interpretable false-positives-per-image metric. The confidence-threshold sweep experiment outlined in Section IV-C is conducted on all 145 test images using the same pipeline for three confidence thresholds: 0.25, 0.50, and 0.75.

IV. RESULTS

A. Aggregate Performance on the Stress-Tested Split

On the reconstructed 145-image split, the detector shows strong aggregate performance despite being evaluated on a test distribution explicitly extended to include the marine conditions likely encountered in ASV operations (see Table II). Average per-image execution time on the evaluation GPU is 3.2 ms of preprocessing, 6.6 ms of inference, and 1.3 ms of post-processing.

These metrics indicate the competence of the verifier but do not specify which classes are responsible for residual errors. Per-class evaluation will address this problem.

B. Per-Class Performance Analysis

In Table III, performance is reported separately for each test-set class. Two evaluation modalities are employed for each class. For the bottles subsets, which have ground-truth annotations, the standard precision, recall, and mAP@0.5 are computed. For the negative-only subsets (open ocean, seaweed, glare), no instances are available for scoring against ground-truth annotations; hence, the table reports the total number of predicted false positives, along with the resulting false-positives-per-image rate.

Bottles (P=0.977, R=0.959). On actual debris imagery, the verifier exhibits excellent performance: precision and recall both exceed 0.95, implying effective detection of bottles if present, as well as a very low rate of spurious bounding boxes. This is the core competency of the model that corresponds to the test-set distribution.

Open ocean (FP/image = 0.610). On 41 featureless/wave-textured images devoid of any debris, the detector predicted 25 false-positive bounding boxes, implying that approximately one false detection occurs for every two images. This result suggests that in practice, each such hallucination could lead to an unnecessary approach of an unoccupied region, consuming energy and time. It appears that the source of hallucinations is wave textures, horizon, and/or specular lighting effects that simulate the bottle features learned during training.

Seaweed (FP/image = 0.667). With a small subset size of 9 images, the detector produced 6 false positives. Such results

were expected and confirm that seaweed is a hard-negative class. Seaweeds observed at oblique angles tend to produce dense, high-contrast silhouettes similar in shape to debris. Because this subset contains only 9 images, the seaweed result should be interpreted carefully; further evaluation on a larger seaweed dataset is needed.

Glare (FP/image = 0.000). All 5 images in this subset generated no predictions by the detector whatsoever. This is a favorable outcome that does not support, in this small subset, the assumption that glare necessarily causes false positives in water-surface analysis. It appears that the model has learned features that distinguish specular effects from boundaries of solid objects. We should emphasize that this conclusion pertains only to the small sample size of 5 images; a single false positive would raise this metric to 0.200. It may be interesting to investigate a larger dataset of glare images in the future.

Hence, per-class metrics make it clear that although aggregate metrics indicate satisfactory performance, they also hide different failure modes. The verifier is precise on actual debris but prone to making false-positive detections on both featureless water and seaweed. Those are two different failure modes, which, as discussed in Section VI, could require different solutions.

C. Confidence-Threshold Sweep Analysis

Another experiment aims to clarify the precision/recall trade-off by varying the confidence threshold between three settings (see Table IV).

Precision and recall values coincide for thresholds $\text{conf}=0.25$ and $\text{conf}=0.50$. This indicates that the majority of predictions consist of confident true positives or weak false positives that would fall below the 0.25 threshold; predictions between 0.25 and 0.50 are rare. When setting the threshold to $\text{conf}=0.75$, the precision increases to 0.965 while recall drops to 0.856, revealing that a more conservative detector leads to lower detection rates with fewer false positives.

This experiment illuminates deployment trade-offs. A cost-effective ASV running autonomously must find a balance between false positives (unnecessary approaches) and false negatives (uninspected debris). A conservative strategy would employ $\text{conf}=0.75$ to avoid wasted maneuvers, while a thorough coverage strategy may opt for $\text{conf}=0.25$ to leverage multi-view voting for suppressing false positives. The confidence-threshold sweep helps mitigate the open-ocean problem revealed in Section IV-B: false positives in the open ocean tend to correspond to low confidence and could be reduced in the confidence-threshold sweep.

D. Failure Mode Examples

In Fig. 2, representative held-out cases illustrating failure modes discussed in the preceding section are provided.

Two major false-positive patterns were observed qualitatively. First, seaweed bunches in oblique and shallow views may have silhouettes reminiscent of bottle-like objects. Second, shoreline or artificial structures appearing in the upper

TABLE III
PER-CATEGORY PERFORMANCE ON THE RECONSTRUCTED TEST SPLIT. BOTTLES SUBSETS ARE SCORED AGAINST GROUND-TRUTH ANNOTATIONS;
NEGATIVE-ONLY SUBSETS REPORT FALSE-POSITIVE BOX COUNTS BECAUSE NO POSITIVE INSTANCES EXIST.

