MINI UAV-BASED LITTER DETECTION ON RIVER BANKS

I. Cortesi[1],* F. Mugnai[1], R. Angelini[1], A. Masiero[1]

[1] Dept. of Civil and Environmental Engineering, University of Florence,
via di Santa Marta 3, Florence 50139, Italy - (irene.cortesi, francesco.mugnai, riccardo.angelini, andrea.masiero)@unifi.it

Commission IV, WG IV/3

KEY WORDS: Litter, Plastic, Deep Learning, Object detection, Mini-UAV

ABSTRACT:

Most of the anthropic pollution arriving to seas and oceans is carried by rivers, leading to a dramatic impact on the aquatic flora and fauna. Hence, several strategies have already been considered to reduce the use of plastic (and non biodegradable) objects, fostering recycling, detect litter in the environment, and catch it. This work aims at implementing a litter detection approach usable also in urban areas, hence considering a mini-UAV (Unmanned Aerial Vehicle) in order to reduce the issues related to flight restrictions. The implemented strategy exploits a high resolution map of the area of interest, a properly trained deep learning litter object detector, and a vision based localization system, which allows to remarkably reduce the positioning error of the UAV navigation system, in order to provide estimates of the litter object positions with an accuracy at decimeter level for objects not too far from locations recognizable in the map.

1. INTRODUCTION

As a matter of fact, most of the anthropic pollution (in particular plastic) arriving to seas and oceans is carried by rivers. Indeed, the river network passes through urban settlements, where litter is typically introduced in the fluvial environment, and then conveyed towards the river mouth and, eventually, into the seas (Lebreton et al., 2017, Schmidt et al., 2017), dramatically impacting on the aquatic flora and fauna (Kurtela et al., 2019).

Since reducing the impact of anthropic pollution on the aquatic environment is of major importance, several strategies have already been implemented to reduce the use of plastic (and non biodegradable) objects, fostering recycling (da Costa, 2021), detect litter in the environment (Topouzelis et al., 2021), and catch it (Gabrys, 2013).

Focusing on litter detection on the aquatic environment (e.g. on seas, rivers, beaches), thanks to the remote sensing community efforts, many methods have been developed during the recent years to such aim. Depending on the available equipment (e.g. Unmanned Aerial Vehicle (UAV), satellite imagery), the search for litter has been investigated at different spatial scales.

Given their flexibility of use, portability and limited costs, the use of UAVs for local surveys, mapping (Nex and Remondino, 2014), and in monitoring applications, is growing almost everywhere in the world. For instance, they can be used to investigate the presence of litter on certain sections of rivers and beaches, enabling the collection of high spatial resolution imagery (and, more in general, multi-sensor data), without any time constraint (unlike satellite images). A quick overview of the off-the-shelf sensors currently available on UAVs include for instance RGB, multi-spectral, and thermal cameras, LiDAR (Light Detection and Ranging), and hyper-spectral sensors.

For what concerns litter detection (plastics, in particular) in the aquatic environment (including also beaches and river banks), (Geraeds et al., 2019) compared the effectiveness for litter detection of visual inspections (done by an expert operator) on RGB images acquired by a UAV flying over a river with those on images collected from a bridge. Periodic operator-assisted visual inspections on RGB images collected by a UAV have been considered in (Merlino et al., 2020), whereas machine learning methods (i.e. Support Vector Machine-based classification (Suykens and Vandewalle, 1999, Hsu and Lin, 2002)) have been introduced in (Martin et al., 2018) in order to automatically detect litter on an orthomosaic obtained from UAV RGB images of a beach. Near InfraRed (NIR) along with RGB imagery has been deployed in (Taddia et al., 2021) in order to filter out vegetation from the areas to be investigated for litter detection on an orthomosaic representing the considered case study area. Multi-spectral imagery has also been successfully used for plastic detection in the satellite case (e.g. for plastic detection on seas and oceans (Biermann et al., 2020, Themistocleous et al., 2020)), and, more recently, on the rivers thanks to UAV-imagery (Cortesi et al., 2022, Cortesi et al., 2021, Topouzelis et al., 2020, Gonçalves and Andriolo, 2022, Maharjan et al., 2022, Cheng et al., 2021).

