Multi-Resolution Remote Sensing Dataset for the Detection of Anthropogenic Litter: A Multi-Platform and Multi-Sensor Approach

Robert Rettig [1,*], Felix Becker [1], Alexander Berghoff [2], Tobias Binkele [2], Wolfram Michael Butter [1], Tilman Floehr [3], Martin Kumm [4], Carolin Leluschko [1], Florian Littau [2], Elmar Reinders [2], Eike Rodenbäck [1], Tobias Schmid [4], Sabine Schründer [3], Sören Schweigert [2], Michael Sinhuber [2], Jens Wellhausen [4], Frederic Stahl [1] and Christoph Tholen [1]

1 German Research Center for Artificial Intelligence, 26129 Oldenburg, Germany; frederic_theodor.stahl@dfki.de (F.S.)
2 Optimare Systems GmbH, 27572 Bremerhaven, Germany; michael.sinhuber@optimare.de (M.S.)
3 everwave GmbH, 52062 Aachen, Germany
4 Department of Engineering Sciences, Jade University of Applied Sciences, 26389 Wilhelmshaven, Germany
* Correspondence: robert.rettig@dfki.de

Abstract

The dataset developed within the PlasticObs+ project aims to facilitate a multi-resolution approach for detecting and quantifying anthropogenic litter through areal images. Traditional detection methods often suffer from narrow, use-case-specific limitations, reducing their transferability. To address this, an image dataset was created featuring various spatial and spectral resolutions. The highest spatial resolution images (ground sampling distance = 0.2 cm) were used to generate a labeled dataset, which was georeferenced for mapping onto coarser-resolution images.

Dataset: https://doi.org/10.5281/zenodo.14844219.

Dataset License: CC-BY

Keywords: anthropogenic litter; plastic litter pollution; litter object detection; dataset annotation; multi-resolution; multi-sensor; multi-class; multispectral data; remote sensing

1. Summary

The dataset presented here was obtained within the project PlasticObs+. Its main focus lies in the efficient combination of overview scans (with a lower spatial resolution but a larger field of view) and a high-resolution, multispectral camera, employing artificial intelligence techniques for the automated detection and precise identification of litter objects in marine environments [1]. Therefore, airborne data collection was performed using two distinct platforms: a fixed-wing research aircraft for large-scale, moderate-resolution surveys and unmanned quadrocopter systems (hereafter referred to as "drone") for high-resolution, site-specific mapping. This approach leveraged the strengths of each platform, with the aircraft enabling rapid coverage of extensive areas and the drone providing detailed imagery of selected sites.

To obtain reliable data to test the multi-platform and multi-sensor approach, a field campaign was conducted from 25 to 28 June 2024 on an open-air festival camping area in

Northern Germany, directly after a festival took place. Due to the limited flying range of the research airplane and the insufficient amount of waste in the nearby areas of Northern Germany, sufficient samples of litter objects could be gathered at the festival location. To ensure data availability, several flights were performed with sensor-equipped drones and an airplane, collecting RGB (red–green–blue) and multispectral data for a total of four measurement days (see Tables 1 and 2).

Table 1. A summary of the metadata for the drone-based dataset.

Name	Resolution	Dates	Time (UTC)	Georeferenced	Type
Label Dataset RGB High-Resolution (660 images)	0.2 cm [1]	25.06.2024	10:35–13:29	Yes	PNG
	0.2 cm [1]	26.06.2024	08:26–13:39	Yes	
	0.2 cm [1]	27.06.2024	06:37–08:45	Yes	
	0.2 cm [1]	28.06.2024	09:44–14:09	No	
Label Dataset Multispectral (blue, green, red, red edge, near-infrared)	2.8 cm	25.06.2024	11:45–13:29	Yes	TIF
	4.7 cm	25.06.2024	12:45–13:05	Yes	
	0.89 cm	26.06.2024	10:16–11:34	Yes	
	2.8 cm	26.06.2024	11:36–11:47	Yes	
	4.7 cm [1]	26.06.2024	11:36–11:47	Yes	
	0.89 cm	27.06.2024	07:28–09:15	Yes	
	2.8 cm	27.06.2024	08:15–08:26	Yes	
	4.7 cm	27.06.2024	08:35–08:43	Yes	
Orthomosaics	2.74 cm	25.06.2024	11:45–13:29	Yes	TIF
	2.79 cm	26.06.2024	11:36–11:47	Yes	
	2.80 cm	27.06.2024	08:15–08:26	Yes	
	2.82 cm	28.06.2024	08:09–08:15	Yes	

[1] Resampled using bilinear interpolation.

