

Detection of Illegal Landfills on Satellite Imagery Using a Multi-agent Framework

Yehor Lytvynov¹, Viktoriia Hnatushenko^{2,3}, Volodymyr Hnatushenko², Christian Heipke³

¹ Dept. Information Technologies and Systems, Ukrainian State University of Science and Technologies, Dnipro, Ukraine –
ihor.litvinov@gmail.com, vvitagnat@gmail.com

² Dept. Information Technologies and Computer Engineering, Dnipro University of Technology, Dnipro, Ukraine –
hnatushenko.V.V@nmu.one

³ Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany –
heipke@ipi.uni-hannover.de

Keywords: landfill; remote sensing; agent architecture; aerospace imagery, machine learning.

Abstract

Illegal waste disposal sites pose significant ecological and public-health risks yet remain difficult to track with traditional field inspections. We propose a multi-agent detection framework that fuses textural, spectral, and contextual cues from medium-resolution satellite imagery for this work. Three specialized agents - Waste-Pile, Road, and Industry detectors - are implemented as YOLO (You Only Look Once) convolutional models that generate partial hypotheses, which are then hierarchically aggregated through rule weights learned from expert-labelled samples. The system provides an interpretable set of object relations, allowing regulators to trace how individual cues contribute to the final decision. The method was validated on an independent test area near Taromske (Dnipropetrovsk region, Ukraine) and corroborated by ground surveys. Joint aggregation raised the posterior probability of the primary target cluster from 0.27 (single-detector confidence) to 0.91, while maintaining robustness to label noise and heterogeneous sensor characteristics. Compared with conventional CNN baselines, the proposed approach delivers three key advantages: explicit explainability of outputs, transferability to 10 m spatial resolution without extensive retraining, and seamless integration of heterogeneous evidence sources. The proposed framework can serve as a cost-effective backbone for regional and national waste-monitoring systems. Future work will focus on near-real-time processing of Sentinel-2 time series, incorporation of hyperspectral and thermal methane indicators to assess remediation stages, and extension of the array of features to other anthropogenic disturbances such as open-pit mining and construction debris.

1. Introduction

Uncontrolled solid-waste dump sites remain one of the most pressing global environmental challenges. According to UNEP estimates (UNEP, 2024), more than two billion tons of municipal solid waste are generated annually, and this figure is steadily rising. Their negative impact includes air and water pollution, soil degradation, and increased greenhouse gas emissions. Efficient detection and monitoring of such sites over large territories is, therefore, a critical component of modern environmental management.

Earth-observation data offer unique opportunities for large-scale waste-site monitoring. However, classical deep-learning approaches based on convolutional neural networks (CNNs) typically require high-resolution imagery, which substantially increases the cost of analysis and restricts geographical coverage (Fraternali et al., 2024). Moreover, CNN models behave as a black box, making verification of results difficult and hindering evidence-based decision-making by regulatory authorities (Gevaert et al., 2022). Even when large training datasets are available, accuracy often decreases, while the lack of a formalized description of spatial and logical relationships among key landfill indicators undermines the robustness of purely pixel-level classifiers (Zhao et al., 2024).

Recent advances in object-detection architectures, particularly the YOLO family of models (Ultralytics, 2025), have demonstrated the ability of convolutional detectors to achieve high accuracy and fast inference across large-scale image collections. YOLO-based detectors have been successfully applied to a wide range of remote sensing tasks, including the identification of industrial facilities (Wan et al., 2024), agricultural fields, and transport infrastructure (Bazrafkan et al., 2025). Nonetheless, their performance remains vulnerable to

noise, resolution constraints, and domain transfer issues in heterogeneous satellite imagery. Most importantly, their decision-making process is not directly interpretable, which limits trust and applicability in environmental regulation.

To overcome these shortcomings, hybrid methods that combine data-driven learning with rule-based reasoning are increasingly explored. We represent each scene with an array of features that captures hierarchical and contextual cues – object categories, geometry, topology, proximity, and multi-scale statistics. In remote sensing, such structured feature arrays allow explicit encoding of spatial relationships between different object types and can directly inform inference or re-weight detector outputs (Zhu et al., 2022). However, the integration of feature-array-based reasoning with modern CNN detectors such as YOLO for landfill detection remains underexplored.

In this study, we introduce a multi-agent framework that couples YOLO-based convolutional detectors with a structured array-of-features representation. The system is designed to operate on medium-resolution satellite imagery, fuse heterogeneous textural, spectral, and contextual cues, exploit CNN architectures such as YOLO for local feature extraction, and produce transparent, human-readable feature arrays with explicit per-feature scores that environmental authorities can readily interpret and audit.

This approach advances the state of the art in landfill identification and monitoring and delivers a practical tool for evidence-based environmental governance.