Despite such a plethora of sensors that can be mounted on a UAV, the available options are quite reduced when considering fluvial environments in urban areas. For instance, the European Union recently imposed restrictions on the UAVs usable for civilian use (EU Regulations 2019/947 and 2019/945), in particular in urban areas. Such regulations are less restrictive for small UAVs, which recently led to an increasing interest of civilian UAV operators for mini-UAVs, e.g. weight $\leq$ 250 gr. Motivated by the latter considerations, this work aims at investigating the feasibility of using a mini-UAV for litter detection on the river banks, in a urban environment.

The paper is organized as follows: first, the considered case study area and the used UAV are described in Section 2, then the implemented strategy is presented in Section 3. Finally, the obtained results are shown in Section 4, and some conclusion are drawn in Section 5.

* Corresponding author

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W1-2022-117-2023

2. CASE STUDY

2.1 Study Area

The study area has been identified as a portion of the Mugnone river, in Florence (Italy). Fig. 1 shows part of the considered area (100 m long).

More than 150 litter objects, indicated with red marks in the figure, have been identified in the area shown in Fig. 1, with a variable spatial density, reaching even more than 9 objects per meter (linear density, e.g. on sections orthogonal to the river direction) Fig. 2.

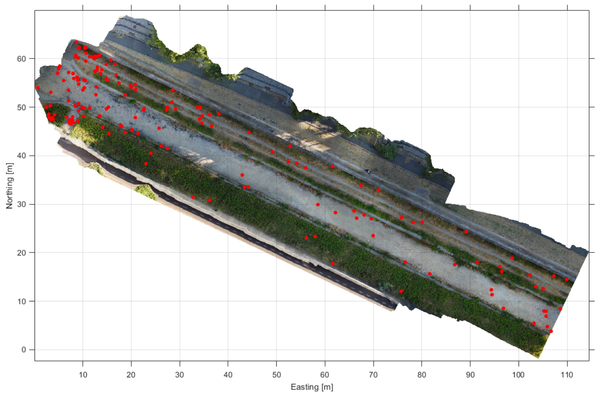


Figure 1. Litter objects (red marks) on the study area.

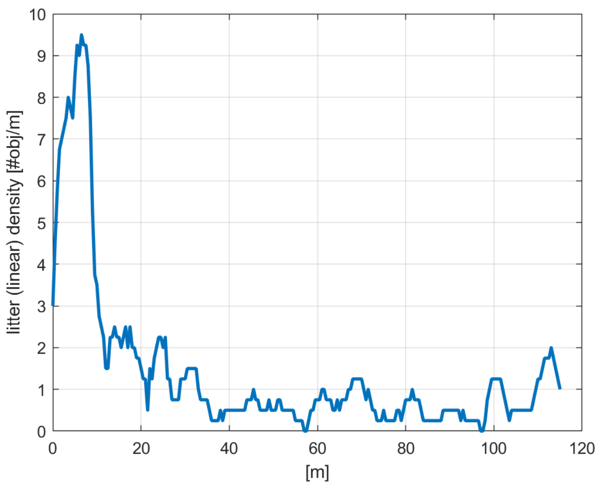


Figure 2. Litter (linear) density along the river direction in the case study area.

2.2 UAV and dataset

The RGB imagery has been acquired by means of a DJI Mini 2 UAV, shown in Fig. 3, flying at a varying altitude over the area of interest.

Table 1 summarizes the specific altitudes (with respect to the river bank level) considered during the data collection associated to the area shown in Fig. 1: on the one hand, the higher the flight altitude the less time is typically required to cover a specific area. On the other hand, the lower the flight altitude the higher is the spatial resolution of an image (i.e. the smaller is the GSD value). Hence, the rationale of such multi-altitude flights is that of investigating the ideal working conditions in such a way to maximize the litter detection performance while minimizing the flight time. Such investigation will be considered in our future work, while in this paper the litter detection is performed only at the lowest altitudes (20-30 m).