Table 2. Airplane-based data.

| Name	Resolution	Dates	Georeferenced	Type
20240625_001_HRVIS_162	-	25.06.2024	No	PNG
	15.2 cm	25.06.2024	Yes	TIF
20240626_002_HRVIS_119	-	26.06.2024	No	PNG
20240628_003_HRVIS_45	-	28.06.2024	No	PNG

The structure of this paper is divided into the data description and the methodology that was used for preprocessing the data. The data description explains the ground truth, drone, and airplane data while addressing other available data sources. The methodology integrates these datasets, and the processing steps involved mapping and analyzing litter distribution across four locations over four days.

2. Data Description

Data taken by drone were separated into label datasets and the orthomosaics. The label dataset as georeferenced based on the orthomosaics to ensure a correct overlay for the scaling approaches. Due to the size constraint of the online storage, the highest resolution for the drone data was set to 0.2 cm. Raw data can be provided upon request.

2.1. Drone Data

Using sensor-equipped drones, four different resolutions were collected from four sampling sites, as shown in Figure 1. A set of high-resolution images was acquired by flying a DJI Mavic 2 Enterprise Advanced drone [2] at approximately 4 m above ground level over individual 2 × 2 m sampling plots (see Figure 2). The integrated camera system was used to capture high-resolution RGB images. The only exception to this was on 27 June 2024, when a drone of type DJI Matrice 210 V2 [3], equipped with the DJI XT2 camera system [4], was used to acquire high-resolution images. These images were resampled to a unified ground sampling distance of 0.2 cm to label the objects with a set of classes, as defined in Table A1. DJI Matrice 210 V2, equipped with the multispectral MicaSense Altum V04 camera system [5], was utilized for data acquisition at three distinct altitudes (20 m, 60 m, and 100 m) according to the specifications outlined in Section 3.1, resulting in the normalized spatial resolutions presented in Table 1.
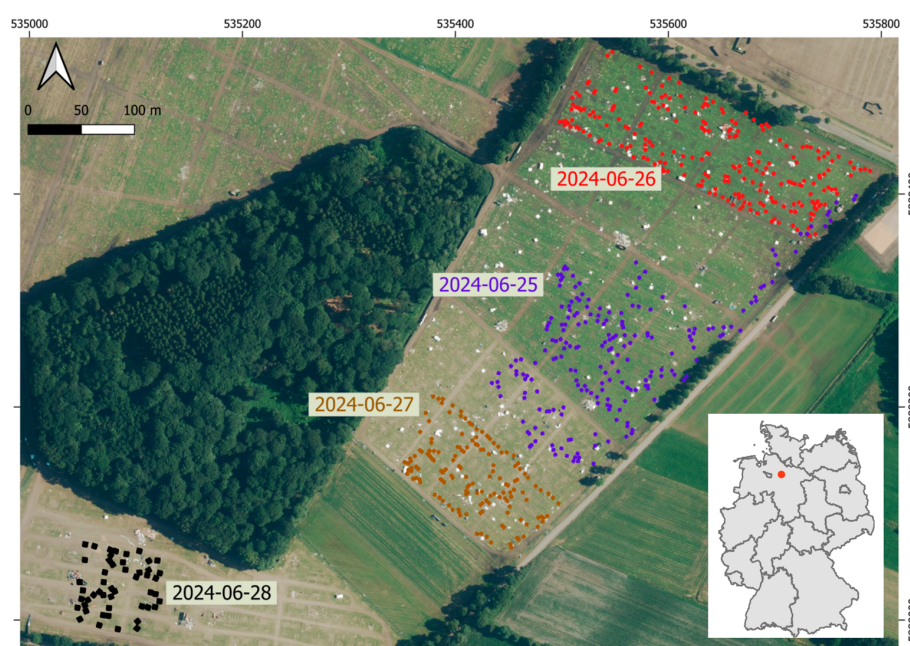
Figure 1. An overview map of the sampling areas during the field campaign (Map projection: UTM 32N). The high-resolution images captured are displayed as points, color-coded according to the date of capture. Background source: ATKIS-DOP [6].
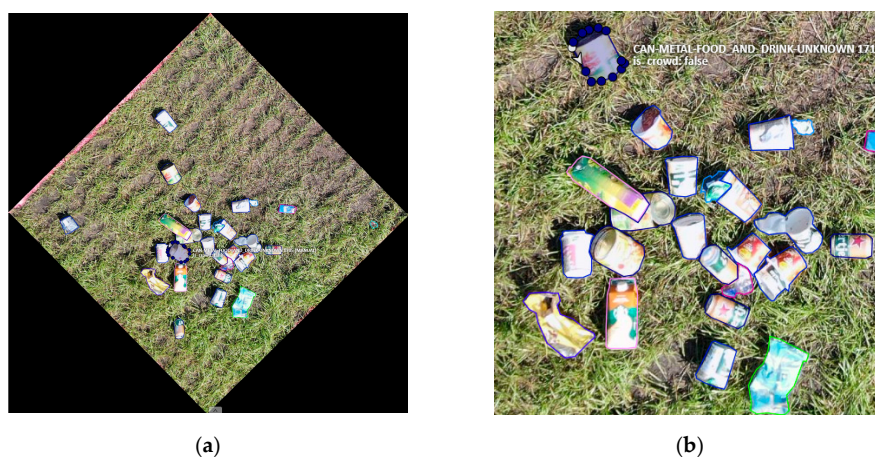
Figure 2. Example of annotation dataset with high-resolution RGB images: (**a**) full PNG image, as georeferenced data, cropped in position of reference system UTM 32N; (**b**) detailed view of labeled data.

