Radiant MLHub
EARTH IMAGERY FOR IMPACT

Dataset Documentation

Dataset Name: Marine Debris Dataset for Object Detection in Planetscope Imagery

Location and boundaries

Overall Location Method

- ☐ Ground collection only
- ☐ Ground collection with boundary drawn using imagery
- ☐ Ground collection with spatial buffer added
- ☒ Boundary drawn from imagery
- ☐ Other _____
- ☐ Unknown

GeoLocation Device

- ☐ Industrial grade GPS (List model) _____
- ☐ Retail grade GPS
- ☐ Mobile Phone GPS
- ☒ N/A
- ☐ Unknown

Ground Boundary Method (Details explained in Appendix A)

- ☐ Live/Continuous point capture of walk-around
- ☐ Manual point capture of walk-around
- ☐ Manual point capture of polygon boundaries (not whole field)
- ☒ Manual point capture for later image annotation
- ☐ Manual point capture for spatial buffer within field
- ☐ Manual point capture while looking at but not in field, with heading recorded
- ☐ Other _____
- ☐ Unknown

Imagery used (Skip if no imagery used)

Sensor: Planetscope

Date(s): 2016-2019

List scenes used in Appendix B

Imagery Annotation methods

- ☐ Boundaries drawn based on a single ground point captured
- ☒ Boundaries drawn/edited based on multiple ground points captured
- ☐ Buffer validated from ground point captured
- ☒ Boundary drawn without ground reference data (Include description of methods in Appendix C)

☐ Pixels annotated without ground reference data (Include description of methods in Appendix C)
☐ Unknown

Boundary inclusion

☐ Captured polygon includes the entire field/area
☒ Captured polygon includes only a sample of the field/area

Classification

Classification Type

☐ Land cover
☐ Crop type
☒ Other: Marine Debris Objects

Classes/fields used

Describe in Appendix D

Ground Referenced Classification

☒ Observation (Describe methods of determination in Appendix E)
☐ Survey/interview with land holder (Describe methods in Appendix E)
☐ Other (Describe methods in Appendix E)

Image Referenced Classification

Describe methods used in Appendix C

Data Properties

Property name	Property Description	Parameters/Allowed responses (optional)
Institutions	Institutions which created the dataset	University of Alabama in Huntsville, Development Seed, NASA
Countries	Countries included in the dataset	Greece, Honduras, Ghana
Year	Year range of dataset	2016-2019
Numerical Class ID	Class identifier as an integer starting from 1. If only one class is being modeled, this value by default equals 1.	1
Categorical Class ID	Class identifier as a string description, e.g. “marine_debris”	marine_debris

Bounding box	Coordinates as [xmin, ymin, ymax, ymax].	variable
Tile name	Corresponding filename with slippy map coordinates and zoom level (x, y, z) separated by underscore.	x_y_z.jpg

Appendix A: Describe the method of geographic ground data collection

Most of the annotated marine debris data was verified by Kikaki et al. between 2016 and 2019 on vessel and diving expeditions in the Bay Islands of Honduras. Marine debris was observed and validated which allowed the team to extract the location coordinates in the WGS'84 coordinate system as listed in Table S1 of the supplementary file by Kikaki et al., 2020. The debris data from Mytilene, Greece were artificially deployed targets containing plastics by Topouzelis et al., 2019 during the Plastic Litter Project. The targets were visible in the Planetscope scenes used for this dataset.

Appendix B: List imagery scenes used for annotation (ideally also included in metadata)

The 43 Planetscope scenes containing marine debris which were used for annotation are included in this list: 20160928_153233_0e16, 20161008_153104_0e3a, 20161008_153105_0e3a, 20170227_203307_0c37, 20170227_203600_1_0c46, 20170227_203601_0c46, 20170322_153132_0e0f, 20170323_152750_0e20, 20170326_153233_0e26, 20170326_153234_0e26, 20170326_153240_0e26, 20170326_154021_0e2f, 20170826_154046_100c, 20170826_154047_100c, 20170927_154610_0e14, 20170927_154611_0e14, 20171007_161314_1_0c65, 20171007_161315_0c65, 20171015_153720_100c, 20180107_153415_1032, 20180224_153958_0f1b, 20180224_153959_0f1b, 20180224_162902_0f2b, 20180313_154258_1008, 20180313_154259_1008, 20180313_154300_1008, 20180313_154301_1008, 20180607_082852_1012, 20180919_155029_1044, 20180919_160750_1020, 20181022_155353_100a, 20181022_155659_101e, 20181023_155746_0f34, 20181023_155747_0f34, 20181024_160201_0f3b, 20181024_160202_0f3b, 20181031_095925_103b, 20181124_155713_1049, 20181124_155714_1049, 20181124_155715_1049, 20181124_155716_1049, 20181215_075902_103e, 20190418_074358_0f3c

