

# UAV-BASED RIVER PLASTIC DETECTION WITH A MULTISPECTRAL CAMERA

I. Cortesi<sup>1\*</sup>, A. Masiero<sup>1</sup>, G. Tucci<sup>1</sup>, K. Topouzelis<sup>2</sup>

<sup>1</sup> Dept. of Civil and Environmental Engineering, University of Florence,  
via di Santa Marta 3, Florence 50139, Italy - (irene.cortesi, andrea.masiero, grazia.tucci)@unifi.it

<sup>2</sup> Dep. of Marine Science, University of the Aegean, University Hill, 81100 Mytilene, Greece - topouzelis@aegean.gr

**KEY WORDS:** Plastic, Machine Learning, Random Forest, Connected Regions, Multispectral Camera, Object detection

## ABSTRACT:

Plastic is the third world's most produced material by industry (after concrete and steel), but people recycle only 9% of plastic that they have used. The other parts are either burned or accumulated in landfills and in the environment, the latter being the cause of many serious consequences, in particular when considering a long-term scenario. A significant part the plastic waste is dispersed in the aquatic environment, having a dramatic impact on the aquatic flora and fauna. This motivated several works aiming at the development of methodologies and automatic or semi-automatic tools for the plastic pollution detection, in order to enable and facilitate its recovery. This paper deals with the problem of plastic waste automatic detection in the fluvial and aquatic environment. The goal is that of exploiting the well-recognized potential of machine learning tools in object detection applications. A machine learning tool, based on random forest classifiers, has been developed to properly detect plastic objects in multi-spectral imagery collected by an unmanned aerial vehicle (UAV). In the developed approach, the outcome is determined by the combination of two random forest classifiers and of an area-based selection criterion. The approach is tested on 154 images collected by a multi-spectral proximity sensor, namely the MAIA-S2 camera, in a fluvial environment, on the Arno river (Italy), where an artificial controlled scenario was created by introducing plastic samples anchored to the ground. The obtained results are quite satisfactory in terms of object detection accuracy and recall (both higher than 98%), while presenting a remarkably lower performance in terms of precision and quality. The overall performance appears also to be dependent on the UAV flight altitude, being worse at higher altitudes, as expected.

## 1. INTRODUCTION

Given the high usability and versatility, plastics are commonly used for many purposes. Plastics are lightweight, malleable, inexpensive and easy to produce materials. Today, it is still not possible to dispose of most of the used plastics in a proper and sustainable way, and it is often dispersed on the environment, leading to a significant impact on flora and fauna (Plastics Europe, 2022, United Nations Environment Program, 2018). In particular, every year 20 million tons of plastic waste are dispersed in the aquatic ecosystem and if the trend is not reversed, it is foreseen that there will be more plastic than fish in the seas and oceans in 2050 (Jong, 2018). Most of the plastic pollution in the oceans comes directly from coastlines or via rivers. The plastic waste transportation through waterways is still not sufficiently well investigated, but it has been estimated that more than 2 million tons come from rivers every year (Meijer et al., 2021, Tasseron et al., 2021). The study of plastic fluvial transportation and sedimentation is important because in this context plastic waste can be detected and removed before it reaches the marine environments.

During the last twenty years, several efforts have been made by the remote sensing community for the plastic pollution detection over large areas. More specifically, most of the research works on this topic focus on the detection of floating litter in seas and oceans through satellite images.

To this aim, the use of the freely accessible Copernicus Sentinel-2 mission data is currently a quite convenient option. Sentinel-2 satellite data are multispectral images are made up

of 13 spectral bands from VNIR to SWIR. Their spatial resolution varies from 10 to 60 m (ESA - Sentinel Online, n.d.). Such data have been exploited by (Topouzelis et al., 2020) and (Themistocleous et al., 2020) to investigate the detection of large artificial plastic targets. Such detection of plastic waste, and its separation from water or aquatic vegetation, is often based on the use of spectral indices such as Normalized Difference Water Index 2 (NDWI2), Plastic Index (PI), Reversed Normalized Difference Vegetation Index (RNDVI) and Floating Debris Index (FDI) (Biermann et al., 2020, Page et al., 2020, Themistocleous et al., 2020).