Category	Images	Instances	Precision	Recall	mAP@0.5	FP count	FP / image
Bottles (all)	90	97	0.977	0.959	0.964	—	—
Open ocean	41	0	—	—	—	25	0.610
Seaweed	9	0	—	—	—	6	0.667
Glare / reflection	5	0	—	—	—	0	0.000
Full set	145	97	0.947	0.930	0.960	—	—

TABLE IV
CONFIDENCE-THRESHOLD SWEEP ACROSS THE WHOLE 145-IMAGE TEST DISTRIBUTION. MAP VALUES IN THIS TABLE ARE RECOMPUTED BY ULTRALYTICS’ EVALUATION PIPELINE AFTER APPLYING THE LISTED CONFIDENCE PRE-FILTER AND SHOULD NOT BE INTERPRETED AS A STANDARD CONFIDENCE-INDEPENDENT AP CURVE.

Conf. thresh.	P	R	mAP@.5	mAP@.5:.95
0.25	0.947	0.930	0.956	0.816
0.50	0.947	0.930	0.940	0.806
0.75	0.965	0.856	0.854	0.770

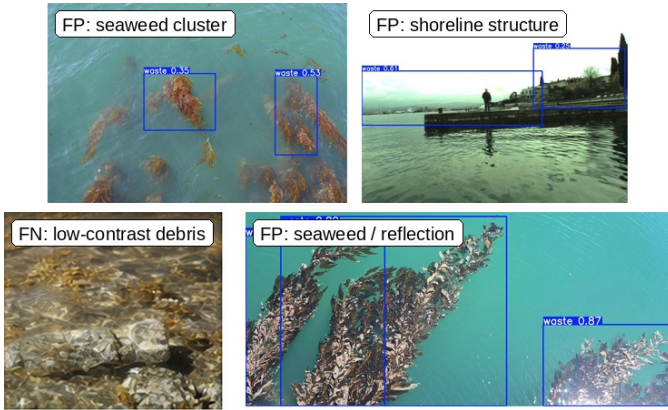


Fig. 2. Failure instances in each category include seaweed-based false positives, shoreline-structure false positives, and low-contrast false negatives. This qualitative categorization agrees with the quantitative per-category statistics reported in Table III. These images are from the cited datasets and image sources [6], [8]–[11].

part of the images from a distance can have a geometry similar to a learned bottle. The primary false-negative pattern arises when debris is camouflaged against low-contrast backgrounds.

V. THERMAL-DELTA EXTENSION

It follows naturally from the results presented in this paper to develop a multimodal fusion of RGB confidence and thermal-delta information, which stands for the temperature difference between a suspicious region and surrounding water bodies. Prior maritime sensing work has examined visible-light and thermal imagery for detecting objects at the sea surface [12]. Due to the possible disparity between thermal characteristics of water and debris, the thermal-delta signal

could potentially reduce visual false positives. This hypothesis is left for experimental verification in future work.

VI. DISCUSSION

A. What do the per-category results show?

Overall, as evident from Table II, the RGB-based verifier works well under stress-test conditions. However, the per-category analysis in Table III reveals that the error incurred by the model is concentrated in two categories (open-ocean hallucination and false positives due to seaweed), not observed in the limited glare subset, and limited for actual debris. In other words, per-category analysis provides insight into what the residual model error is and what it is not. Reporting an aggregate precision/recall evaluation of the verifier in lieu of per-category metrics would fail to capture the most significant result for low-cost ASV deployment: the rate of open-ocean hallucinations.

B. Relevance to low-cost ASV operation

From the ASV’s point of view, precision and recall represent different mission costs. A false positive instance on featureless water incurs the cost of an unnecessary approach maneuver, which drains battery power and time the platform cannot recover. Assuming the rate of 0.6 hallucinations per featureless-water image, an unmitigated single-camera verifier would cause frequent false approaches. On the contrary, a false negative means the absence of confirmation for a real debris candidate, i.e., the mission expends approach resources without a recorded detection. Since these two errors impose different costs to mission success, it is important to characterize them independently and per category to take informed architectural decisions regarding confidence thresholds, voting policies, and training-data augmentation.

C. Architectural responses to failure modes

One key advantage of per-category analysis is its direct connection to architectural responses, each of which is discussed below.