Figure 3. DJI Mini 2 drone.

Altitude [m]	GSD [cm]
20	1.0
30	1.6
40	2.1
60	3.1
80	4.2

Table 1. Ground Sampling Distance (GSD) as function of the flight altitude with respect to the river bank level.

Furthermore, a different survey campaign has been conducted in order to obtain a reliable, accurately geo-referenced, 3D reconstruction and an orthophoto of the whole area of interest (500 m long, along the river direction).

3. METHODS

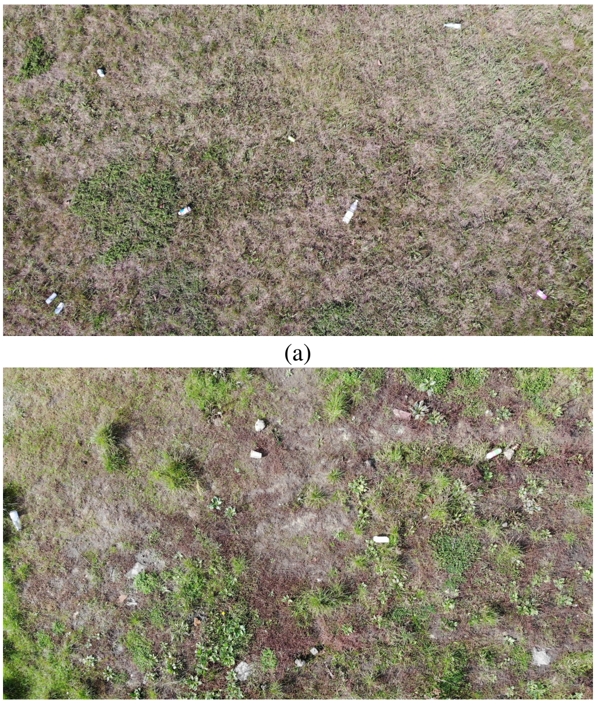
The proposed method is based on the use of a deep learning approach to detect litter objects from mini-UAV RGB imagery. The rationale is that the mini-UAV could be used to fly either automatically or remotely piloted over the area of interest. Then, the acquired imagery could be transmitted to a computing station and processed in near real-time.

The resolution of the acquired imagery is reduced to no more than 1 Mpixel in order to enable near real-time transmission and analysis. Furthermore, it is worth to notice that this work only investigates the analysis part, while the imagery transmission will be considered as a future development of the proposed approach.

Quick object detection has been implemented using a Yolo v4 network (Bochkovskiy et al., 2020): transfer learning from a Yolo v4 network originally pre-trained with the COCO (Common Objects in Context) dataset (Lin et al., 2014) has been done with some hundreds litter images, taken from a public database (Kraft et al., 2021). Fig. 4 shows a couple of examples of images taken from such public database.

Since the positions measured by the mini-UAV navigation system may be affected by an error at meter-tens of meter level, the visual information included in the acquired imagery is used in order to improve the positioning accuracy of the UAV.

A high resolution map of the area of interest is assumed to be available as an orthophoto obtained by the data acquired in a previous UAV-photogrammetry survey campaign. Such orthophoto is assumed to be accurately georeferenced, e.g. at centimeter level of accuracy, by using tens of control points properly distributed over the considered area, surveyed with geodetic GNSS receivers.



(a)

(b)

Figure 4. Examples of images taken from the training database.

Some additional information, such as the river bed and the river bank areas, is assumed to be available on such high resolution map, and hence also usable in order to select only the areas of interest (e.g. the river banks) to be inputted in the classification procedure.

The video frames acquired by the mini-UAV are processed by an ad hoc developed SLAM (Simultaneous Localization and Mapping, (Leonard and Durrant-Whyte, 1991)) algorithm.

The mini-UAV camera is assumed to be pre-calibrated, and the video frame are down-scaled to approximately 0.5 Mpixel before being inputted in the SLAM algorithm, in order to make the procedure executable in real-time.

To be more precise, the UAV imagery is assumed to be acquired with camera in nadir orientation, and hence, the nowadays quite standard visual ORB-SLAM approach proposed in (Mur-Artal et al., 2015) is modified in order to take into account of such prior information.