Instead, the detection of smaller plastic objects, usually over smaller areas, should be done by different instruments, ensuring the representation of the region of interest at a higher resolution. Since Unmanned Aerial Vehicles (UAV) allow the quite quick and easy acquisition of high-resolution spatial data over relatively large areas, they appear as a quite ideal option for plastic detection in local areas, such as rivers, beaches, coastlines or lakes, equipping the UAV with proper sensors, such as RGB and multispectral cameras (Martin et al., 2018). Plastic detection in this imagery is typically done by means of machine learning tools, and, recently, with deep learning-based approaches (Jakovljevic et al., 2020, Wolf et al., 2020).

Given the important role of waterways in the plastic waste transportation to the seas, this paper focuses on detection of plastic objects in river environments. Similarly to the satellite remote sensing approaches, this work aims at plastic litter detection by properly employing multispectral information in a machine learning approach (random forest-based detection), ((Belgiu and Drăguț, 2016)), described in Section 3. The proposed approach has been tested on multi-spectral imagery, col-

\* Corresponding author

lected by means of a proximity multi-spectral sensor, namely the MAIA-S2 camera, mounted on a UAV, the DJI Matrice 300, while flying over a portion of the Arno river (Italy). A more detailed description of the case study is provided in Section 2, whereas the obtained results are reported in Section 4, and some conclusions are drawn in Section 5.

## 2. CASE STUDY AND MATERIALS

### 2.1 Study Area

The study area has been identified in the locality of Prulli (Reggello, Florence, Italy) on the Arno river. A portion of the river, approximately 100 meter long, has been selected for its characteristics, such as the width, the depth of the stream bed, and the possibility of reaching the centre of the watercourse via some ruins of an ancient bridge (Ponte di Annibale).

Plastic samples (bottles, flacons and shoppers) were introduced into the river, and tied to the ancient bridge through a transparent fishing line, in such a way to ease the object recovery at the end of the data collection. The fishing line length was different for each plastic object, and each of them had a limited freedom of movement. Figure 1 shows the region of interest, i.e. where the UAV flew (red rectangle) and the place where the fishing lines were anchored (yellow circle).



Figure 1. Study area: Arno river in the locality of Prulli (Italy).

### 2.2 Dataset

The imagery has been acquired by means of the multispectral camera MAIA-S2, developed by SAL Engineering (Modena, Italy) and EOPTIS (Trento, Italy). This camera is provided with nine different optical sensors, with the same bands of the Sentinel-2 Satellite (Table 1) (SAL Engineering and EOPTIS, 2018). In particular, two are red edge bands (S5, S6), whereas the last three are in the near infrared (S7, S8, S9). The native camera resolution is  $1208 \times 960$ , the nominal focal length is 7.5 mm, and hence a field of view of  $35^\circ \times 26^\circ$ .

The MAIA-S2 camera was mounted on a DJI Matrice 300, which flew at low speed over the area of interest, while collecting one image per second. The UAV flew at different alti-

tudes, ranging from 20 m to 80 m from the ground, in order to compare the plastic detection results varying the flight height.

Band	Start WL [nm]	Stop WL [nm]	Color
S1	433	453	Violet
S2	457.5	522.5	Blue
S3	542.5	577.5	Green
S4	650	680	Red
S5	697.5	712.5	Red Edge 1
S6	732.5	747.5	Red Edge 2
S7	773	793	NIR 1
S8	784.5	899.5	NIR 2
S9	855	875	NIR 3

Table 1. Wavelength (WL) intervals of MAIA S-2 bands.

Image locations were acquired by means of an external GNSS receiver mounted on the UAV. This receiver worked in RTK mode, exploiting corrections provided by a GNSS base-station located approximately 100 m far from the area of interest.

An irradiance light sensor (ILS) was mounted on the top of the DJI Matrice 300 (as shown in Figure 2), enabling automatic radiometric correction of the multispectral imagery.