On 28 June 2024, the data acquisition method had to be changed, and no ground control points (GCPs) were taken for the individual sampling plots. As a result, georeferencing the high-resolution dataset was not possible for this date.

2.2. Airplane Data

The aerial data acquisition procedure was conducted using the fixed-wing research aircraft "Jade One", a Diamond HK36-TTC ECO motor glider, operated by the Jade University of Applied Sciences. This specialized platform is designed for scientific missions and features two certified underwing pods, which can be outfitted with scientific equipment. For the measurements presented in this study, the airplane was equipped with a VIS Line Scanner, a sensor developed by Optimare Systems GmbH, modified to suit the needs of the PlasticObs+ project [7]. It is a single-line RGB sensor with a resolution of 4096 px and an acquisition rate of up to 500 Hz. The field of view of the VIS Line Scanner is 94 degrees, yielding a ground resolution of up to 15 cm at a 1000 ft airplane altitude. The obtained lines are color-corrected, timestamped, and corrected for the roll movement of the airplane by utilizing the location and orientation information obtained with an IMU (Inertial Measurement Unit).

On 25 June 2024, stacked VIS Line Scanner imagery was successfully georeferenced using a third-degree polynomial transformation based on 39 GCPs. This approach resulted in a mean error of 18.77 px in the georeferenced raster outputs. In contrast, imagery acquired on other days exhibited pronounced roll movements, attributed to the elasticity of the airplane wings, which led to substantial image distortion. While georeferencing these datasets was theoretically possible, the resulting spatial accuracy was found to be insufficient for reliable analysis, and, thus, georeferenced products were not generated for those dates. A summary of the airplane datasets can be seen in Table 2.

2.3. Additional Data

The DOP20 dataset includes a digital orthomosaic ;Tile ID: 325345888; acquisition date: 26 June 2024, obtained from the NI-LGLN OpenGeodata portal [6]. This dataset offers orthomosaics at a spatial resolution of 20 cm, comprising four distinct spectral bands (blue, green, red, near-infrared), with a geometric accuracy of $+/- 0.4$ m.

