OPERATIONAL RELIABILITY OF UAV-BASED WASTE SITE DETECTION FOR ENVIRONMENTAL MONITORING
/
FIABILITATEA OPERAȚIONALĂ A IDENTIFICĂRII ZONELOR CU DEȘEURI UTILIZÂND UAV-URI ÎN MONITORIZAREA MEDIULUI

Adrian IOSIF [1)], Edmond MAICAN[1)], Sorin-Ștefan BIRIȘ[1)], Gabriel-Alexandru CONSTANTIN[1)], Mihai-Gabriel MATACHE[2)]

[1)]Department of Biotechnical Systems, Faculty of Biotechnical Systems Engineering, National University of Science and Technology Politehnica Bucharest, 060042 Bucharest, Romania

[2)]INMA Bucharest/Romania

Tel.: 0723700628; Email: aiosif@yahoo.com

DOI: https://doi.org/10.35633/inmateh-79-82

***Keywords:** environmental monitoring; illegal dumping detection; UAV inspection; change detection; deep learning; operational reliability.*

ABSTRACT

Illegal dumping in peri-urban and natural environments poses persistent challenges for environmental authorities, particularly when large areas must be inspected with limited personnel and budget. Reliable and scalable monitoring approaches are essential for practical environmental management. Unmanned Aerial Vehicles (UAVs) offer rapid spatial coverage, yet it remains unclear how complex the analysis must be to reliably detect waste sites in practice. This controlled pilot study evaluated the operational reliability of three UAV-based image analysis approaches (global statistical screening, classical change detection with colour filtering, and deep learning object recognition) under controlled field conditions. The field experiment used controlled contamination conditions so that the evaluation reflected practical detection reliability rather than algorithm benchmarking. The statistical method could not reliably identify localised waste, while change detection achieved moderate reliability (~72%; 29 of 40 evaluated acquisitions) but required reference imagery. The deep learning model achieved the highest site detection reliability (~93%; 37 of 40 evaluated acquisitions) and worked without baseline images. Its main advantage was improved object completeness rather than better identification of contaminated areas. These observations showed that monitoring workflow design matters more than increasing algorithmic complexity. A tiered strategy that combines broad screening with focused high-accuracy analysis can reduce inspection effort without compromising the results of inspection. The evaluated pilot framework supported the selection of monitoring approaches according to available resources, and the findings apply to the controlled, low-altitude field conditions (15 m above ground level, single test site) examined here rather than to fully validated multi-site operation.

REZUMAT

Depozitarea ilegală a deșeurilor în zone periurbane și în mediul natural reprezintă o provocare continuă pentru autoritățile de mediu, în special atunci când trebuie inspectate suprafețe extinse cu personal și bugete limitate. În acest scop, sunt necesare metode de monitorizare fiabile și scalabile pentru o gestionare practică a mediului. Vehiculele aeriene fără pilot (UAV) oferă o acoperire spațială rapidă, însă rămâne încă neclar cât de complexă trebuie să fie analiza pentru a detecta în mod fiabil siturile de deșeuri. Acest studiu-pilot, desfășurat în condiții controlate, a evaluat fiabilitatea operațională a trei abordări de analiză a imaginilor bazate pe UAV (analiza statistică globală, detecția clasică a schimbărilor cu filtrare de culoare și recunoașterea obiectelor prin învățare profundă) în condiții de teren controlate. Experimentul a utilizat condiții de contaminare controlate, astfel încât evaluarea să reflecte fiabilitatea practică a detecției, și nu performanța algoritmilor în sine. Metoda statistică nu a putut identifica în mod fiabil deșeurile localizate, în timp ce analiza detecției schimbărilor a atins o fiabilitate moderată (~72%; 29 din 40 de achiziții evaluate), dar a necesitat imagini de referință. Modelul de învățare profundă a obținut cea mai ridicată fiabilitate în detectarea siturilor cu deșeuri (~93%; 37 din 40 de achiziții evaluate) și a funcționat fără imagini de referință. Principalul său avantaj a fost îmbunătățirea prin completare a setului de obiecte detectate, mai degrabă decât o identificare mai bună a zonelor contaminate. Aceste observații arată că proiectarea fluxului de monitorizare este mai importantă decât creșterea complexității algoritmice. O strategie pe niveluri, care combină o analiză generală cu o analiză focalizată pe obținerea unei acuratețe mai mari, poate reduce efortul de inspecție fără a compromite rezultatele inspecției.

Cadrul de lucru evaluat sprijină selectarea metodelor de monitorizare în funcție de resursele disponibile, iar rezultatele sunt valabile pentru condițiile de teren controlate, la altitudine redusă (15 m deasupra solului, un singur amplasament de test), examinate în acest studiu-pilot.

INTRODUCTION

Illegal dumping in peri-urban areas, waterways, agricultural fields, and natural reserves remains a common environmental problem. Ground-based inspections are labour-intensive, time-consuming, and often provide limited spatial coverage.

Illegal dumping affects soil, groundwater, biodiversity, and public health and is regulated within several European and international policy frameworks. The European Union Waste Framework Directive (Directive 2008/98/EC) establishes the legal basis for waste prevention, monitoring, and remediation obligations within Member States. Locating illegal disposal sites is required for regulatory compliance and enforcement. Waste monitoring is also aligned with the United Nations Sustainable Development Goals, particularly SDG 11 (Sustainable Cities and Communities) and SDG 12 (Responsible Consumption and Production), which promote environmentally sound waste management. Scalable and cost-efficient inspection methods therefore support environmental governance and sustainable urban management (*European Parliament and Council Directive, 2008; United Nations, 2015*).

UAV surveys combined with machine learning provide a practical alternative to ground inspections by enabling rapid coverage and automatic identification of potential waste hotspots. UAVs offer high spatial resolution and have become an established tool in environmental monitoring (*Manfreda et al., 2018*). However, their effectiveness depends not only on detection algorithms but also on acquisition strategy and survey design (*Torresan et al., 2017*). Recent studies have explored deep learning approaches for waste detection from aerial imagery. Reported results show promising detection accuracy but also reveal practical limitations such as dataset size requirements, variability of acquisition conditions, and lack of standardized datasets (*Youme et al., 2021; Sliusar et al., 2022; Mei et al., 2023; Wang et al., 2024*).

Experimental applications further illustrate these constraints. Deep neural networks have been used for marine litter detection and coastal debris mapping, where performance depends strongly on spatial resolution and scene complexity (*Fallati et al., 2019; Moy et al., 2018*). Other studies focus on spatial distribution mapping rather than individual object identification, highlighting the role of UAV surveys in environmental assessment workflows (*Gonçalves et al., 2020*).