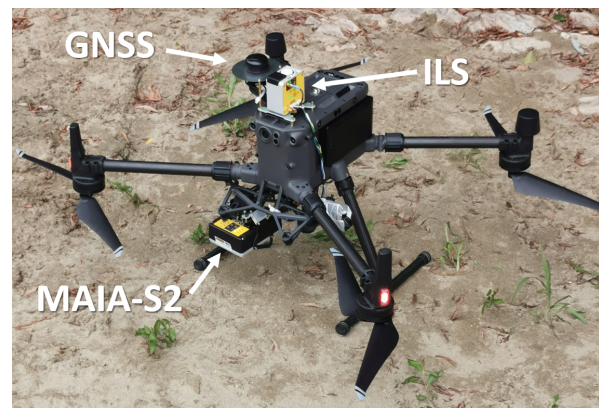


Figure 2. MAIA-S2 multispectral camera, GNSS external antenna and ILS sensor mounted on DJI Matrice 300.

Geometric corrections and co-registration of the bands were computed by means of the MAIA image-processing software, provided by SAL Engineering along with the camera. Overall, the dataset is composed by 1268 images at different altitudes (from 20 m to 80 m, as shown in Table 2). Nevertheless, this work focuses on the comparison of the results obtained in the 30 m and 80 m cases.

Altitude [m]	# of images	GSD [cm]
20	205	1.0
30	119	1.6
40	895	2.1
60	44	3.1
80	35	4.2

Table 2. Number of collected images and Ground Sampling Distance (GSD) as functions of the flight altitude with respect to the ground.

## 3. METHOD

The proposed method for plastic detection is based on the use of a multi-step random forest approach, where the final results are

obtained by means of the cascade of two random-forest classifiers (named classifier A and B hereafter) and of an area-based selection criterion. It is worth to notice that the method proposed in this work is an evolution of the one presented in (Cortesi et al., 2021), in particular improving the previous results by including the above mentioned two-step procedure.

Ideally, a unique classifier, as in (Cortesi et al., 2021), able to properly distinguish “Plastic” from “Other” pixels would be preferable. However, according to the experimental tests performed with the collected data (see Table 2), the performance of such a in certain cases is unsatisfying. Hence, a multi-step classification procedure has been used, instead.

The analysis starts with a pixel-based classification, where only two classes are considered in the classification process, “Plastic” and “Other”, aiming at distinguishing pixels related to plastic objects from all the others.

Both classifier A and B are pixel-classifiers: they receive as input the co-registered nine bands of a single multispectral image, and they analyse pixel by pixel the nine-band measurements, aiming at distinguishing plastics from the other materials by their peculiar spectral response.

Despite being similar in terms of the general goal (i.e. detecting plastic pixels), classifier A and B differ for their specific classification ability, which have been achieved by using different trainings.

To be more precise, classifier A is well suited for ensuring an optimal performance in terms of avoiding false negatives (i.e. pixels incorrectly classified as “Other”) in the pixel plastic detection. Differently, classifier B ensures a better performance in terms of reducing the number of false positive pixels, i.e. pixels incorrectly classified as “Plastic”.

Since plastic objects may not visible in many images, and since the number of plastic pixels in an image is usually small in an image, the total number of “Other” pixels is huge with respect to the total number of “Plastic” pixels. Consequently, the direct use of classifier A and B on the multispectral images typically lead to a much worse performance of A, due to the much higher number of false positives.

Examples of the results obtained by classifier A and B on one of the images acquired at 30 m height can be seen in Figure 3 - 6: Figure 3 and Figure 5 show the results obtained on the image by using classifier A and B, respectively, whereas 4 and 6 are the zooms of the previously mentioned figures on one of the plastic objects. These figures compare the data inputted in the classifier (shown as RGB-like images, obtained by bands S4, S3 and S2, see Table 1) with the obtained results (plastics are shown in white, whereas “Other” pixels in blue).

By comparing the results in Figure 3 - 6, the better ability of B in avoiding false negatives is quite apparent, whereas A reduces the amount of false positives (usually appearing as salt and pepper-like noise), as expected.

The different classifier behaviors derive from a different training:

- A: 500 trees have been used in the random forest classifier, trained by randomly extracting from the dataset 100k “Plastic” and 500k “Other” pixels. The use of unbalanced

training datasets for the two classes is motivated by: 1) the quite small amount of plastic pixels available in the images, which limits the maximum number of plastic training pixels, and 2) the need of a reasonably large number of “Other” pixels to decently represent the characteristics of the huge amount of pixels in this class. The use of balanced training classes is foreseen in the future developments of this work in order to yield to a more robust classifier performance.