Sentinel-2 L2A satellite data can be downloaded from the Copernicus Data Space Ecosystem Browser [8]. For the Area of Interest (UTM 32N: x_{min} 535846, x_{max} 535856, y_{min} 5889194, y_{max} 5889204), cloud-free observations for exemplary dates on 20 June 2024, 25 June 2024, and 27 June 2024 are provided. Geometric accuracy depends on applied refining steps.

Additionally, weather data from DWD stations 4745 and 4275 [9] can be included to provide essential context and support for interpreting satellite and orthomosaic data. Incorporating this weather information allows researchers to account for the environmental conditions present during data collection periods.

3. Methods for Data Acquisition and Processing

In this study, we employed an integrated methodology that combined data acquired from multiple platforms and sensors to detect anthropogenic litter. By fusing these diverse datasets, comprehensive multi-resolution mapping was achieved, with the aim of enabling robust and scalable analytical outcomes. Four sampling locations were considered on four consecutive days, as shown in Figure 1.

3.1. Drone Data Acquisition

Drone imagery was acquired with both front (longitudinal) and side (lateral) overlaps set to 80%. This high degree of overlap was selected to ensure comprehensive coverage

and continuity between adjacent images, which is essential for generating detailed and accurate three-dimensional (3D) surface models. The substantial overlap enables each ground feature, including targets and objects of varying heights, to be captured from multiple perspectives at all considered heights. The drone flight plan was created with DJI Pilot, optimized upon the custom camera settings for the multispectral camera system Micasense Altum V04 [5]. The 3D maps were generated with Pix4D version 4.6.4 [10] and radiometrically calibrated using the option "Camera, Sun Irradiance, and Sun Angle using DLS IMU" and images from the reference panel. After generating the orthomosaics, georeferencing was performed by utilizing QGIS version 3.38.2 [11] to map targets and objects on the ground.

The georeferencing of drone-acquired imagery was conducted through an iterative process, beginning with datasets of the lowest spatial resolution, typically collected at an altitude of 100 m. For this initial step, prepared GCPs measured with Real-Time Kinematic GPS were utilized, ensuring comprehensive coverage of the sampling area for each survey day. Subsequently, higher-resolution datasets captured at lower altitudes were aligned using the previously georeferenced maps as references, with manual identification of corresponding ground control points to refine the spatial accuracy.

Throughout the georeferencing workflow, the mean pixel error associated with the transformation remained low (e.g., 4.2×10^{-12} px). Any spatial shifts resulting from camera distortions were systematically addressed, and, where necessary, the georeferencing procedure was repeated to guarantee the consistent alignment of features across all levels of image detail. For downstream analysis, the thermal channel, characterized by its lower spatial resolution, was excluded from the set of georeferenced orthomosaics, allowing the focus to remain on the assessment of high-resolution multispectral data.

3.2. Annotation of the Drone Data

For labeling the waste objects, the single high-resolution RGB images, taken at a low altitude, were first georeferenced to be mapped on the other flight altitudes datasets and masked afterward by consistent 2×2 m shapes.

The further processing of the orthomosaics, including masking by the 2×2 m shapes, normalizing using the full sensor range, resampling to 0.2 cm, and reformatting, were performed with R version 4.3.2 using package terra and sf [12]. The processed images, as displayed in Figure 2a, were formatted to PNG and uploaded to an annotation tool [13], labeled with predefined classes, and checked for inconsistencies; see Appendix A: Table A1.

Consistent label alignment during the integration of multi-resolution datasets was ensured using a developed annotation harmonization framework. In this framework, bounding box coordinates of the COCO-format annotations were scaled via affine transformations across spatial resolutions. This involved programmatically adjusting annotation coordinates using scale factors derived from ground sampling distance ratios while preserving topological relationships between features [14].