2. Literature review

Between 2022 and 2025, a series of publications significantly advanced automated landfill detection using Earth-observation (EO) data. A comprehensive survey (Fraternali et al., 2024)

synthesized more than 120 papers and highlighted two dominant trends:

- a shift from manual visual inspection to deep-learning pipelines,
- a move from very-high-resolution (VHR) imagery to cost-efficient medium-resolution sensors such as Sentinel-2 and Landsat-8.

This dual trend reflects both technological maturity and the growing demand for scalable monitoring frameworks capable of covering large geographic regions at lower cost.

Among CNN-based models, CascadeDumpNet (Zhang et al., 2024) attracted considerable attention for combining AutoML search with a cascade of hybrid blocks, reaching an F1-score of 0.91 on 0.3 m GSD images. On a global scale, a landfill map derived from more than 1.2 million VHR scenes represented the first statistically significant demonstration of a correlation between waste distribution and socio-economic factors. Meanwhile, spectral-indicator studies have reported consistent ecological signals, with an average NDVI decrease of 20–25% in landfill surroundings. Training on Sentinel-2 NDVI stacks has been shown to markedly improve inter-regional transferability of models, confirming the role of vegetation indices as stable auxiliary indicators.

The importance of medium resolution imagery was underscored by the PML4DC@ICLR 2022 workshop (Rajkumar et al., 2022), where CNN architectures achieved a mean IoU of 0.33–0.40 in semantic landfill segmentation even with limited annotated datasets. A subsequent thematic review in Remote Sensing (2025) emphasized the persistent “small-object problem” (<50 pixels) and called for more explicit integration of contextual features such as proximity to roads or industrial facilities. Research has also diversified the range of landfill signatures under study, from thermal anomalies to methane plumes. For instance, PRISMethaNet automates CH₄ detection in PRISMA hyperspectral imagery, paving the way for integrated monitoring of toxic emissions.

A particularly dynamic line of work has been the adaptation of the YOLO family of object detectors to landfill detection (Ultralytics). Originally designed for real-time recognition in natural images, YOLO has undergone successive refinements that are directly relevant to EO. YOLOv8 introduced a decoupled head for classification and regression, improving detection of small objects typical of medium-resolution imagery. YOLOv9 integrated DPCnv layers and programmable gradient techniques, leading to faster convergence and higher robustness in heterogeneous conditions. YOLOv10 streamlined the backbone with lightweight efficiency blocks, enabling large-scale monitoring with lower computational cost. Most recently, YOLOv11 incorporated transformer-inspired attention mechanisms and dynamic anchor assignment, yielding notable improvements in recall for irregularly shaped landfills and better generalization across diverse geographic regions. In practical EO applications, YOLO-based pipelines have been employed not only for direct dumpsite identification but also for mapping industrial facilities (J. Gao, 2021), road networks (B. Cui, 2025), and other contextual signals that co-occur with waste. Comparative studies show that YOLO models trained on mixed-resolution datasets outperform generic CNN classifiers by 5–10 percentage points in F1-score when applied to Sentinel-2 and Landsat imagery. Nevertheless, YOLO detectors remain sensitive to spectral confusion with bare soil or demolition debris, and their decision boundaries are opaque to end users — factors that reinforce the need for methods that combine efficiency with interpretability.

Interpretability has thus become a central research theme. Gradient-based visualizations such as Grad-CAM and LIME provide intuitive hot-spot maps of predicted landfills but often

impose a 3–5% accuracy decrease. To mitigate the black-box effect while preserving performance, recent work has shifted toward explicit, human-readable representations — here, framed as an array of features, that is a smaller version of formal grammars and graph-based models (Zhu et al., 2006). In this paradigm, each candidate landfill (or image tile) is described by a structured feature array that enumerates heterogeneous cues — textural descriptors, spectral indices (e.g., NDVI drop), geometric/morphological attributes, and contextual distances to roads or industrial sites — each with a calibrated confidence score and provenance. The feature array makes dependencies and evidence contributions transparent without relying on production rules; stakeholders can audit which specific features (and sensors) drove the final decision.

Complementarily, deep-learning semantic segmentation on Sentinel-2 provides strong baselines for land cover mapping: on summer-2023 scenes from Ukraine, five U-Net variants achieved ~80–89.2% overall accuracy; residual-block U-Nets topped accuracy while a LeakyReLU variant offered faster inference, supporting operational updates of thematic maps (Hnatushenko and Honcharov, 2024).

Finally, applied projects illustrate the practical impact of these methodological advances. In 2025, a UNDP pilot in Guatemala combined an automated dumpsite classifier with participatory community mapping, substantially accelerating landfill inventory compared with conventional methods (UNDP, 2025).

In summary, despite substantial progress in CNN- and YOLO-based approaches, key challenges remain: namely interpretability, resilience to missing or noisy inputs, and consistent performance on medium-resolution EO data. The array-of-features fusion framework proposed in this study addresses these gaps by formally aggregating heterogeneous indicators into a transparent probabilistic model, thereby advancing the state of the art in scalable and accountable landfill monitoring.