Beyond object detection, UAVs have also been used for volumetric estimation and structural monitoring of waste sites. Applications include landfill volume modelling, settlement analysis, and combined UAV–terrestrial laser scanning workflows for improved topographic characterisation (*Filkin et al., 2022; Baiocchi et al., 2019; Son et al., 2020*).

Plastic waste detection has received particular attention across aquatic and terrestrial environments. Approaches range from deep semantic segmentation and hyperspectral discrimination to multisensor fusion and thermal infrared imaging (*Jakovljevic et al., 2020; Balsi et al., 2021; Maharjan et al., 2022; Topouzelis et al., 2019; Goddijn-Murphy et al., 2022; Yang et al., 2019*). These studies demonstrate that detection performance is influenced not only by algorithm choice but also by sensor characteristics, spatial resolution, and acquisition conditions. In UAV-based environmental applications, spatial resolution and acquisition geometry must be matched to the monitoring objective. Ground sampling distance and revisit frequency directly affect detectability, particularly for small or partially reflective objects such as plastic waste (*Pîrjol, 2018*).

Although UAV and AI tools are increasingly used, authorities still face uncertainty regarding how much analytical complexity is actually required in routine monitoring. In many situations, simpler approaches may already provide sufficient information, particularly when inspection resources are limited (*Gheres et al., 2025*).

This study therefore examines monitoring methodology rather than proposing a new detection algorithm. Three levels of analysis are compared in an operational context: basic statistical screening, change detection using classical computer vision techniques, and deep learning–based object recognition. The aim is to determine the minimum level of analysis needed for reliable site identification and to support the selection of appropriate monitoring strategies.

The main novelty of this work is therefore twofold and is stated here explicitly. First, the study provides an operational comparison of three distinct levels of image analysis (global statistical screening, classical change detection with colour filtering, and deep learning object recognition) applied to the same imagery, acquired with the same flight parameters and assessed with the same decision criterion, instead of comparing values reported on different benchmark datasets.

Second, the study separates two performance concepts that are usually reported together: site-level inspection triggering, that is, whether a survey correctly decides that a field inspection is needed, and object-level detection performance, that is, how completely the individual waste items are identified. This separation is what allows the minimum sufficient level of analysis to be identified for a given monitoring objective, and it is the element that distinguishes the present work from conventional algorithm benchmarking.

The dataset is used as a calibration scenario rather than a large training corpus. Controlled contamination conditions allow consistent comparison between methods while minimizing environmental variability.

MATERIALS AND METHODS

UAV Platform and Image Acquisition

The UAV system used in this study is a DJI Mini 2 (SZ DJI Technology Co., Ltd., Shenzhen, China), a lightweight quadrotor platform weighing 249 g, equipped with a 1/2.3" CMOS sensor capable of capturing 12 MP still images and 4K video.

Fig. 1 - DJI Mini 2 UAV platform used for image acquisition in this study

All imagery was acquired over the orchard garden of the Faculty of Biotechnical Systems Engineering, National University of Science and Technology Politehnica Bucharest. The drone was operated at a maximum altitude of 15 m above ground level (AGL), providing a spatial resolution of approximately 0.5 cm per pixel. Two reference images were captured: one depicting the undisturbed scene (baseline) and one after controlled placement of PET bottles and aluminium cans of varying sizes, colours, and shapes. This controlled setup enables direct verification of detection outputs across the three evaluated methods.

Colour-Based Image Representation. RGB Colour Model

The RGB colour space was used for initial statistical screening based on channel histogram comparison. Due to its sensitivity to illumination variation, RGB alone is not suitable for reliable outdoor segmentation but serves as a baseline reference for global scene changes (*Stan et al., 2025*).

The statistical screening criterion was defined as follows. For each of the three red–green–blue (RGB) channels the 256-bin histogram was normalised and compared between the baseline and the post-contamination image with the OpenCV correlation metric (cv2.HISTCMP_CORREL). A scene was flagged as changed when the correlation coefficient of at least one channel fell below 0.98. This threshold was derived from the data and not assumed: 15 repeated acquisitions of the undisturbed scene produced baseline-to-baseline channel correlations between 0.962 and 0.998 (mean 0.987), so 0.98 corresponds approximately to the lower quartile of the natural, illumination-driven variability of the site. The correlation criterion was preferred over histogram intersection or the chi-square distance because it is normalized, dimensionless and directly comparable between the three channels, which is the property required for a single global screening decision. These 15 undisturbed acquisitions are the same frames that form the uncontaminated part of the 40-frame evaluation set defined in the Experimental Setup subsection. The 0.98 correlation threshold was therefore calibrated on the evaluation imagery itself and not on an independent image set, and this is taken into account when the corresponding results are interpreted.

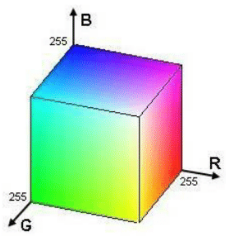

Fig. 2 - Conceptual illustration of RGB channel composition used as baseline reference for histogram comparison

HSV and HSL Colour Representations

The HSV and HSL colour representations were evaluated for colour-based segmentation of plastic waste. Both separate chromatic information from intensity, reducing sensitivity to illumination variations in outdoor scenes. In practice, HSV provided more stable threshold intervals for the tested objects, while HSL produced less consistent separation under shadow and reflective conditions. Consequently, HSV was selected for the segmentation stage.

The numerical hue–saturation–value (HSV) intervals applied in the segmentation stage were derived empirically from the acquired imagery rather than assumed. A total of 120 regions of interest containing polyethylene terephthalate (PET) bottles or aluminium cans and 120 regions of interest containing only vegetation, bare soil or shadow were sampled manually from the post-contamination frames. For the waste regions of interest, the 5^{th} - 95^{th} percentile range of the hue, saturation and value channels was computed (OpenCV convention, hue in the range 0-179) and rounded, which produced the two intervals used later for validation: dark blue (H 70-130, S 40-255, V 30-255) and cyan or light blue (H 50-100, S 40-255, V 100-255). The lower saturation bound of 40 removes desaturated soil and dry vegetation and the lower value bounds remove deep shadow pixels; these two classes were the dominant sources of false colour matches in the sampled background regions of interest. The 240 regions of interest were sampled from frames that also belong to the 40-frame evaluation set, so the hue–saturation–value intervals and the 5% coverage criterion were likewise calibrated on the evaluation imagery rather than on independent data. It should also be stated explicitly that these two intervals deliberately target only the blue and cyan chromatic cues that dominate the staged polyethylene terephthalate bottles: the validation stage was never intended to cover the complete range of colours and material classes placed in the scene, and it is applied as a colour-consistency filter on candidate regions rather than as a general waste classifier.