Additionally, employing the same framework, the annotation categories in the labeling dataset were clustered according to their corresponding material types (Metal, Paper, Plastic, Others, as illustrated in Appendix A: Table A2), as well as into a binary one-class categorization. This multi-class approach enabled a broad use-case and the research-specific usage of the annotation dataset. For objects which could not accurately attached to a specific object, material, category, or material class, the term "others" was used. The object classes for "material" and "binary" were each clustered from the annotation for all object classes. For example, the object classes "X...-Plastic-X..." were merged into the material class "Plastic".

*3.3. VIS/Airplane Data Acquisition*

The VIS Line Scanner recorded RGB line data at 4096 px with a variable frame rate of up to 500 Hz. The frame rate was dynamically adjusted with respect to the airplane speed and altitude to ensure that individual pixels corresponded to rectangular areas on the ground. The sensor was operated on board using a laptop to coordinate data acquisition and timing. Using the timing information, lines from the respective channels could be combined to obtain traditional images. As the individual color channels on the chip were physically separated, their recording times had to be staggered to overlap the individual resulting images. The sensor was oriented to record lines on the ground perpendicular to the flight direction. This maximized the covered ground area by reducing recording overlap, but this also introduced a sensitivity with respect to the roll movement of the airplane. As the sensor was mounted rigidly inside the cargo pod, the rolling movement of the airplane resulted in distortions of the recorded images. To address this, the airplane's navigation data were recorded using the Nginuity DAQAHRS-IMU [15]. Based on this information, the individual scan lines were shifted to correct for the airplane's roll motion, thereby producing geometrically undistorted images. The 25 June 2024 scene was georeferenced and aligned with drone orthomosaics using the methodology detailed in Table 2. Imagery from subsequent dates exhibited significant roll movements due to wing elasticity, causing substantial distortion. While georeferencing remained technically feasible, the resulting spatial accuracy was deemed insufficient for reliable analysis. Consequently, no georeferenced products were generated for these dates.

4. Experimental Verification

The verification of the published dataset was conducted within the framework of the PlasticObs+ project, which developed a two-stage artificial intelligence system for the detection of plastic waste in diverse environments [1]. The system integrates a Variational Autoencoder for rapid onboard anomaly detection as the first stage and a Mask R-CNN-based segmentation module as the second stage for the high-resolution multispectral detection of objects. This modular architecture enabled the system to achieve both speed and precision and supported adaptability to future sensor technologies.

Performance validation was carried out using drone imagery acquired at different altitudes. Models trained on drone data recorded at 20 m altitude (GSD = 9 mm), which closely matched the high-resolution sensor technology for the airplane, exhibited particularly high detection performance with F1 scores of 0.80. Models trained from drone images at a 4 m altitude (GSD = 0.2 cm) also achieved strong results, with the model 04m_binary_RGB reaching an F1 score of 0.77 (see Figure 3). These findings suggest that segmentation approaches are less effective at lower resolutions (e.g., 60 m with an F1 score < 0.6), and that high-resolution sensor technologies benefit significantly from the two-stage detection process.

The interdisciplinary and adaptable architecture of the PlasticObs+ system bridges the gap between basic research and practical environmental monitoring. The system is well suited to support efficient cleanup operations in threatened ecosystems, as well as commercial applications such as post-event cleanup management. The main emphasis of this dataset is on scalability and interoperability across multiple sensor systems, hence the addition of satellite and DOP20 data to the airplane and drone dataset. Thus, further integration with other datasets and use cases can increase the detection capacity, applicability, and generalizability.
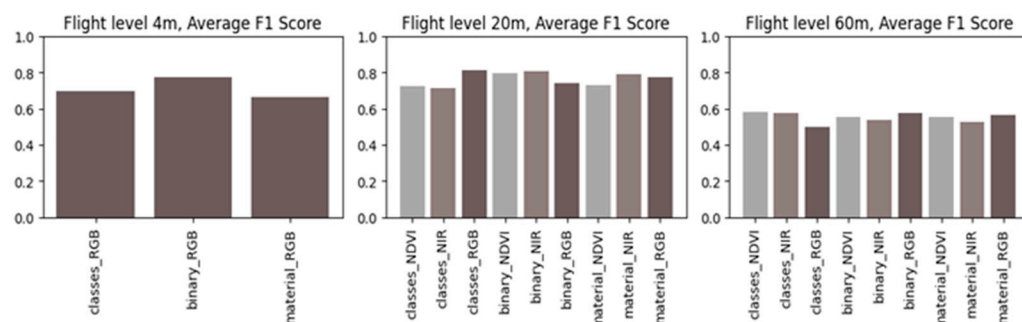
Figure 3. The F1 scores for the detection of waste objects with the label dataset RGB and multispectral combinations at different altitudes colour-coded according to the channels used.