Including the nadir-camera view assumption in the SLAM algorithm allows to implement a two-point/three-point algorithm in order to determine a reasonable initial assessment of the relative camera pose between successive frames (Troiani et al., 2014, Masiero and Vettore, 2016), hence speeding up the RANSAC (Random Sample Consensus (Fischler and Bolles, 1981)) step, and the overall procedure. To be more specific, the two-point algorithm can be used in the UAV case assuming also a constant flight altitude (on the considered frames) (He and Habib, 2016, He et al., 2018). Indeed, the latter assumption allows to reduce the number of independent variables to be estimated when considering the relative pose between two camera frames (He and Habib, 2016, He et al., 2018):

$$p_2^\top [t]_\times R p_1 \tag{1}$$

where the equation has been written assuming that the lens distortions have already been corrected, K is the intrinsic matrix, m_i are the coordinates of a point in the i-th camera frame and $p_i = K^{-1} m_i$ are the corresponding normalized coordinates, R and t are the relative orientation and translation between the two cameras, which, assuming the z axis to be aligned with the vertical direction, in these specific working conditions can be written as:

$$t = \begin{bmatrix} t_x \\ t_y \\ 0 \end{bmatrix} \tag{2}$$

and

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where θ is the rotation angle around the z axis.

Finally, $[t]_\times$ is the skew-symmetric matrix that allows to compute the cross product of t with another vector as a matrix multiplication, i.e. $[t]_\times b = t \times b$ for any vector b. The reduced number of parameters to be estimated in the essential matrix E, obtained as $E = [t]_\times R$, is quite apparent. An additional constraint can also be imposed taking into account that E can only be assessed up to a scaling factor.

It is worth to notice that the vision SLAM approach allows only to track the UAV position in a local reference system: introducing some external information is needed in order to determine the geographic coordinates of the UAV. In this work two approaches are considered to such aim:

a) using the GPS measurements of the receiver mounted on the mini-UAV

b) exploit the already available orthophoto in order to determine some landmarks to be used for accurately georeferencing the UAV.

The performance of the GPS-assisted approach, in a), is clearly affected by the positioning error of the receiver mounted on the UAV itself. Such an error in certain cases may be unacceptable, hence b) has been implemented in order to improve the overall positioning performance.

The rationale in b) is that of properly recognizing corresponding points in the UAV images and on the previously computed orthophoto. This is done by exploiting ORB (Oriented FAST and rotated BRIEF (Rublee et al., 2011)) feature matching. Furthermore, an outlier rejection step is implemented by exploiting the already estimated geometry: indeed, the SLAM algorithm already provides 3D information about the camera locations and certain 3D points in the scene, despite in a local reference system.

In this work, the outlier rejection step is implemented by considering just the horizontal point locations (e.g. those on the orthophoto and those according to the SLAM). Nevertheless, the extension to exploiting 3D information will be considered in our future works.

Given a set of corresponding 2D points $\{p_{o,i}\}$ in the orthophoto and $\{p_{s,i}\}$ in the SLAM reference system, i.e. just the planar (x, y) coordinates in the SLAM reference system, then the georeferencing transformation is obtained as follows:

$$\{\hat{R}, \hat{t}, \hat{s}\} = \arg\min_{R,t,s} \sum_i ||p_{o,i} - (sRp_{s,i} + t)||^2 \quad (4)$$

4. RESULTS

The procedure for litter detection and localization described in the previous section is validated on a portion of the case study area shown in Section 2.

First, Table 2 shows the performance of the deep-learning based litter detection approach on the considered dataset, where precision, recall, accuracy and F1-score are defined as follows in terms of the number of false/true positives, (FP/TP) false/true negatives (FN/TN):

$$precision = \frac{TP}{TP + FP} \quad (5)$$

$$recall = \frac{TP}{TP + FN} \quad (6)$$

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

$$F1 - score = \frac{2(recall \times precision)}{recall + precision} \quad (8)$$

	precision	recall	accuracy	F1-score
DL detector	85.7%	80%	70.6%	82.8%

Table 2. Waste deep-learning detector.