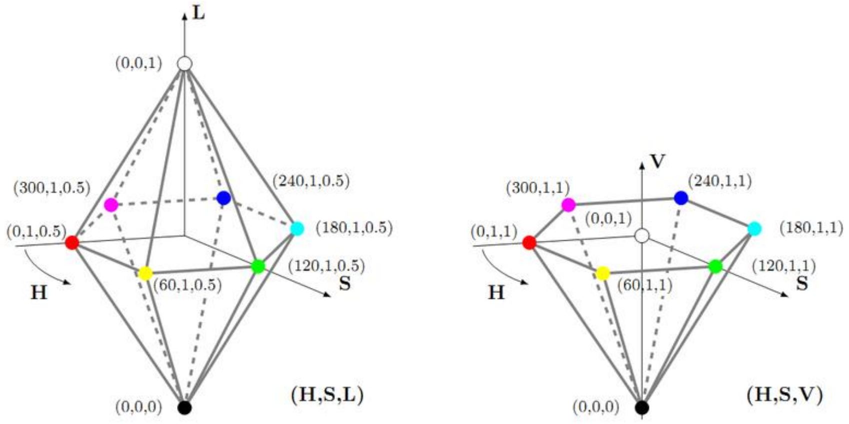

Fig. 3 - Comparison of HSL and HSV representations used during segmentation testing.
HSV showed more stable threshold behaviour under field conditions
(original figure, produced by the authors)

Neural Network Architecture: YOLOv8

A YOLOv8 single-stage object detection model was used to evaluate deep learning–based recognition performance. The model was retrained on a domain-specific dataset to detect PET bottles and aluminium cans under field conditions.

The pretrained YOLOv8 model (COCO dataset, 80 classes) does not include PET bottles or aluminium cans as dedicated categories. Retraining was therefore required to adapt the detector to the specific waste classes considered in this study. The YOLOv8s (small) variant was selected as a balance between detection performance and computational efficiency.

Dataset Preparation for YOLOv8 Retraining

A dedicated dataset was constructed consisting of 1130 images of PET bottles and aluminium cans placed in natural outdoor settings representative of field inspection conditions. Objects covered multiple brands, sizes, colours, and deformation states. Images were captured at varying distances (0.5–3 m), viewing angles, and illumination conditions to introduce intra-class variability while maintaining controlled environmental context. Annotations were created in YOLO format using Roboflow with two classes: *pet_bottle* and *aluminium_can*. Data augmentation included horizontal and vertical flips, HSV variation (±15%), and mosaic augmentation.

The dataset was split image-wise into 791 training images (70%), 226 validation images (20%) and 113 test images (10%), with no scene and no physical object appearing in more than one subset, and it was used to calibrate the detection workflow rather than to train a general-purpose detector. Diverse backgrounds, lighting conditions, and object appearances were intentionally included.

Fig. 4 - Selection of field-collected PET bottle and aluminium can images used for YOLOv8 retraining
(original images, acquired by the authors)

Retraining was performed for 100 epochs using the YOLOv8s (small) variant with a batch size of 16 and an initial learning rate of 0.01 reduced through cosine annealing. Training was executed on an NVIDIA T4 GPU (Google Colab). The model checkpoint with the highest validation mAP@0.5 was used for inference.

The validation subset was used exclusively for checkpoint selection (highest validation mean average precision at an intersection-over-union of 0.5, reached at epoch 87 of 100), and the 113-image test subset was processed only once, after the final checkpoint had been fixed, giving a mean average precision at an intersection-over-union of 0.5 of 0.83. The 40 unmanned aerial vehicle frames used for the operational evaluation reported in the Results section were fully independent from model training and from model selection: they were acquired in a separate flight session at 15 m above ground level in nadir orientation, whereas all 1130 dataset images were acquired at close range (0.5–3 m) on different dates and in different parts of the site. No evaluation frame, and no crop derived from an evaluation frame, was used at any stage of training, augmentation, validation or checkpoint selection.

The experimental site was selected to ensure controlled environmental conditions rather than geographic representativeness. The orchard contains mixed vegetation, variable illumination, and reflective surfaces typical of peri-urban dumping areas. Using a single controlled location isolates detection behaviour from geographic variability and is consistent with pilot validation procedures prior to multi-site deployment.

Experimental Setup

In environmental inspection scenarios, detection performance is evaluated in terms of operational reliability rather than pure computer vision accuracy. The objective is to determine whether a method can reliably trigger a site inspection when waste is present, not to optimize object-level classification metrics. Accuracy values are therefore interpreted as indicators of inspection reliability.

Site-level detection reliability was defined as the correct classification of a monitored area as contaminated or uncontaminated. A detection was considered successful if at least one waste object was correctly identified within the contaminated scene, triggering an inspection flag. To reduce random variability, each scenario was repeated with slight UAV repositioning (±0.5 m lateral shift). Each waste layout was captured 5 times and the undisturbed condition was captured 15 times, all within short temporal intervals and with identical flight parameters.

The evaluation set was identical for the three methods and consisted of 40 unmanned aerial vehicle frames acquired in the same session: 25 frames of the contaminated area (5 distinct spatial layouts of the waste objects, each captured 5 times) and 15 frames of the undisturbed area. Each method was therefore applied 40 times, which gives 40 site-level decisions per method and 120 site-level decisions in total.

For the change detection method, which requires a reference, each of the 25 contaminated frames was paired with a baseline frame from the same session, and the 15 undisturbed frames were processed as baseline-to-baseline pairs so that false alarms could be counted under identical processing conditions.

Site-level detection reliability R was computed as the proportion of correct site-level decisions over the complete evaluation set: R = (TPsite + TNsite) / Ntotal, with Ntotal = 40. Here TPsite is the number of contaminated frames (out of 25) in which at least one waste object was correctly localised, thereby raising the inspection flag, and TNsite is the number of undisturbed frames (out of 15) in which no inspection flag was raised. The denominator is therefore the total number of evaluated acquisitions and not the number of waste objects present in the scene; object counts are used only for the object-level metrics reported separately in the subsection Object-Level Detection Metrics.

Concerning the independence of the data used for threshold selection, the three methods were not evaluated under exactly the same conditions, and this difference should be clearly acknowledged. For the two classical approaches, the decision thresholds (the 0.98 histogram-correlation threshold, the two hue–saturation–value (HSV) ranges, and the 5% coverage criterion) were established using frames and regions of interest from the same 40-frame dataset used for evaluation. No separate calibration flight was carried out. Therefore, the reliability values reported below for the RGB histogram comparison and the OpenCV change-detection method with HSV filtering should be regarded as calibration-scenario results rather than as performance obtained on an independent test set. In contrast, the retrained YOLOv8 model was evaluated under fully independent conditions. Its training and validation data, as well as the data used for checkpoint selection, consisted of close-range images acquired during separate sessions, as described in the Dataset Preparation subsection. This difference in the evaluation procedure actually gives the classical methods an advantage. Consequently, the better performance of the deep learning approach reported in the Results section can be considered a conservative result rather than one caused by a more favourable evaluation setup. At the same time, this experimental design reflects a realistic operational situation, since threshold-based methods are commonly adjusted to the specific conditions of a site before deployment.

