Marine Litter Exposé

Ekaterina Gikalo

May 24, 2023

1 MARIDA Summary

The MARIDA (Marine Debris Detection and Analysis) article provides an exploration of the application of satellite imagery and computer vision algorithms in the context of marine debris detection. The study specifically focuses on utilizing multispectral satellite images obtained from the Sentinel-2 satellite to develop models capable of accurately identifying and classifying marine debris.

1.1 Dataset

The MARIDA Dataset used in the study consists of 1381 patches, containing a total of 837,357 annotated pixels. Each annotation belongs to one of the 15 classes: Marine Debris, Dense Sargassum, Sparse Sargassum, Natural Organic, Material Ship, Clouds, Marine Water, Sediment-Laden Water, Foam, Turbid Water, Shallow Water, Waves, Cloud Shadows, Wakes, Mixed Water. The pixel-level class distribution can be seen in Figure 1. The patches were derived from 63 Sentinel-2 scenes acquired between 2015 and 2021 with data collected from eleven countries. Each pixel in the dataset is assigned a confidence level of annotation, categorized as high, moderate, or low.

During data preprocessing, Rayleigh reflectance values were extracted at a 10 m resolution for 11 Sentinel-2 bands using the ACOLITE atmospheric processor, excluding the Vapour band (Band 9) and Cirrus band (Band 10).

1.2 Machine learning

For the machine learning techniques, the initial 15 classes were grouped into 11 classes by merging certain water classes into a single super-class. This decision was based on the semantic similarity and similar spectral profiles observed among these classes.

To evaluate the performance of various models for marine debris detection, the MARIDA article considers four different approaches. The first three of them are different Random Forest models: Random Forest with spectral signatures RF_{SS}, with spectral indices RF_{SS+SI} and gray-level co-occurrence matrix $RF_{SS+SI+GLCM}$. The extracted spectral indices were NDVI, NDWI, FAI, FDI, Shadow Index (SI), Normalized Difference Moisture Index (NDMI), Bare Soil Index (BSI) and NRD. To compute the GLCM features, Rayleigh corrected RGB composites were converted to grayscale images, which consequently were quantized in 16 bins-level. The selected GLCM features were Contrast, Dissimilarity, Homogeneity, Energy, Correlation and Angular Second Moment. For those features extraction, a window of size 13 x 13 was used.

The fourth approach is the U-Net architecture. The first input layer of U-Net was modified to adapt to the 11 Rayleigh reflectance Sentinel-2 bands, and the final classification layer was changed to output the 11 MARIDA classes. 4 down-sampling and up-sampling blocks were used, as well as 16 hidden channels produced by the initial down-sampling block. The annotators' confidence level information wasn't used for the training.

1.3 Results

The results of the predicted classes can be seen in the Figure 2. Among the mentioned models, the $RF_{SS+SI+GLCM}$ model consistently achieves the highest scores across all evaluation metrics, followed by RF_{SS+SI} and RF_{SS} and U-Net.

S2 Tile	MD	DenS	SpS	NatM	Ship	Cloud	MWater	SLWater	Foam	TWater	SWater	Waves	CloudS	Wakes	MixWater	# of pixels	# of S2 scenes
16PCC	1496	2048	574	78	3322	62082	60169	285886	712	99501	3960	3417	3585	5929	191	532950	19
16PDC	143	49	226	78	96	13507	15258	85449	334	24923	2251	0	883	253	75	143525	6
16PEC	129	222	645	193	485	11678	19341	11	86	27080	3782	108	1733	1115	51	66659	6
16QED	0	474	691	0	90	4098	1719	0	0	0	5910	0	1841	221	0	15044	2
18QWF	0	0	0	0	0	0	324	0	0	0	0	1461	0	0	0	1785	1
18QYF	1112	4	200	154	408	7977	1360	0	0	0	1038	0	314	48	58	12673	13
18QYG	90	0	0	7	0	373	222	0	0	831	277	0	106	0	15	1921	1
19QDA	0	0	21	3	11	0	110	0	0	5	40	0	0	0	0	190	1
30VWH	27	0	0	0	36	3505	24393	0	0	0	0	0	1975	0	0	29936	1
36JUN	46	0	0	0	625	3500	600	0	0	300	0	0	0	18	0	5089	1
48MXU	208	0	0	0	71	5807	194	0	0	382	45	0	489	15	12	7223	2
48MYU	24	0	0	0	223	0	291	0	0	10	48	0	0	611	0	1207	2
48PZC	24	0	0	0	298	4108	2079	0	48	4129	0	0	765	171	1	11623	3
50LLR	41	0	0	3	27	402	485	0	41	0	18	841	0	72	5	1935	1
51PTS	38	0	0	20	17	0	35	0	0	0	0	0	0	0	0	110	2
51RVQ	17	0	0	0	0	363	163	0	0	0	0	0	37	0	0	580	1
52SDD	4	0	0	328	94	0	2416	1591	4	451	0	0	0	37	2	4927	1
Total pixels	3399	2797	2357	864	5803	117400	129159	372937	1225	157612	17369	5827	11728	8490	410	**837377**	**63**
Perc. %	0,41	0,33	0,28	0,1	0,69	14,02	15,42	44,54	0,15	18,82	2,07	0,70	1,40	1,01	0,05	100	