- B: Random forest classifier B has been set and trained similarly to A, but inserting an additional set of 90k pixels in the “Other” training dataset, randomly extracted from certain critical areas (rocks, sunglint, foam, etc.). These additional pixels allow classifier B to improve its ability in properly recognizing the pixels of such critical areas.

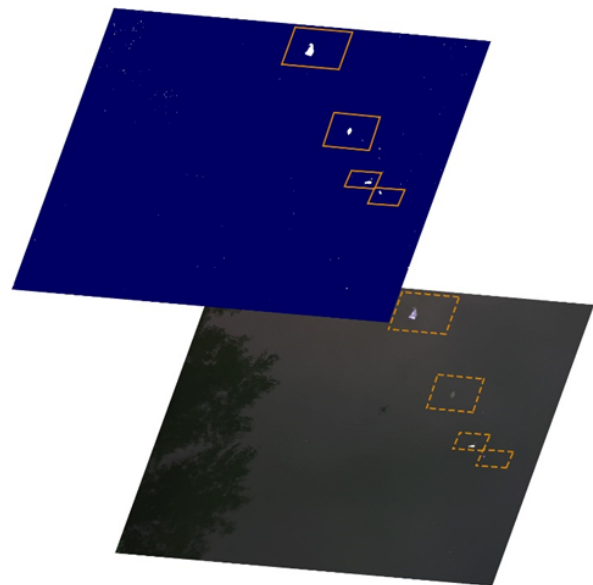


Figure 3. Example of results obtained with classifier A: RGB-like image (bottom) and classification results (top).

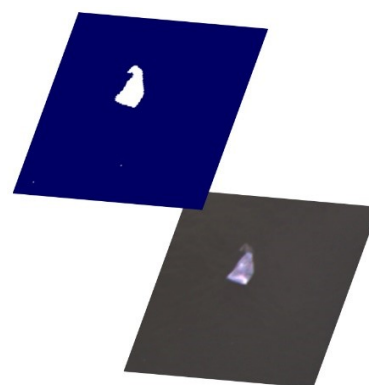


Figure 4. Zoom of Figure 3 on one the plastic samples.

Given the peculiar characteristics of the above presented classifiers, B appears to be a reasonable first step for a plastic detection procedure, allowing the identification of most (but not all) of the plastic pixels, while keeping the number of false positives lower with respect to A.

Once some plastic pixels have been found by B, A is used in order to improve the plastic pixel detection just in the neighboring

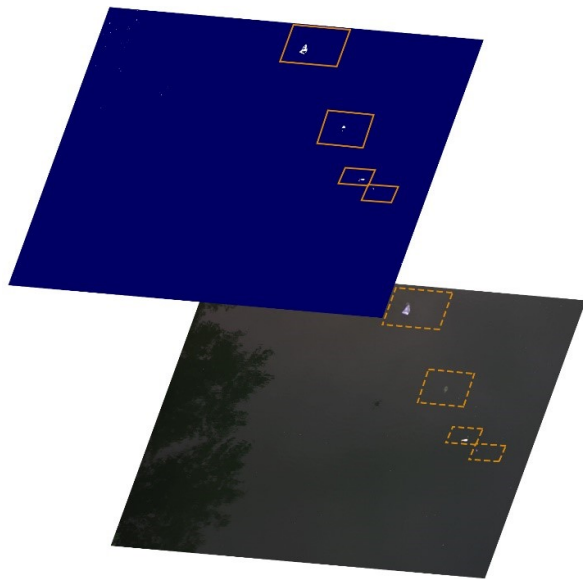


Figure 5. Example of results obtained with classifier B: RGB-like image (bottom) and classification results (top).

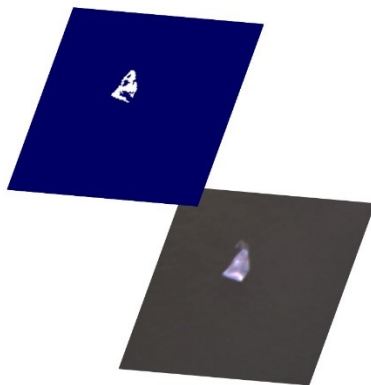


Figure 6. Zoom of Figure 5 on one of the plastic samples.

area of the already found pixels. In practice, the pixels in a spatial buffer around any previously detected plastics are inputted in classifier A. Then, the results of A and B are merged: B outcome is used in the “Other”-only areas, whereas A outcome is used to better discriminate “Plastic” and “Other” pixels where plastic is present.