Radiant MLHub
EARTH IMAGERY FOR IMPACT

Appendix C: Describe how boundaries and classes were determined without ground reference data

Planet imagery was validated for visual marine debris detections using the Planet Explorer. Determining marine debris visually is relatively straightforward as it has a frequently uniform color and aggregate shape due to the presence of vegetation and ocean convergence dynamics near coastal regions, respectively. Planetscope scenes for the same location were checked on the closest dates before and after the ground validated scenes containing marine debris and the location of the debris objects were different. The change in location of suspected marine debris objects confirmed that they are floating debris. Furthermore, a study included dates of river discharges in Honduras which were corroborated with marine debris locations (Kikaki et al., 2020). Certain dates for marine debris events in Ghana were collected using social media by Biermann et al., 2020. Using a combination of visual inspection, verification of movement of the floating objects, and proximity to known river discharge locations, the marine debris locations were determined without ground reference data. Each marine debris object in our collection of Planetscope scenes was manually labeled using Image Labeler, a labeling tool developed by NASA's Interagency Implementation and Advanced Concepts Team (IMPACT).

Appendix D: List all top-level classes or the classification guidance used

The only class labeled in this dataset is marine debris. Bounding boxes were drawn around marine debris for the object detection model. The only top-level class used is marine debris.

Appendix E: Describe methods for determining classes based on direct/ground observation

Marine debris in this dataset comprises multiple object classes which include floating vegetation such as seaweed and algae, woody materials, timber, sea foam, pumice, and plastics. The floating debris were observed or deployed directly in the studies by Kikaki et al. and Topouzelis et al. as described in Appendix A.

Include any additional information/extra space as Appendix F+

The dataset consists of the 256x256 pixel tiled images of marine debris sorted by the parent scene, the geojsons of marine debris objects merged on the parent scene, and the labels.npz file containing the bounding box coordinates and class ID of marine debris objects in each tile.

Sample code for reading a labels.npz file:

```
import numpy as np
# Read an example NPZ file
x = np.load('20181124_155713_1049/labels.npz', mmap_mode='r')
# Print all of the tile entries in the NPZ
print(x.files)
# Print the tile name of the first entry
print(x.files[0])
# Output: 16762-29718-16'
# Print the bounding box coordinates and numerical class ID of the tile's associated annotations
```

```
print(x['16762-29718-16'])

# Output: array([[131,   8, 193,  77,   1], [159,  73, 214, 150,   1]])
# The nested array contains entries displaying [xmin, ymin, xmax, ymax, numerical class ID]
# There can be more than one bounding box for a single tile.
```

References

1. Biermann, L., Clewley, D., Martinez-Vicente, V. *et al.* Finding Plastic Patches in Coastal Waters using Optical Satellite Data. *Sci Rep* **10,** 5364 (2020). https://doi.org/10.1038/s41598-020-62298-z

2. Kikaki, A.; Karantzalos, K.; Power, C.A.; Raitsos, D.E. Remotely Sensing the Source and Transport of Marine Plastic Debris in Bay Islands of Honduras (Caribbean Sea). *Remote Sens.* **2020**, *12*, 1727. https://doi.org/10.3390/rs12111727.

3. Konstantinos Topouzelis, Apostolos Papakonstantinou, Shungudzemwoyo P. Garaba, Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018), International Journal of Applied Earth Observation and Geoinformation, Volume 79, 2019, Pages 175-183, ISSN 0303-2434, https://doi.org/10.1016/j.jag.2019.03.011.