All experiments were conducted in the orchard garden of the Faculty of Biotechnical Systems Engineering. Following the baseline image acquisition, PET bottles (ranging from 0.5 L to 2 L, in transparent, blue, green, and branded varieties) and aluminium cans (330 mL, assorted brands) were distributed across a defined ground area. The UAV was repositioned using identical flight parameters (15 m AGL, nadir orientation) for post-contamination acquisition.

Fig. 5 - Baseline UAV image acquired at 15 m AGL over the experimental site (no waste present)
(original image, acquired by the authors)

Fig. 6 - Post-contamination UAV image showing PET bottles and aluminium cans distributed across the test area. Waste objects appear as reflective elements contrasting with the surrounding vegetation
(original image, acquired by the authors)

RESULTS

Method 1: RGB Histogram Channel Comparison

RGB channel histograms were computed using OpenCV (256 intensity bins per channel) and compared between baseline and contaminated scenes to assess global distribution shifts.

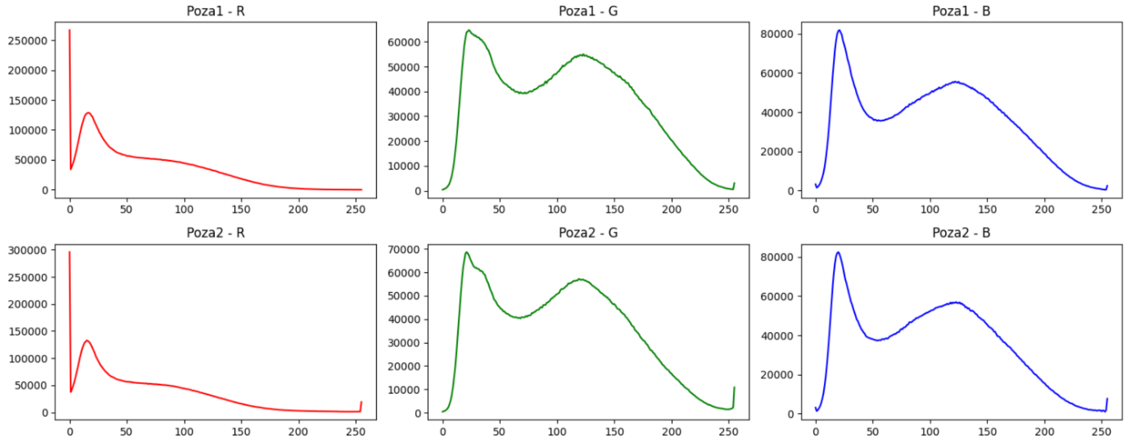

Fig. 7 - RGB channel histograms for the baseline image (Row 1) and contaminated image (Row 2). The distributions are nearly identical, indicating limited sensitivity to localised waste objects

(Original figure, produced by the authors)

The histogram pairs (Fig. 7) show minimal differences between the two scenes. This reflects the small spatial footprint of the waste objects relative to the full image and the reflective or semi-transparent nature of PET bottles, which blend with surrounding vegetation. As a result, global RGB statistics were insufficient for detecting localised contamination under these conditions and serve primarily as a baseline reference for comparison with more advanced methods.

Expressed in the terms defined in the Experimental Setup subsection, the correlation criterion flagged none of the 25 contaminated frames (0 of 25) and correctly classified 11 of the 15 undisturbed frames as unchanged, so the site-level reliability was (0 + 11) / 40 = 11/40 = 27.5%. The waste objects covered only 0.4-0.6% of the frame area, and the mean channel correlation between a baseline frame and its contaminated counterpart was 0.994, that is, higher than the correlation observed between two repeated captures of the undisturbed scene (0.962–0.998). This is the direct numerical reason why the global statistical approach is non-conclusive rather than merely imprecise: the signal produced by the waste is smaller than the natural variability of the scene. As stated in the Experimental Setup subsection, the 0.98 threshold was calibrated on these same frames, so this value describes calibration-scenario performance and not an independent test; a threshold transferred from another site would not be expected to perform better.

Method 2: OpenCV Change Detection with HSV Filtering

The second method combines temporal change detection (pixel-level image differencing) with spatial object filtering based on HSV colour masking, a classical approach used in environmental and surveillance monitoring (*Radke et al., 2005*). Similar threshold-based segmentation strategies have been applied in UAV imagery analysis for vegetation detection (*Torres-Sánchez et al., 2015*).

The processing steps were as follows:

Step 1 – Image alignment and differencing. The two images were resized to identical dimensions and the absolute pixel-wise difference was computed using cv2.absdiff(), producing a map highlighting changed regions.

Geometric alignment was obtained in two stages. First, both acquisitions were taken from the same programmed waypoint with the aircraft in hover-hold, at 15 m above ground level and with the gimbal oriented nadir, so that the residual horizontal displacement logged by the onboard global navigation satellite system and vision positioning remained within ±0.5 m; frames whose logged attitude deviated by more than 2° in pitch or roll from the baseline frame were discarded before processing. Second, the post-contamination frame was resized to the baseline dimensions and shifted by an integer pixel offset obtained by maximising the normalised cross-correlation of the two grayscale images over a ±30-pixel search window, computed only on the outer 15% of the frame, a border region that contains no waste objects. The residual misalignment after this correction was 1–2 pixels, that is, approximately 0.5–1 cm on the ground, which is smaller than the smallest object footprint considered. Only a translational correction was applied; sub-pixel and rotational residuals were not compensated, which is why feature-point registration is proposed as an improvement in the final part of this section. The ±0.5 m quoted above is the position-hold envelope specified for the platform and not the displacement actually observed; at the stated ground sampling distance it would correspond to about 100 pixels, whereas the correlation search window of ±30 pixels covers approximately ±15 cm. The two values are reconciled by the acceptance procedure applied to the frames: the residual translational offset measured between each post-contamination frame and its baseline, before correction, was 6–24 pixels (approximately 3–12 cm on the ground) with a mean of 13 pixels (approximately 6.5 cm), so every accepted pair lay inside the search window. Three acquisitions exceeded the ±30-pixel window and were rejected and re-flown before the evaluation set was assembled; consequently all 40 frames used in the evaluation satisfy this criterion, and the search window was sized on the measured offsets rather than on the worst-case positioning specification.