For instance, this procedure leads to the substitution of the B results shown in Figure 6 with the A outcome in Figure 4 thus ensuring a more effective identification of the plastic boundaries. The results obtained by using classifier B, A and the cascade B+A can also be compared in Figure 7 and, once zoomed, in 8.

Given the pixel Ground Sample Distance (GSD) size (Table 2), once entered in the fluvial environment usually the area of a floating macro-plastic object is quite larger than a pixel. Since the classification errors resulting from the cascade of B and A are often salt and pepper-like noise, an area-based selection of the detected plastic areas can be used in order to reduce the number of false positives.

First, the image regions associated to different plastic objects are identified by computing the connected regions, by means of a flood-fill algorithm (Fishkin and Barsky, 1985). Then,

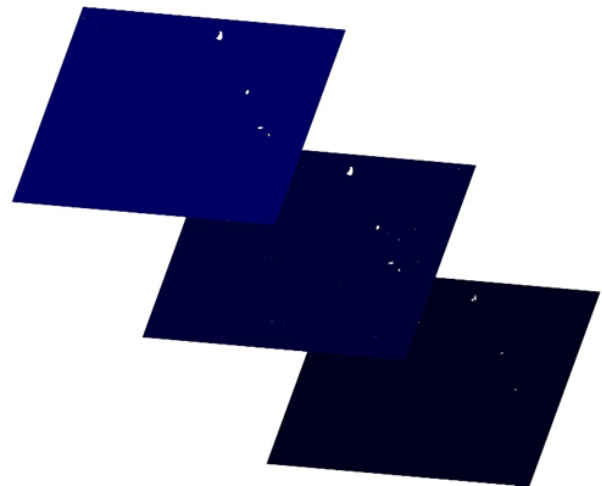


Figure 7. Comparison of the results obtained with classifier B+A, A, and B (from top to bottom).

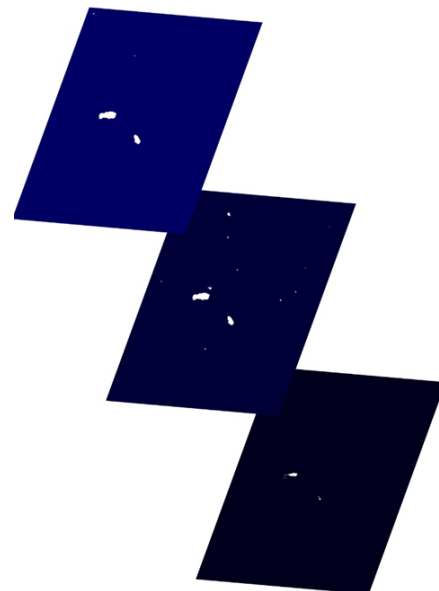


Figure 8. Zoom of Figure 7: comparison of the results obtained with classifier B+A, A, and B (from top to bottom).

the plastic objects with areas lower than a certain threshold, to be defined as a design parameter of the procedure, are discarded. The overall procedure is summarized in the block diagram shown in Figure 9.

Figure 10 compares the areas of the plastic objects and of the false positives (just after applying the B+A cascade) in the 30 m height case. The quite remarkable area difference in this dataset partially justifies the adoption of the area-based selection, as explained above. This choice can clearly have a negative impact on the detection performance when the size of the plastic objects dispersed in the considered region is smaller than the considered threshold area.

#### 4. RESULTS AND DISCUSSION

Table 3 shows the numerical classification results, obtained by applying B and A, at two different flight altitudes, i.e. 30 m and 80 m from the ground. This comparison aims at highlighting

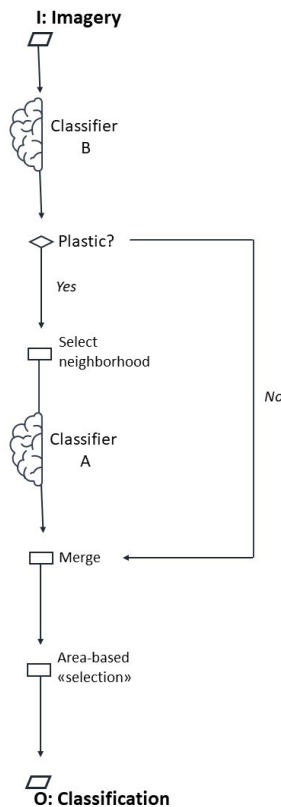


Figure 9. Block diagram of process of classification.

the dependence of the classification results with respect to the GSD value.