False positives on featureless ocean – confidence thresholding. From the thresholding experiments performed in Section IV-C, it follows that raising the operating confidence from 0.25 to 0.75 increases precision but reduces recall by about 7 percentage points. As an improvement to the energy-constrained mission, the trade-off is worth considering. The

threshold is a deployment parameter, not a fixed model property.

False positives on seaweed – multi-view quorum. False positives caused by seaweed are fundamentally viewpoint-dependent. Indeed, while a camera sees the deceptive silhouette of an oblique seaweed bunch as a bottle-like object, another camera from a different viewpoint might miss this trick. Specifically, the three-view quorum designed for the verification pipeline (Fig. 1) is intended to serve as a safeguard against the issue. A good sign is that this mitigation does not involve training a more complex classifier; it is purely architectural.

Glare behavior in the limited sample – no glare-specific mitigation indicated by this limited sample. The glare negative subset produced no false positives in the 5-image sample, suggesting that the learned features may distinguish these specular reflections from solid objects. No additional glare-specific mitigation was indicated by this limited sample, allowing system design efforts to focus on failure modes that more clearly require mitigation. This result should not be interpreted as general glare robustness without evaluation on a larger glare dataset.

Low-contrast false negatives – training-data expansion. The final category, false negatives due to debris being camouflaged in a cluttered or sediment-rich background, cannot be addressed by means of either confidence thresholding or voting. This is a feature-salience issue, not viewpoint or confidence; the cameras cannot reach consensus about something that they do not see. Thus, the architectural response to this type of failure involves expanding the training dataset with images containing background clutter and sediment. This motivates future training-data expansion with cluttered, low-contrast, and sediment-rich water backgrounds.

D. Limitations and future work

While this experiment is focused on the evaluation of the RGB verifier, future work should deal with the multimodal thermal-delta extension and thorough target-platform evaluation. Furthermore, the use of smaller subsets, such as 9 images in the case of seaweed and 5 in the case of glare, might limit how confidently each category can be evaluated; a bigger number of images in each category should be used as well. In addition, this study is limited to the detection of bottles, meaning that it is not yet clear whether the verifier generalizes to bags, foam, fishing gear, and other types of debris. Future work might consider the following research directions: increasing the diversity in training class distribution, adding water-surface hard-negative examples, comparing YOLOv8n with other lightweight detector architectures, and performing end-to-end ASV experiments that include onboard timing and realistic water-surface conditions.

VII. CONCLUSION

In this paper, we evaluated a lightweight RGB verifier of marine debris based on the YOLOv8n model, in the context of marine environmental monitoring. To approach

the framework of the deployment-oriented diagnosis of the verifier, we reconstructed the test set to include hard negatives in featureless water areas, seaweed, and glare. Analyzing the results per-category rather than aggregating them yielded more insights into the performance and failure modes. The verifier demonstrated impressive performance on real-life debris ($P = 0.977$, $R = 0.959$), produced no false positives on the limited glare sample, and suffered from false positives both on seaweed ($FP/image = 0.667$) and open ocean ($FP/image = 0.610$). The sweep of confidence thresholds provided a tunable precision-recall trade-off to improve the verifier under energy constraints. Furthermore, each observed failure pattern suggested a corresponding architectural mitigation: thresholding for open ocean, multi-view quorum for seaweed, no additional measures indicated for glare under this sample, and training-data expansion for low-contrast negatives. Multimodal thermal-delta extension is a natural step forward.