5. User Notes

In the use case of litter detection within the PlasticObs+ project, while working with georeferenced multi-resolution remote sensing datasets, which are taken at different times, it is necessary to test for homogeneity, e.g., the moving or disappearing of lightweight objects. Small displacements, based on inconsistencies among lenses and patterns while creating orthomosaics, should be emphasized. Therefore, the annotations for images with coarser resolutions are double-checked, so that misregistered or moved waste pieces are sorted out. Since newly introduced objects (e.g., bags or paper) that appear in the marked sampling areas between consecutive flights often cannot be reliably distinguished from pre-existing features due to sensor resolution limitations, these ambiguous instances were assigned the label "Others-Others-Others-Unknown" in this study.

While working on multi-resolution data, one objective was to process high-resolution images with annotations by adapting them down to lower spatial resolutions. Challenges were encountered due to some missing coverage of orthomosaics; thus, some lower-resolution masked images were unavailable. Therefore, a conditional data validation logic was added to the PyTorch (Version 2.4.1) dataloader to ensure compatibility between images and annotations. Annotations were automatically excluded when their corresponding images were unavailable, and vice versa. This bidirectional validation ensured uninterrupted dataloader iteration, preventing runtime errors while maintaining script execution efficiency.

Abbreviations

The following abbreviations are used in this manuscript:

GCP	Ground control point
IMU	Inertial Measurement Unit
RGB	Red-green-blue
RTK	Real-Time Kinematic

Appendix A

Table A1. Frequency of labels at 0.2 cm for all classes.

Label (Object_Type-Material_Type-Main_Category-Condition [1])	Frequency
Others-Others-Others	2348
Piece-Paper-Unknown	1135
Piece-Plastic-Unknown	931
Can-Metal-Food_and_Drink	804
Cup-Plastic-Food_and_Drink	755
Wrapper-Plastic-Food_and_Drink	598
Bags_transparent-Plastic-Cleaning_and_Cosmetic	339
Bags_transparent-Plastic-Food_and_Drink	217
Bottle-Plastic-Food_and_Drink	160
Tetra_Pak-Plastic-Food_and_Drink	152
Bags_non_transparent-Plastic-Cleaning_and_Cosmetic	139
Piece-Cardboard-Logistic_and_Transport	130
Box-Cardboard-Logistic_and_Transport	115
Fabric-Plastic-Unknown	113
Bottle-Glas-Food_and_Drink	101

Table A1. *Cont.*

Label (Object_Type-Material_Type-Main_Category-Condition [1])	Frequency
Bags_non_transparent-Plastic-Food_and_Drink	88
Cap-Plastic-Food_and_Drink	86
Cup-Cardboard-Food_and_Drink	74
Container-Plastic-Food_and_Drink	74
Tray-Plastic-Food_and_Drink	63
Bowl-Plastic-Food_and_Drink	62
Box-Cardboard-Food_and_Drink	61
Textiles-Textiles-Unknown	58
Piece-Metal-Unknown	57
Shoe-Textiles-Clothing	53
Canister-Plastic-Food_and_Drink	42
Food-Organic-Food_and_Drink	35
Tray-Cardboard-Food_and_Drink	33
Foam-Plastic-Logistic_and_Transport	30
Bottle-Plastic-Cleaning_and_Cosmetic	28
Lid-Plastic-Food_and_Drink	24
Container-Plastic-Cleaning_and_Cosmetic	23
Rope-Unknown-Unknown	20
Straw-Plastic-Food_and_Drink	15
Medical_package-Plastic-Cleaning_and_Cosmetic	13
Sponge-Plastic-Cleaning_and_Cosmetic	13
Piece-Rubber-Unknown	9
Bottle-Metal-Cleaning_and_Cosmetic	9
Coal-Others-Unknown	9
Lid-Metal-Food_and_Drink	7
Piece-Organic-Natural	7
Piece-Glas-Unknown	6
Pipe-Plastic-Construction	5
Piece-Lumber-Construction	3
Furniture-Lumber-Others	3
Other_Net-Plastic-Unknown	3
Trouser-Textiles-Clothing	2
Piece-Ceramic-Unknown	1