Step 2 – Noise reduction and binarisation. The difference map was converted to grayscale and smoothed with a Gaussian blur (kernel 5×5, σ=10) to reduce high-frequency noise and minor registration artifacts. A fixed threshold (T=60) generated a binary mask in which values above the threshold corresponded to significant changes.

Step 3 – Contour extraction and filtering. External contours were extracted using cv2.findContours(). Contours with bounding box areas smaller than 1000 pixels were discarded. At the tested resolution of approximately 0.5 cm per pixel this threshold corresponds to a ground footprint of about 250 cm² and was set to suppress the small change blobs produced by vegetation movement between the two acquisitions. As a direct consequence, an isolated 330 mL aluminium can, whose bounding box covers approximately 310 pixels, falls below the threshold and is retained only when it merges with an adjacent object into a larger change region; this is one of the identified causes of the moderate object-level recall reported below.

Step 4 – HSV colour validation. For each candidate bounding box, the corresponding Region of Interest (ROI) was converted to HSV. Two mask intervals targeted typical waste colours: dark blue (H:70-130, S:40-255, V:30-255) and cyan/light blue (H:50-100, S:40-255, V:100-255). If more than 5% of ROI pixels satisfied either mask, the object was classified as PET waste and marked with a bounding box. The 5% coverage value was determined from the same sampled regions of interest used to derive the HSV intervals: true waste regions showed between 12% and 48% of pixels inside the two masks, whereas vegetation, soil and shadow regions remained below 2%, so any value between 3% and 10% separates the two populations; 5% was retained as the midpoint of that interval. This validation stage is intentionally restricted to blue and cyan colour cues and is not intended to detect the complete range of waste colours and classes staged in the experiment. Candidate regions corresponding to transparent, green or metallic objects are consequently rejected here even when the differencing step localised them correctly, which is a further identified cause of the moderate object-level completeness discussed below.

Fig. 8 - Output of the OpenCV change detection pipeline. Overlapping bounding boxes are visible due to the absence of Non-Maximum Suppression (NMS). Detection reliability under controlled conditions was approximately 72% (29 of 40 evaluated acquisitions)
(original figure, produced by the authors)

The method localised the main waste cluster and most individual objects, performing substantially better than the histogram-based approach. However, several limitations were observed. Illumination changes between acquisitions (e.g., sun angle or cloud cover) affected chromatic consistency and reduced stability.

The absence of Non-Maximum Suppression produced overlapping bounding boxes in partially occluded areas. False positives occurred when non-waste blue objects were present. In addition, the approach requires a clean baseline image of the same scene, limiting its use to monitoring scenarios where prior reference imagery exists.

In numerical terms, the change detection pipeline raised a correct inspection flag in 21 of the 25 contaminated frames and correctly returned no flag in 8 of the 15 undisturbed frames, which gives a site-level reliability of (21 + 8) / 40 = 29/40 = 72.5%, reported throughout as approximately 72%. The 7 false alarms on undisturbed frames were produced by illumination changes on reflective leaf surfaces between the two acquisitions of a pair, and the 4 missed contaminated frames corresponded to layouts containing only small isolated cans, which fall below the contour area threshold defined in Step 3. This value is likewise a calibration-scenario result, since the hue–saturation–value intervals and the coverage criterion were derived from the same imagery.

Method 3: Custom-Retrained YOLOv8 Neural Network

YOLOv8 was retrained on the 1130-image PET and aluminium can dataset described in Section 2.4. Inference was performed using the Ultralytics Python API with a confidence threshold of 0.3 and an Intersection-over-Union threshold of 0.7 for duplicate suppression. Detected objects were annotated with bounding boxes and saved to folder runs/detect/predict/.

Under the tested conditions, the retrained YOLOv8 reached approximately 93% site detection reliability (23 of the 25 contaminated frames were correctly flagged and 14 of the 15 undisturbed frames correctly returned no flag, that is (23 + 14) / 40 = 37/40 = 92.5%), representing the highest reliability among the evaluated methods. The model remained stable in the presence of partial occlusion, varied object orientations, and reflective surfaces that reduced the performance of the colour-based methods.

PET bottles of different sizes and colours were correctly identified in the test scene. The built-in NMS mechanism within the YOLO inference pipeline prevented duplicate bounding boxes. A representative detection example under field conditions is presented in Fig. 9.

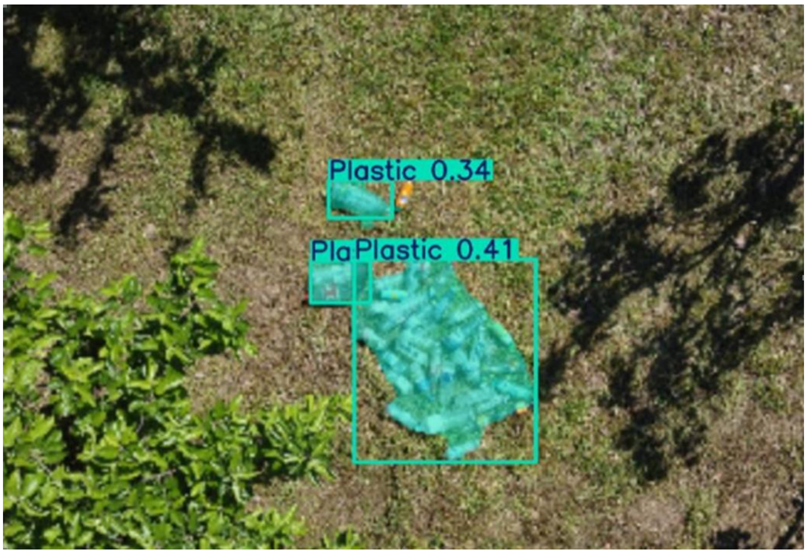

Fig. 9 - Site-level detection result obtained with the retrained YOLOv8 model. All but one waste object were identified, sufficient to trigger a positive inspection flag under the defined reliability criteria
(original figure, produced by the authors)

Unlike the change-detection approach, YOLOv8 does not require a reference image. Detection can therefore be performed on single acquisitions, which is relevant for surveys in areas without prior coverage. However, the approach requires dedicated computational resources for training, and its performance depends on the diversity and representativeness of the training dataset.

Object-Level Detection Metrics

Beyond site-level reliability, object-level detection was assessed using standard object detection metrics. The retrained YOLOv8 model reached a precision of approximately 0.90 and a recall of approximately 0.75, resulting in an F1-score of about 0.82. Mean average precision was approximately mAP@0.5≈0.80 and mAP@0.5:0.95≈0.52. These metrics were computed over the 25 contaminated evaluation frames, which contain 168 annotated waste objects in total (112 PET bottles and 56 aluminium cans).