Furthermore, Fig. 5 shows an example of the litter detection, whereas Fig. 6 shows a false positive and a false negative example of detection errors.

Figure 5. Example of litter detection on an image of the case study dataset.

Then, the UAV positioning performance is tested on the same dataset. First, the reliability of the nadir camera orientation assumption is checked: in order to obtain a reasonable assessment of the real camera orientation, a set of $\approx$ 150 camera frames have been processed with Agisoft Metashape, and, once properly georeferenced, their optical axis directions have been compared with the vertical (downward) direction, obtaining the results shown in Table 3.

(a)

(b)

Figure 6. Examples of detection errors: (a) false positive, (b) false negative.

	deviation from the vertical [deg]
mean	1.7
median	1.7
standard deviation	0.4
median absolute dev.	0.3

Table 3. Deviation from the vertical.

Then, the positioning performance is assessed for both the a) and b) approaches considered in the previous section. The results obtained for the a) and b) cases are reported in the first and second row, respectively, of Table 4. It is worth to notice that m and D in Table 4 indicate the median error and the median absolute deviation. Table 4 reports their values along the x and y axes, and their $2D$ value.

Furthermore, the last two rows in Table 4 show the results obtained by the SLAM approach when limiting the analysis to locations close (less than 5 m) to a point matched with the corresponding one in the orthophoto, and far (more than 20 m) far from any matched point.

	m_x [cm]	D_x [cm]	m_y [cm]	D_y [cm]	m_{2D} [cm]	D_{2D} [cm]
a) GPS	66	56	198	26	209	62
b) SLAM	53	63	30	22	61	67
b) close	29	12	21	5	36	13
b) far	143	6	87	12	168	13

Table 4. Localization error.

5. CONCLUSIONS

The paper considers the problem of litter detection on river banks and localization when using a mini-UAV, which can be a quite convenient solution in urban environments.

The litter detection is based on the use of a properly trained Yolo v4 network, and the obtained results are quite encouraging (precision = 85.7%, recall = 80%). Nevertheless, our future work will be dedicated to an improvement in particular in terms of reduction of the false negatives. To be more specific, in accordance with the results obtained in our tests, the trained network is not enough sensitive when dealing with certain transparent plastic objects, as shown in Fig. 6(b). It is also worth to notice that a similar criticality has already been observed also when dealing with plastic litter detection by using UAV multispectral imagery (Cortesi et al., 2022).

For what concerns the positioning results, Table 4 shows that, in the considered example, the UAV GPS-based approach leads to some meters of 2D positioning error, as expected (the positioning error may be even larger in other cases). Instead, the SLAM approach exploiting points matched with the corresponding ones in the orthophoto allows to reach sub-meter errors.

To be more precise, the localization error of the proposed approach is quite dependent on the distance from the points matched with the orthophoto ones, ranging from few decimeters to even more than a meter. This behavior is quite standard for any SLAM-based approach, where the assessed positions are expected to drift from the correct ones if reliable positioning updates are not available.

To conclude, it is worth to notice that the characteristic size of the litter objects is at decimeter level, hence the obtained localization error level is expected to be sufficient in real applications when the set of points matched with the orthophoto is reasonably well distributed over the area of interest.

Future investigations will also be dedicated to the analysis of the system performance, in terms of both litter detection and localization error, when varying the UAV flight altitude.

REFERENCES

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.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M., 2020. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Cheng, Y., Zhu, J., Jiang, M., Fu, J., Pang, C., Wang, P., Sankaran, K., Onabola, O., Liu, Y., Liu, D. et al., 2021. Flow: A dataset and benchmark for floating waste detection in inland waters. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10953–10962.

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.

Cortesi, I., Masiero, A., Tucci, G., Topouzelis, K., 2022. UAV-based river plastic detection with a multispectral camera. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 855–861.

da Costa, J. P., 2021. The 2019 global pandemic and plastic pollution prevention measures: Playing catch-up. *Science of The Total Environment*, 774, 145806.