Despite the obtained accuracy and recall are good at both the altitudes, the high number of detected false positives leads to an unsatisfactory performance in terms of precision and quality. Realizing the main causes of such false positives is of key importance to determine how to improve the overall approach performance. To such aim, Fig. 11 compares the spectral signatures of:

- (a) a transparent sample, visible in more than 450 images,
- (b) water, assessed by around 1 M pixels from just one image,
- (c) sunglint, extracted from about 20 images.

The characterization of the spectral signatures in Fig. 11 resembles those presented in (De Giglio et al., 2021), even though using different bands, i.e. different portions of the electromagnetic spectrum.

While water and transparent plastic signatures are quite different (Fig. 11(a) and (b)), hence justifying the use of multispectral information to detect plastics in fluvial/maritime environment,

	30 m	80 m
Accuracy [%]	100.0	98.8
Precision [%]	3.1	7.3
Quality [%]	3.1	7.3
Recall [%]	100.0	98.8

Table 3. Object detection performance of the cascade of classifier B and A.

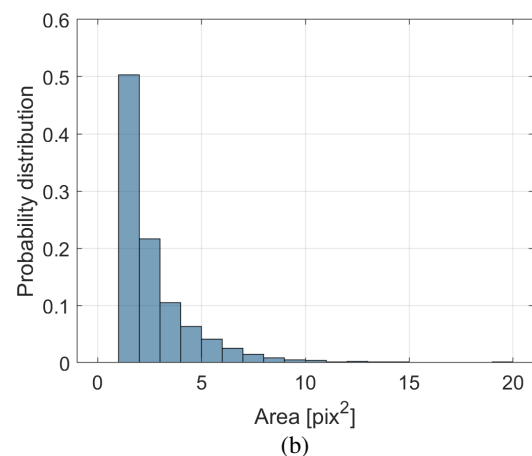
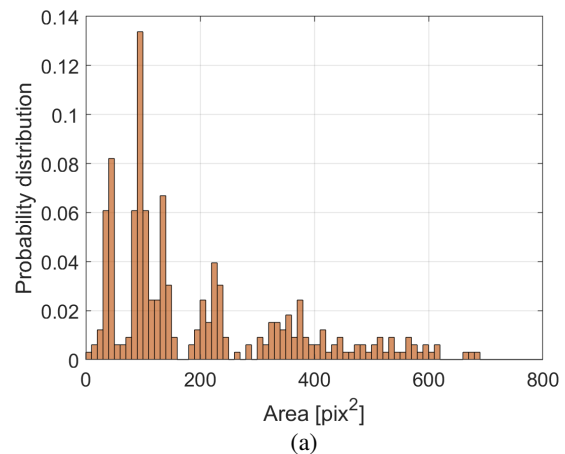


Figure 10. (a) plastic object and (b) false positive (after applying the B+A cascade) areas in the 30 m height dataset.

the latter is apparently similar to the sunglint one (Fig. 11(a) and (c)). Being sunglint hardly distinguishable from certain plastics in the considered portion of the electromagnetic spectrum, such similarity can be considered as the main cause of the previously mentioned false positives.

Although sunglint and certain plastic samples have very similar spectral signatures, their size is usually quite different: Figure 10 compares the distributions of the areas (in the multispectral images acquired at 30 m altitude) of the plastic objects (a), and of the sunglint (b).

The above consideration motivated the introduction of an area-based object selection step, i.e. selecting only those objects with an area above a certain threshold (which should be properly set in order to ensure a satisfying performance of the approach), leading to a remarkable reduction of the false positives (see Table 4), in particular in the 30 m altitude case, while only moderately reducing accuracy and recall.