The model returned 140 detections above the confidence threshold, of which 126 were true positives and 14 false positives, while 42 annotated objects were not detected; this gives a precision of 126/140 = 0.90, a recall of 126/168 = 0.75 and an F1-score of 0.82. Recall was lower than site-level reliability because small, partially occluded, or highly reflective objects were not always individually detected, even when the contaminated area was correctly classified. Such behaviour is consistent with monitoring tasks where confirming contamination presence is more relevant than detecting every individual item.

To evaluate robustness, inference was repeated under identical flight parameters using multiple image captures acquired within short temporal intervals. Variability in detection performance remained within ±3-5% for precision and recall (computed as the range of the precision and recall values across the five repeated captures of each of the five waste layouts), indicating stable behaviour under consistent environmental conditions. While the study does not address multi-site generalisation, the limited variability indicates stable performance under the tested conditions.

Comparative Performance Summary

Table 1 summarizes the key characteristics and operational monitoring characteristics of the three methods evaluated in this study. The site-level reliability values of the two classical methods are calibration-scenario values in the sense defined in the Experimental Setup subsection, whereas the YOLOv8 value was obtained with a model trained, validated and selected on independent imagery.

Table 1

Comparative summary of the three waste detection methods

Method	Core Technique	Site Detection Reliability	Baseline Required	Operational Processing Feasibility
RGB Histogram	Global colour distribution comparison	<30% (non-conclusive)	Yes	Yes
OpenCV Change Detection + HSV	Pixel differencing + HSV colour masking	~72%	Yes	Partially
YOLOv8 (retrained)	Custom deep learning detector	~93%	No	Yes

Note: the column Operational Processing Feasibility was rated with two criteria measured on the evaluation set: (a) the mean processing time per frame on the field configuration used in this study, a laptop with an Intel Core i7 processor, 16 GB of memory and no dedicated graphics card, and (b) the need for additional inputs or manual tuning during processing. The rating Yes means under 30 s per frame with no manual intervention (RGB histogram comparison, 0.8 s per frame; YOLOv8 inference, 2.1 s per frame on the processor and 0.35 s per frame on the NVIDIA T4 graphics card). The rating Partially means that the processing chain completes but requires a co-registered reference frame and manual re-tuning of the HSV intervals whenever object colours or illumination change (OpenCV change detection, 4.2 s per frame in addition to the alignment step). To ensure that the same definition is applied consistently to all three methods, the need for a reference image is not included in this rating. Instead, this requirement is reported separately in the *Baseline Required* column. The rating therefore considers only the requirements of the processing chain once the input images are available. Based on this definition, the RGB histogram comparison is rated *Yes*. Although it uses a reference frame, it compares global histograms and therefore does not require geometric co-registration or parameter adjustment during operation. The OpenCV change-detection method is rated *Partially*, not because it requires a reference image, but because the two images must be accurately aligned before comparison. In addition, the HSV intervals may need to be manually adjusted when object colours or illumination conditions change. YOLOv8 is rated *Yes* because it processes a single frame using fixed model parameters. Thus, the three ratings are based on the same criteria: processing time, co-registration effort, and the need for manual parameter tuning. This makes the ratings directly comparable across the three methods.

The contrast between object-level metrics and site-level reliability highlights the distinction between algorithm benchmarking and practical inspection needs. Missing individual items has limited impact on operational decisions if contamination presence is correctly identified. In this context, the combination of high precision and moderate recall is appropriate for screening scenarios in which UAV surveys are followed by targeted ground verification.

The subsections above report the direct experimental findings of the study. The subsections that follow (Analysis of Method Limitations, Toward Improved Waste Detection Systems, Environmental Monitoring Implications and Study Scope and Generalizability) contain the interpretation of those findings, the limitations of the experimental design and the directions proposed for future work, and are grouped here so that they remain clearly separated from the measured results.

Analysis of Method Limitations

The RGB histogram analysis method failed primarily due to the optical properties of the waste material. Transparent PET bottles at 15 m altitude act as highly reflective surfaces, reflecting the surrounding vegetation's spectral signature, thus making their contribution to the global colour histogram negligible. *Balsi et al. (2021)* reported similar challenges with transparent plastics in hyperspectral marine remote sensing, and the acquisition-dependent behaviour described by *Honkavaara et al. (2013)* suggests that global colour statistics are insufficient for localised waste mapping regardless of spectral resolution.

The OpenCV change detection pipeline overcomes the transparency problem by exploiting temporal changes (the appearance of new objects) rather than absolute colour. However, the method inherits the fundamental limitation of requiring a clean baseline image, which is unavailable in many operational contexts (e.g., post-event disaster assessments, areas with no prior UAV coverage). The fixed HSV threshold intervals are also susceptible to limitation when waste colours deviate from the assumed blue/cyan range – as would occur with red, yellow, or transparent bottles, or at different altitudes where atmospheric scattering modifies apparent hue. The absence of NMS is a practical limitation producing visually cluttered output; this could readily be addressed by applying cv2.dnn.NMSBoxes() to the detected bounding box list.

The retrained YOLOv8 model showed higher reliability than the classical methods but requires a domain-specific training corpus. Although the dataset contains 1130 training images, it remains limited in environmental diversity, as acquisitions were performed in a single site under controlled conditions. The model may struggle with out-of-distribution waste types (unusual colours, highly degraded bottles, dense overlapping clusters). Scale variation between training (ground-level photography) and inference (15 m AGL aerial) may also introduce domain shift, which was partially mitigated by including perspective-augmented crops in the dataset. Similar limitations have been widely reported for small object detection tasks, where limited pixel footprint and complex backgrounds reduce detection recall (*Ren et al., 2018; Kellenberger et al., 2018*). In operational monitoring, avoiding false alarms is often more critical than detecting every individual item. High precision reduces unnecessary field interventions, while moderate recall remains acceptable when aerial screening is followed by ground verification (*Gheorghe et al., 2024*).

The baseline-referenced approach is better suited to targeted surveillance contexts, such as periodically inspected protected areas, riverbanks, or infrastructure corridors. In these cases, reference imagery is typically available from previous inspection flights, making change-based detection a practical operational strategy.

While the experimental validation was conducted in a single controlled site, the environmental characteristics of the orchard – heterogeneous vegetation, variable illumination, reflective surfaces, and partial occlusion – represent common features of peri-urban dumping locations, which are known challenges in UAV-based remote sensing due to varying acquisition geometry and illumination conditions (*Honkavaara et al., 2013*). Detection performance mainly depends on ground sampling distance (GSD) and object size in pixels rather than on location (*Torresan et al., 2017*). At the tested altitude (15 m AGL, ~0.5 cm per pixel), standard PET bottles occupy sufficient pixel area to ensure detectability. Performance would likely decrease at higher altitudes or in dense vegetation. This relationship can be approximated from the ratio between object size and pixel size. The results should therefore be considered applicable to similar low-altitude UAV surveys.