https://doi.org/10.1371/journal.pone.0262247.t003

Figure 1: MARIDA's class distribution at pixel-level.

	RF_{SS}			RF_{SS+SI}			$RF_{SS+SI+GLCM}$			U-Net		
Class	**IoU**	**PA**	F_1	**IoU**	**PA**	F_1	**IoU**	**PA**	F_1	**IoU**	**PA**	F_1
MD	0.55	0.91	0.71	**0.67**	**0.92**	**0.8**	0.65	**0.92**	0.79	0.33	0.7	0.5
DenS	0.87	0.92	**0.93**	**0.88**	**0.93**	**0.93**	0.87	**0.93**	**0.93**	0.6	0.6	0.75
SpS	0.53	0.91	0.69	0.69	**0.92**	0.82	**0.83**	0.9	**0.91**	0.66	0.89	0.79
NatM	**0.31**	**0.47**	**0.47**	0.17	0.27	0.29	0.18	0.31	0.31	0.02	0.02	0.04
Ship	0.54	0.72	0.7	0.47	0.7	0.64	**0.67**	**0.82**	**0.8**	0.62	0.76	0.76
Clouds	0.75	0.85	0.86	0.74	0.82	0.85	**0.84**	**0.86**	**0.91**	0.62	0.62	0.76
MWater	0.66	0.82	0.79	0.65	0.83	0.79	**0.75**	**0.93**	**0.86**	0.61	0.88	0.76
SLWater	**1**	**1**	**1**	0.99	**1**	**1**	0.99	**1**	**1**	0.99	0.99	**1**
Foam	0.23	0.29	0.37	0.31	0.48	0.47	**0.6**	**0.74**	**0.75**	0.55	0.55	0.71
TWater	0.74	0.78	0.85	0.8	0.83	0.89	**0.88**	0.92	**0.94**	0.84	**0.95**	0.91
SWater	0.08	0.25	0.16	0.13	0.33	0.23	0.3	0.37	0.46	**0.45**	**0.67**	**0.62**
Average	0.57	0.72	0.69	0.59	0.73	0.7	**0.69**	**0.79**	**0.79**	0.57	0.69	0.69

https://doi.org/10.1371/journal.pone.0262247.t004

Figure 2: Evaluation scores obtained by RF_{SS}, RF_{SS+SI}, $RF_{SS+SI+GLCM}$ and U-Net for each class.

Regarding the classification of marine debris, RF_{SS+SI} demonstrates the highest scores, while adding the GLCM features does not improve the classification performance results.

2 Research questions

1. How does the performance of the U-net model change when applied to satellite images acquired by different satellites, such as Landsat or MODIS, and how can the model be adapted for cross-platform compatibility?

2. Can the $RF_{SS+SI+GLCM}$ and U-Net models be further improved by incorporating pre-processing techniques, such as cloud masking and denoising algorithms?

3. What are the limitations and challenges of using the ACOLITE Dark Spectrum Fitting algorithm for marine debris detection in satellite imagery, and how do these limitations impact the accuracy and reliability of the results?

4. What are the potential methods for incorporating annotators' confidence level information into the learning process of U-Net, to improve the classification of challenging cases?

5. How does the performance of the $RF_{SS+SI+GLCM}$ model compare to other deep learning architectures, such as Mask R-CNN or EfficientDet, for marine debris detection in satellite imagery?