	30 m	80 m
Accuracy [%]	95.1	92.9
Precision [%]	89.7	33.9
Quality [%]	85.8	33.1
Recall [%]	95.1	92.9

Table 4. Object detection performance of the proposed approach.

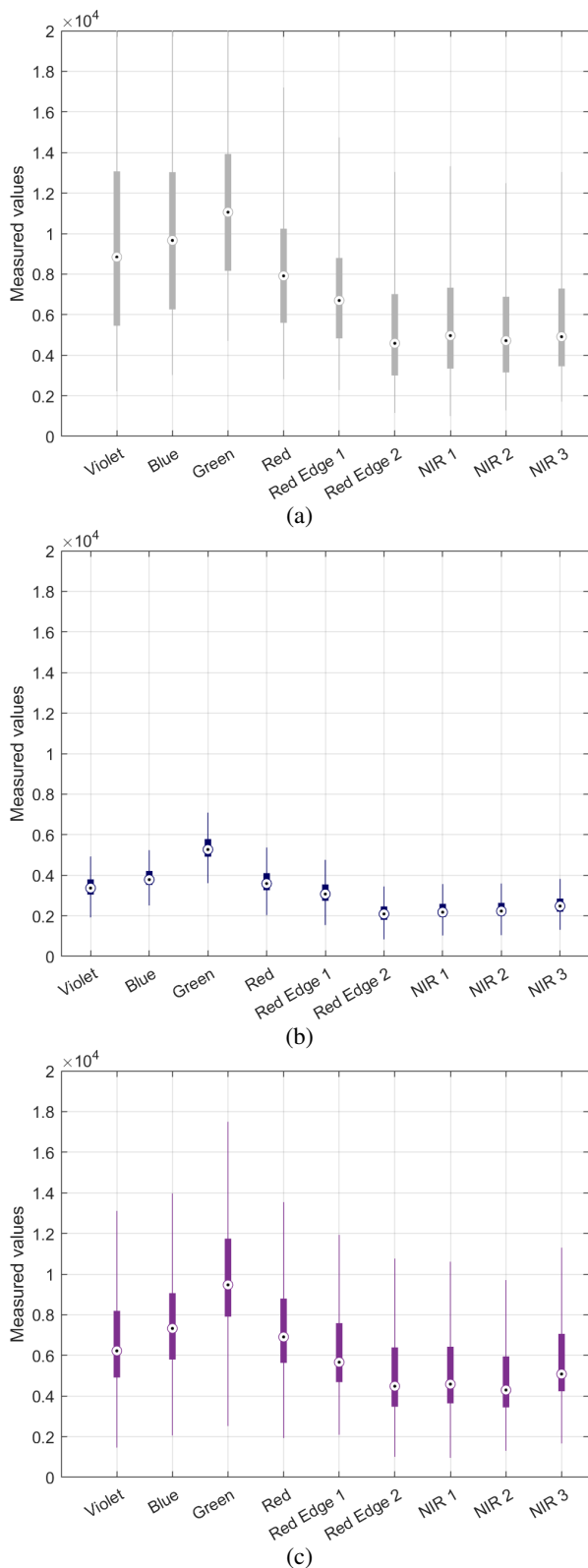


Figure 11. Examples of spectral signatures (a) transparent plastic object, (b) water, (c) sunglint.

Despite improving the performance at both the considered altitudes, such improvement is clearly much less apparent in the 80 m case. This is because the sunglint effect causes the false detection of areas of similar values in pixel at both the altitudes, whereas the plastic object area reduces significantly with the

increase of the flight altitude (and hence of the GSD), making them less distinguishable.

Consequently, the effectiveness of the implemented area-based selection step reduces when increasing the flight altitude. As a side effect, plastic objects smaller than the used area threshold cannot be detected: hence, such threshold should be carefully chosen, and, in general, it cannot be used when the plastic objects to be detected are of size comparable with the pixel size.

Given the presence of rather indistinguishable plastic objects and sunglint in the acquired multispectral images, overcoming the need of an area-based selection step requires the introduction of some different information. For instance, in analogy with (Topouzelis et al., 2020), the use of different bands in the infrared spectrum are expected to be useful: this will be investigated in our future work.