Toward Improved Waste Detection Systems

Several improvements could further strengthen the proposed framework. Expanding the YOLOv8 training dataset beyond the current 1130 images to include aerial imagery acquired at multiple altitudes and additional waste classes (e.g., glass, cardboard, textiles) would improve generalisation. Instance segmentation models such as YOLOv8-seg could also enable pixel-level delineation and more precise estimation of contaminated surface area. The OpenCV pipeline could be refined through feature-point-based image registration (e.g., ORB with RANSAC homography) to reduce misalignment artifacts that occasionally generate spurious detections.

Multispectral or hyperspectral sensing may further enhance discrimination by exploiting the spectral contrast between plastics and natural materials, as previously reported by *Balsi et al. (2021)*.

From an operational perspective, deployment on embedded platforms such as NVIDIA Jetson devices could enable near real-time onboard inference during flight, reducing post-processing delays.

The proposed framework may support municipal dump site monitoring, GIS-based documentation of illegal dumping, emergency response assessments, and other environmental surveys using comparable UAV workflows.

Environmental Monitoring Implications

Environmental authorities rarely require exhaustive object inventories during initial inspections; instead, rapid identification of suspicious locations is typically sufficient to trigger field intervention, similar to zone-based environmental risk mapping approaches used in remote sensing (*Rozario et al., 2018*). The present results indicate that moderate-complexity approaches can detect contamination presence reliably, while deep learning mainly improves object-level completeness rather than site classification.

In practice, environmental surveillance programs may benefit from a tiered monitoring structure: wide-area screening using lower-cost detection followed by targeted high-accuracy analysis only where needed. This approach aligns with adaptive monitoring strategies discussed in environmental management literature (*Nichols and Williams, 2006*) and can reduce inspection effort and operational time. UAV imagery also provides georeferenced visual documentation that may support inspection reporting and enforcement procedures. Such tiered structures are consistent with established environmental decision-support practice (*Matthies et al., 2007*).

Study Scope and Generalizability

The dataset size reflects the methodological focus of the study rather than an attempt at statistical generalisation. While large datasets are necessary for training broadly transferable detectors, pilot environmental monitoring studies typically aim to evaluate workflow feasibility, acquisition constraints, and detection behaviour under realistic field conditions. The results should therefore be interpreted within the context of controlled validation rather than multi-site generalisation.

CONCLUSIONS

This study examined the feasibility of UAV-assisted waste monitoring workflows rather than benchmarking artificial intelligence performance. Inspection reliability was found to depend more on operational design – such as baseline availability, inspection objectives, and intervention criteria – than on algorithmic complexity alone. All results reported here were obtained under controlled, low-altitude field conditions (a single orchard site, 15 m above ground level, nadir acquisition, a ground sampling distance of approximately 0.5 cm per pixel and deliberately placed PET bottles and aluminium cans) and the conclusions are limited to that scope.

RGB histogram comparison was ineffective for detecting transparent PET waste against vegetation. OpenCV-based change detection with HSV filtering reached approximately 72% detection reliability under controlled conditions (29 of 40 evaluated acquisitions), a calibration-scenario value obtained with thresholds tuned on the same imagery, and proved useful when reference imagery was available, although sensitive to illumination variability and colour assumptions. The retrained YOLOv8 model achieved the highest site-level reliability (~93%; 37 of 40 evaluated acquisitions) and remained stable under occlusion, reflection, and colour variability without requiring a baseline image.

The main outcome of the study is an evaluated pilot monitoring framework linking analytical complexity to inspection reliability under the controlled conditions tested. The findings offer practical guidance for structuring UAV-based waste monitoring systems and may support decision-support workflows in which rapid aerial screening informs targeted field intervention.

Future developments may include expanding the dataset, exploring onboard inference, and integrating multispectral sensing. Overall, while higher analytical complexity improves object-level completeness, reliable contamination detection can already be achieved at moderate complexity levels under the conditions tested here, depending on operational objectives. Method selection should therefore follow monitoring needs rather than detection accuracy alone.

REFERENCES

[1] Baiocchi, V., Napoleoni, Q., Tesei, M., Servodio, G., Alicandro, M., Costantino, D. (2019). UAV for monitoring the settlement of a landfill. *European Journal of Remote Sensing*, 52(sup1), 41–52. https://doi.org/10.1080/22797254.2019.1683471

[2] Balsi, M., Moroni, M., Chiarabini, V., Tanda, G. (2021). High-resolution aerial detection of marine plastic litter by hyperspectral sensing. *Remote Sensing*, 13(8), 1557. https://doi.org/10.3390/rs13081557

[3] European Parliament and Council (2008). Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on Waste and Repealing Certain Directives. *Official Journal of the European Union*, L312, 3–30.

[4] Fallati, L., Polidori, A., Salvatore, C., Saponari, L., Savini, A., Galli, P. (2019). Anthropogenic Marine Debris assessment with Unmanned Aerial Vehicle imagery and deep learning: A case study along the beaches of the Republic of Maldives. *Science of the Total Environment*, 693, 133581. https://doi.org/10.1016/j.scitotenv.2019.133581

[5] Filkin, T., Sliusar, N., Huber-Humer, M., Ritzkowski, M., Korotaev, V. (2022). Estimation of Dump and Landfill Waste Volumes Using Unmanned Aerial Systems. *Waste Management*, 139, 301–308. https://doi.org/10.1016/j.wasman.2021.12.029

[6] Gheorghe, G.-V., Dumitru D.-N., Ciupercă, R., Mateescu M., Mantovani S.-A., Prisacariu E., Harabagiu A. (2024). Advancing Precision Agriculture with UAV'S: Innovations in Fertilization. *INMATEH - Agricultural Engineering*, 74(3). https://doi.org/10.35633/inmateh-74-89

[7] Gheres, M.-I., Mariasiu, F., Csato, A., Sechel, I.-C. (2025). Main Directions of Application of Artificial Intelligence in Agriculture: A Review. *INMATEH - Agricultural Engineering*, 77(3). https://doi.org/10.35633/inmateh-77-52

[8] Goddijn-Murphy, L., Williamson, B.J., McIlvenny, J., Corradi, P. (2022). Using a UAV Thermal Infrared Camera for Monitoring Floating Marine Plastic Litter. *Remote Sensing*, 14(13), 3179. https://doi.org/10.3390/rs14133179

[9] Gonçalves, G., Andriolo, U., Pinto, L., Bessa, F. (2020). Mapping Marine Litter Using UAS on a Beach-Dune System: a Multidisciplinary Approach, *Science of the Total Environment*, 706, 135742. https://doi.org/10.1016/j.scitotenv.2019.135742