REFERENCES

- [1] J. R. Jambeck *et al.*, “Plastic waste inputs from land into the ocean,” *Science*, vol. 347, no. 6223, pp. 768–771, Feb. 2015, doi: 10.1126/science.1260352.
- [2] National Oceanic and Atmospheric Administration, “What is the Great Pacific Garbage Patch?,” NOAA Ocean Service, Jun. 16, 2024. [Online]. Available: <https://oceanservice.noaa.gov/facts/garbagepatch.html>. Accessed: May 1, 2026.
- [3] L. Lebreton *et al.*, “Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic,” *Scientific Reports*, vol. 8, Art. no. 4666, Mar. 2018, doi: 10.1038/s41598-018-22939-w.
- [4] A. L. Andrady, “Microplastics in the marine environment,” *Marine Pollution Bulletin*, vol. 62, no. 8, pp. 1596–1605, 2011, doi: 10.1016/j.marpolbul.2011.05.030.
- [5] L. Biermann, D. Clewley, V. Martinez-Vicente, and K. Topouzelis, “Finding plastic patches in coastal waters using optical satellite data,” *Scientific Reports*, vol. 10, Art. no. 5364, Apr. 2020, doi: 10.1038/s41598-020-62298-z.
- [6] B. Bovcon, J. Muhovic, J. Pers, and M. Kristan, “The MaSTr1325 dataset for training deep USV obstacle detection models,” in *Proc. 2019 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, Macau, China, Nov. 2019, pp. 3431–3438, doi: 10.1109/IROS40897.2019.8967909.
- [7] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLO,” version 8.x, Ultralytics, 2023. [Online]. Available: <https://github.com/ultralytics/ultralytics>. Accessed: May 1, 2026.
- [8] RoboProject, “bottles Dataset,” version 2, Roboflow Universe, Roboflow, open-source object detection dataset, CC BY 4.0, generated Jul. 11, 2025. [Online]. Available: <https://universe.roboflow.com/roboproject-0aofk/bottles-svzvd-w6q4v/dataset/2>. Accessed: Jul. 2025.
- [9] D. Makeev, “Reflection in water, glare on water. img 010,” “Reflection in water, glare on water. img 011,” “Reflection in water, glare on water. img 013,” and “Reflection in water, glare on water. img 016,” Wikimedia Commons, photographs, CC BY-SA 4.0; U. Semenova, “Bright sun shining over the ocean at sunset,” Unsplash, photograph, Sep. 23, 2025, Unsplash License. [Online]. Available: https://commons.wikimedia.org/wiki/File:Reflection_in_water_glare_on_water_img_010.jpg; https://commons.wikimedia.org/wiki/File:Reflection_in_water_glare_on_water_img_011.jpg; https://commons.wikimedia.org/wiki/File:Reflection_in_water_glare_on_water_img_013.jpg; https://commons.wikimedia.org/wiki/File:Reflection_in_water_glare_on_water_img_016.jpg; <https://unsplash.com/photos/bright-sun-shining-over-the-ocean-at-sunset-R8K0vzUk9yk>. Accessed: May 2026.

- [10] Bjoertvedt, "Laminaria saccharina IMG 1029 listafyr.JPG," Wikimedia Commons, photograph, CC BY-SA 4.0; P. de Lange, "Lessonia tholiformis," iNaturalist photo 31394825, no rights reserved; D. Rabich, "Cape Town (ZA), Sea Point Promenade, Küste, Algen – 2024 – 3625.jpg," Wikimedia Commons, photograph, CC BY-SA 4.0; Nandaro, "Goleta Beach (2013) 11.JPG," Wikimedia Commons, photograph, CC BY-SA 3.0 Unported. [Online]. Available: https://commons.wikimedia.org/wiki/File:Laminaria_saccharina_IMG_1029_listafyr.JPG; <https://www.inaturalist.org/photos/31394825>; [https://commons.wikimedia.org/wiki/File:Cape_Town_\(ZA\),_Sea_Point_Promenade,_K%C3%BCste,_Algen--_2024--_3625.jpg](https://commons.wikimedia.org/wiki/File:Cape_Town_(ZA),_Sea_Point_Promenade,_K%C3%BCste,_Algen--_2024--_3625.jpg); [https://commons.wikimedia.org/wiki/File:Goleta_Beach_\(2013\)_11.JPG](https://commons.wikimedia.org/wiki/File:Goleta_Beach_(2013)_11.JPG). Accessed: May 2026.
- [11] A. Vinayak, "A bunch of seaweed floating on top of a body of water," Unsplash, photograph, Jun. 29, 2024; K. Holloway, "A large flock of seaweed floating on top of a body of water," Unsplash, photograph, Aug. 9, 2024; A. Castillo, "A body of water surrounded by grass and rocks," Unsplash, photograph, May 6, 2023; K. K., "Clear ocean water over seaweed and rocks," Unsplash, photograph, Apr. 4, 2026; N. Almendrada, "Seaweed and calm ocean under a clear blue sky," Unsplash, photograph, May 4, 2026. [Online]. Available: https://unsplash.com/photos/a-bunch-of-seaweed-floating-on-top-of-a-body-of-water-9lC_P3lwJwo; <https://unsplash.com/photos/a-large-flock-of-seaweed-floating-on-top-of-a-body-of-water-le27yiFKTWI>; https://unsplash.com/photos/a-body-of-water-surrounded-by-grass-and-rocks-OdO-V_uBdsM; <https://unsplash.com/photos/clear-ocean-water-over-seaweed-and-rocks-E4zHC7ZHFTA>; <https://unsplash.com/photos/seaweed-and-calm-ocean-under-a-clear-blue-sky-pbizyfsqtaQ>. Accessed: May 2026.
- [12] C. D. Rodin and T. A. Johansen, "Detectability of objects at the sea surface in visible light and thermal camera images," in *Proc. 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, Kobe, Japan, May 2018, pp. 1–10, doi: 10.1109/OCEANSKOB.2018.8559310.