## 5. CONCLUSIONS

This paper presented a multi-step approach macroplastic detection in river ecosystems, by exploiting multispectral imagery acquired by an UAV. The main step of the proposed method is based on the combination of two machine learning classifiers (i.e. the cascade of two random forest classifiers).

The obtained results proved to be quite promising, especially for images acquired at quite low altitudes, whereas a degradation of the performance has been shown when increasing the acquisition altitude.

In accordance with similar works for plastic detection from satellite imagery, the introduction of information at certain additional bands in the infrared spectrum are expected to improve the overall detection performance of the algorithm, for instance making plastics more distinguishable from sunglint. This aspect will be considered in our future investigations, along with the use of other machine learning tools, such as deep learning approaches, in order to improve the plastic detection results.

## REFERENCES

Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114, 24–31.

Biermann, L., Clewley, D., Martinez-Vicente, V., Topouzelis, K., 2020. Finding plastic patches in coastal waters using optical satellite data. *Scientific reports*, 10(1), 1–10.

Cortesi, I., Masiero, A., De Giglio, M., Tucci, G., Dubbini, M., 2021. Random Forest-Based River Plastic Detection with a Handheld Multispectral Camera. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 9–14.

De Giglio, M., Dubbini, M., Cortesi, I., Maraviglia, M., Parisi, E. I., Tucci, G., 2021. Plastics waste identification in river ecosystems by multispectral proximal sensing: a preliminary methodology study. *Water and Environment Journal*, 35(2), 569–579.

ESA - Sentinel Online, n.d. Sentinel user guides. (13 January 2022). <https://sentinels.copernicus.eu/web/sentinel/home>.

- Fishkin, K. P., Barsky, B. A., 1985. An analysis and algorithm for filling propagation. *Computer-generated images*, Springer, 56–76.
- Jakovljevic, G., Govedarica, M., Alvarez-Taboada, F., 2020. A deep learning model for automatic plastic mapping using unmanned aerial vehicle (UAV) data. *Remote Sensing*, 12(9), 1515.
- Jong, F., 2018. More plastic than fish in the sea by 2050. will advertising be the solution?: The effects of ambient advertising on attitudes and recycling behavior. Master's thesis, University of Twente.
- Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M. F., Duarte, C. M., 2018. Use of unmanned aerial vehicles for efficient beach litter monitoring. *Marine pollution bulletin*, 131, 662–673.
- Meijer, L. J., van Emmerik, T., van der Ent, R., Schmidt, C., Lebreton, L., 2021. More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean. *Science Advances*, 7(18), eaaz5803.
- Page, R., Lavender, S., Thomas, D., Berry, K., Stevens, S., Haq, M., Udugbezi, E., Fowler, G., Best, J., Brockie, I., 2020. Identification of tyre and plastic waste from combined copernicus sentinel-1 and-2 data. *Remote Sensing*, 12(17), 2824.
- Plastics Europe, 2022. Marine litter (12 january 2022). <https://plasticseurope.org/sustainability/circularity/waste-management-prevention/marine-litter/> .
- SAL Engineering and EOPTIS, 2018. Maia, the multispectral camera. (12 january 2022). <https://www.spectralcam.com/maia-tech/>.
- Tasserou, P., Van Emmerik, T., Peller, J., Schreyers, L., Biermann, L., 2021. Advancing Floating Macroplastic Detection from Space Using Experimental Hyperspectral Imagery. *Remote Sensing*, 13(12), 2335.
- Themistocleous, K., Papoutsas, C., Michaelides, S., Hadjimitsis, D., 2020. Investigating detection of floating plastic litter from space using Sentinel-2 imagery. *Remote Sensing*, 12(16), 2648.
- Topouzelis, K., Papageorgiou, D., Karagaitanakis, A., Papakonstantinou, A., Arias Ballesteros, M., 2020. Remote Sensing of Sea Surface Artificial Floating Plastic Targets with Sentinel-2 and Unmanned Aerial Systems (Plastic Litter Project 2019). *Remote Sensing*, 12(12).
- United Nations Environment Program, 2018. The state of plastics. world environment day outlook.
- Wolf, M., van den Berg, K., Garaba, S. P., Gnann, N., Sattler, K., Stahl, F., Zielinski, O., 2020. Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q). *Environmental Research Letters*, 15(11), 114042.