[10] Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., Litkey, P., Mäkynen, J., Pesonen, L. (2013). Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. *Remote Sensing*, 5(10), 5006–5039. https://doi.org/10.3390/rs5105006

[11] 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. https://doi.org/10.3390/rs12091515

[12] Kellenberger, B., Marcos, D., Tuia, D. (2018). Detecting Mammals in UAV Imagery: Best Practices to Address a Substantially Imbalanced Dataset with Deep Learning. *Remote Sensing of Environment*, 216, 139–153. https://doi.org/10.1016/j.rse.2018.06.028

[13] Maharjan, N., Miyazaki, H., Pati, B.M., Dailey, M.N., Nakamura, T., Shrestha S., Yoshida, Y. (2022). Detection of River Plastic Using UAV Sensor Data and Deep Learning. *Remote Sensing*, 14(13), 3049. https://doi.org/10.3390/rs14133049

[14] Manfreda, S., McCabe, M.F., Miller, P.E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., Ben Dor, E., Helman, D., Estes, L., Ciraolo, G., et al. (2018). On the Use of Unmanned Aerial Systems for Environmental Monitoring. *Remote Sensing*, 10(4), 641. https://doi.org/10.3390/rs10040641

[15] Matthies, M., Giupponi, C., Ostendorf, B. (2007). Environmental decision support systems: Current issues, methods and tools. *Environmental Modelling & Software*, 22(2), 123–127. https://doi.org/10.1016/j.envsoft.2005.09.005

[16] Mei, A., Baiocchi, V., Mattei, S., Zampetti, E., Pai, H.-J., Tratzi, P., Ragazzo, A.V., Cuzzucoli, A., Mancuso, A., Bearzotti, A., Fontinovo, G., Grosso, M., Chu, C.-Y., Bianconi, D. (2023). Conceptualization of a Satellite, UAS and UGV Downscaling Approach for Abandoned Waste Detection and Waste to Energy Prospects. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-1/W1-2023, 287–293. https://doi.org/10.5194/isprs-archives-XLVIII-1-W1-2023-287-2023

[17] Moy, K., Neilson, B., Chung, A., Meadows, A., Castrence, M., Ambagis, S., Davidson, K. (2018). Mapping Coastal Marine Debris Using Aerial Imagery and Spatial Analysis. *Marine Pollution Bulletin*, 132, 52–59. https://doi.org/10.1016/j.marpolbul.2017.11.045

[18] Nichols, J.D., Williams, B.K. (2006). Monitoring for Conservation. *Trends in Ecology & Evolution*, 21(12), 668–673. https://doi.org/10.1016/j.tree.2006.08.007

[19] Pîrjol, P. (2018). Aspects Regarding the Use of Remote Sensing in Air Surveillance Missions. *Buletinul Universității Naționale de Apărare "Carol I"*, pp. 31–37.

[20] Radke, R.J., Andra, S., Al-Kofahi, O., Roysam, B. (2005). Image Change Detection Algorithms: A Systematic Survey. *IEEE Transactions on Image Processing*, 14(3), 294–307. https://doi.org/10.1109/TIP.2004.838698

[21] Ren, Y., Zhu, C., Xiao, S. (2018). Small Object Detection in Optical Remote Sensing Images via Modified Faster R-CNN. *Applied Sciences*, 8(5), 813. https://doi.org/10.3390/app8050813

[22] Rozario, P.F., Madurapperuma, B.D., Wang, Y. (2018). Remote Sensing Approach to Detect Burn Severity Risk Zones in Palo Verde National Park, Costa Rica. *Remote Sensing*, 10(9), 1427. https://doi.org/10.3390/rs10091427

[23] Sliusar, N., Filkin, T., Huber-Humer, M., Ritzkowski, M. (2022). Drone Technology in Municipal Solid Waste Management and Landfilling: A Comprehensive Review. *Waste Management*, 139, 1–16. https://doi.org/10.1016/j.wasman.2021.12.006

[24] Son, S.W., Kim, D.W., Sung, W.G., Yu, J.J. (2020). Integrating UAV and TLS Approaches for Environmental Management: A Case Study of a Waste Stockpile Area. *Remote Sensing*, 12(10), 1615. https://doi.org/10.3390/rs12101615

[25] Stan, A., Bănică, C.-K., Gândescu, C.-H., Simion, C., Cotorobai, D.-M., Simion, V., Muraru-Ionel, C. (2025). Design and Experiment of Multi-Purpose and Modular Sensor Robotic Platform for Real-Time Assessing and Monitoring of Cea Farms Using Point Cloud, LIDAR, RGB and Thermal Imaging. *INMATEH - Agricultural Engineering*, 77(3). https://doi.org/10.35633/inmateh-77-01

[26] Topouzelis, K., Papakonstantinou, A., Garaba, S.P. (2019). Detection of Floating Plastics from Satellite and Unmanned Aerial Systems. *International Journal of Applied Earth Observation and Geoinformation*, 79, 175–183. https://doi.org/10.1016/j.jag.2019.03.011

[27] Torresan, C., Berton, A., Carotenuto, F., Di Gennaro, S. F., Gioli, B., Matese, A., Wallace, L. (2017). Forestry Applications of UAVs in Europe: A Review. *International Journal of Remote Sensing*, 38(8–10), 2427–2447. https://doi.org/10.1080/01431161.2016.1252477

[28] Torres-Sánchez, J., López-Granados, F., Peña, J.M. (2015). An Automatic Object-based Method for Optimal Thresholding in UAV Images: Application for Vegetation Detection in Herbaceous Crops. *Computers and Electronics in Agriculture*, 114, 43–52. https://doi.org/10.1016/j.compag.2015.03.019

[29] United Nations (2015). Transforming Our World: The 2030 Agenda for Sustainable Development. United Nations General Assembly Resolution A/RES/70/1, New York.

[30] Wang, B., Xing, Y., Wang, N., Chen, C.L.P. (2024). Monitoring Waste from Uncrewed Aerial Vehicles and Satellite Imagery Using Deep Learning Techniques: A Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 20064–20079. https://doi.org/10.1109/JSTARS.2024.3488056

[31] Yang, Q., Liu, M., Zhang, Z., Yang, S., Ning, J., Han, W. (2019). Mapping Plastic Mulched Farmland for High Resolution Images of Unmanned Aerial Vehicle Using Deep Semantic Segmentation. *Remote Sensing*, 11(17), 2008. https://doi.org/10.3390/rs11172008

[32] Youme, O., Bayet, T., Dembele, J.M., Cambier, C. (2021). Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV. *Procedia Computer Science*, 184, 197–204. https://doi.org/10.1016/j.procs.2021.05.037