[1] The last placeholder, "Unknown", designed for a potential annotation of the condition, is left out for readability in this table. For other frequencies, depending on the resolution available for the individual days, please see the dataset folder "Frequencies".

Table A2. Frequency of label at 0.2 cm for clustered material type categories.

Label [1]	Frequency
1_Plastic	3974
300_Others	2655
5_Paper	1548
7_Metal	877

[1] Due to the nature of AI architectures, the classes for training are usually translated into consecutive integers for class interpretation and prediction. To prevent confusion while merging and interpreting datasets, we created a unique ID for each category, ensuring that assigned classes were harmonized, even if other datasets contained more or other material type categories.

References

1. Tholen, C.; Wolf, M.; Leluschko, C.; Zielinski, O. Machine Learning on Multisensor Data from Airborne Remote Sensing to Monitor Plastic Litter in Oceans and Rivers (PlasticObs+). In Proceedings of the OCEANS 2023—Limerick, Limerick, Ireland, 5–8 June 2023; IEEE: Limerick, Ireland, 2023; pp. 1–7.
2. DJI Mavic 2 Enterprise Advanced. Available online: https://www.dji.com/uk/support/product/mavic-2-enterprise-advanced (accessed on 10 March 2025).
3. DJI Matrice 210 V2. Available online: https://www.dji.com/uk/support/product/matrice-200-series-v2 (accessed on 10 March 2025).
4. DJI Zenmuse XT2. Available online: https://www.dji.com/uk/support/product/zenmuse-xt2 (accessed on 10 March 2025).
5. Micasense Altum V04. Available online: https://support.micasense.com/hc/en-us/articles/360010025413-Altum-Integration-Guide (accessed on 10 March 2025).
6. NI-LGLN ATKIS-DOP. Available online: https://opengeodata.lgln.niedersachsen.de (accessed on 10 March 2025).
7. VIS Line Scanner. Available online: https://www.optimare.de/fileadmin/optimare/pdf/Products/Product_FEK_VIS_210414jg.pdf (accessed on 10 March 2025).
8. Copernicus Data Space Ecosystem Browser. Available online: https://dataspace.copernicus.eu/browser/ (accessed on 10 March 2025).
9. DWD (Deutscher Wetterdienst) Meteorological Data. Available online: https://www.dwd.de/DE/leistungen/cdc/cdc_ueberblick-klimadaten.html (accessed on 10 March 2025).
10. PIX4Dmapper (4.6.4), Professional Photogrammetry Software for Drone Mapping. Available online: https://www.pix4d.com/product/pix4dmapper-photogrammetry-software (accessed on 15 March 2025).
11. QGIS (3.38.2), Geographic Information System. Available online: http://qgis.org (accessed on 15 March 2025).
12. R (4.3.2). A Language and Environment for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 15 March 2025).
13. CVAT (2.16.1). Computer Vision Annotation Tool. Available online: https://zenodo.org/doi/10.5281/zenodo.3497105 (accessed on 31 March 2025).
14. Rettig, R.; Becker, F. Adapting Annotation Datasets. Available online: https://github.com/DFKI-NI/Adapting_Annotation_Datasets (accessed on 26 February 2025).
15. Nginuity DAQAHRS. Available online: https://www.nginuity.com/page/Nginuity_Products/DAQ-SERIES/DAQAHRS/ (accessed on 20 March 2025).