Fischler, M., Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.

Gabrys, J., 2013. Plastic and the work of the biodegradable. *Accumulation*, Routledge, 222–241.

Geraeds, M., van Emmerik, T., de Vries, R., bin Ab Razak, M. S., 2019. Riverine plastic litter monitoring using unmanned aerial vehicles (UAVs). *Remote Sensing*, 11(17), 2045.

Gonçalves, G., Andriolo, U., 2022. Operational use of multispectral images for macro-litter mapping and categorization by Unmanned Aerial Vehicle. *Marine Pollution Bulletin*, 176, 113431.

He, F., Habib, A., 2016. Automated relative orientation of UAV-based imagery in the presence of prior information for the flight trajectory. *Photogrammetric Engineering & Remote Sensing*, 82(11), 879–891.

He, F., Zhou, T., Xiong, W., Hasheminnasab, S. M., Habib, A., 2018. Automated aerial triangulation for UAV-based mapping. *Remote Sensing*, 10(12), 1952.

Hsu, C.-W., Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, 13(2), 415–425.

Kraft, M., Piechocki, M., Ptak, B., Walas, K., 2021. Autonomous, Onboard Vision-Based Trash and Litter Detection in Low Altitude Aerial Images Collected by an Unmanned Aerial Vehicle. *Remote Sensing*, 13(5).

Kurtela, A., Antolović, N. et al., 2019. The problem of plastic waste and microplastics in the seas and oceans: impact on marine organisms. *Croatian Journal of Fisheries*, 77(1), 51–56.

Lebreton, L., Van Der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., Reisser, J., 2017. River plastic emissions to the world's oceans. *Nature communications*, 8(1), 1–10.

Leonard, J., Durrant-Whyte, H., 1991. Simultaneous map building and localization for an autonomous mobile robot. *Intelligent Robots and Systems '91. 'Intelligence for Mechanical Systems, Proceedings IROS '91. IEEE/RSJ International Workshop on*, 1442–1447 vol.3.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C. L., 2014. Microsoft coco: Common objects in context. *European conference on computer vision*, Springer, 740–755.

Maharjan, N., Miyazaki, H., Pati, B. M., Dailey, M. N., Shrestha, S., Nakamura, T., 2022. Detection of River Plastic Using UAV Sensor Data and Deep Learning. *Remote Sensing*, 14(13), 3049.

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.

Masiero, A., Vettore, A., 2016. Improved Feature Matching for Mobile Devices with IMU. *Sensors*, 16(8), 1243.

Merlino, S., Paterni, M., Berton, A., Massetti, L., 2020. Unmanned aerial vehicles for debris survey in coastal areas: Long-term monitoring programme to study spatial and temporal accumulation of the dynamics of beached marine litter. *Remote Sensing*, 12(8), 1260.

Mur-Artal, R., Montiel, J. M. M., Tardos, J. D., 2015. ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE transactions on robotics*, 31(5), 1147–1163.

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: a review. *Applied Geomatics*, 6(1), 1–15.

Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. Orb: An efficient alternative to sift or surf. *2011 International conference on computer vision*, IEEE, 2564–2571.

Schmidt, C., Krauth, T., Wagner, S., 2017. Export of plastic debris by rivers into the sea. *Environmental science & technology*, 51(21), 12246–12253.

Suykens, J. A., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293–300.

Taddia, Y., Corbau, C., Buoninsegni, J., Simeoni, U., Pellegrinelli, A., 2021. UAV Approach for Detecting Plastic Marine Debris on the Beach: A Case Study in the Po River Delta (Italy). *Drones*, 5(4), 140.

Themistocleous, K., Papoutsa, 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).

Topouzelis, K., Papageorgiou, D., Suaria, G., Aliani, S., 2021. Floating marine litter detection algorithms and techniques using optical remote sensing data: A review. *Marine Pollution Bulletin*, 170, 112675.

Troiani, C., Martinelli, A., Laugier, C., Scaramuzza, D., 2014. 2-point-based outlier rejection for camera-IMU systems with applications to micro aerial vehicles. *2014 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 5530–5536.