3. Landfill detection method

Chapter 3 presents the proposed method for landfill detection based on medium resolution Earth observation data. The approach integrates YOLO-based convolutional agents with a structured array-of-features representation that couples low-level texture recognition with explicit contextual evidence. Section 3.1 outlines the multi-stage pipeline—confidence thresholding, intra- and inter-agent clustering, and consolidation of site-level “landfill ecosystems” by merging per-instance feature arrays into a unified record. Section 3.2 details the YOLO architecture used for detection and the resulting per-object feature vectors that populate the array. A lightweight probabilistic fusion then aggregates these features into a posterior score for each candidate site. The method delivers computational efficiency via real-time detectors and explainability via human-readable feature arrays with per-feature contributions and provenance. This hybrid design addresses the limitations of conventional CNN models and establishes a robust, transparent framework for large-scale environmental monitoring.

3.1 Overview

The proposed workflow comprises five principal stages:

1. Task dispatch and primary analysis - a main software module retrieves image-analysis tasks from the processing queue and distributes them to several independent analytical services responsible for, e.g., segmentation, classification or contextual verification. Each service processes the input scene and forwards its partial results to

an aggregation module that fuses the evidence and produces an initial landfill-detection conclusion.

- This set of indicators captures both the core of a landfill and its typical context, ensuring that the description remains compact yet informative.

Table 1 defines an array of features that can be used as model parameters

Terminal	Agent detects	Contains
WASTE_PILE	Characteristic texture/relief of a waste mound (chaotic clusters of bright and dark fragments)	Core element of the object
ROAD	Access road near the predicted core	Nearly all dumpsites are serviced by transport
INDUSTRY	Industrial facility, quarry, or warehouse nearby	Often a source or recipient of waste

Table 1 – Terminal agents, detection cues, and semantic roles in the landfill-ecosystem model.

Formally, the array is defined as:

$$T = \{ \text{WASTE_PILE, ROAD, INDUSTRY} \}$$

The key element of the core is the classical WASTE_PILE detector: a segmentation model isolates chaotic waste structures that sharply contrast with the surrounding surface. Once, such a pattern is detected, the system immediately searches for a ROAD, representing transport access. The presence of an access road significantly increases the likelihood that the object is indeed a dumpsite rather than, for example, a natural landslide. The additional presence of an INDUSTRY element - such as a quarry, plant, or storage facility—often indicates a stable flow of waste and the supporting infrastructure for its accumulation.

- Confidence thresholding - the outputs of individual agents are examined, and any detections whose confidence scores fall below a predefined threshold are discarded. The outputs of individual detection agents (e.g., WASTE_PILE, ROAD, INDUSTRY) are not used directly but are first subjected to confidence filtering. Each detector produces a probability score (*prob*) associated with a bounding box or segmentation mask. To reduce the number of false positives, only those detections whose confidence exceeds a predefined threshold (θ) are retained for further processing. The threshold values are empirically calibrated for each agent type: for instance, the WASTE_PILE detector requires a relatively low threshold to maximize sensitivity to small dumpsites, whereas ROAD and INDUSTRY detectors typically employ stricter thresholds to minimize spurious detections caused by linear soil features or non-industrial buildings.

This step ensures that the algorithm operates only on reliable evidence, thereby reducing the risk of false positives (e.g., treating random soil patches as waste piles). At the same time, the use of class-specific thresholds retains: each agent can be tuned to balance precision and recall according to its role in the landfill context. For example, missing a weak INDUSTRY signal may not critically affect detection, but missing a low-confidence WASTE_PILE would eliminate the landfill hypothesis.

- After confidence thresholding, the detectors of each agent may still produce multiple overlapping or closely spaced detections that correspond to the same physical object. To avoid redundant evidence and to create semantically meaningful units, we apply intra-agent clustering. In this step, spatially adjacent detections of the same type (e.g., several WASTE_PILE masks or consecutive ROAD segments) are merged into a unified cluster. Clustering is performed using spatial criteria: if the distance between the centroids of two detections falls below a threshold, or if their bounding boxes exhibit an intersection-over-union (IoU) larger than a predefined limit, they are assigned to the same cluster. This ensures that extended objects such as long access roads or industrial complexes are represented by a single symbol, rather than fragmented detections. The confidence score of a cluster is recalculated using a Bayesian summation rule for independent evidence (Wu et al, 2024).

$$P = 1 - \prod_{i=1}^N (1 - p_i)$$

This expression represents the probability that at least one of the detections corresponds to the target object. Unlike simple averaging, Bayesian summation reinforces the effect of multiple consistent detections, while reducing the influence of noise.

- The final stage of the pipeline integrates evidence across heterogeneous agents. While intra-agent clustering consolidates detections of the same type, inter-agent clustering groups detections of different types (e.g., WASTE_PILE, ROAD, INDUSTRY) that fall within a defined spatial proximity radius. The rationale is that a dumpsite rarely exists in isolation: it typically co-occurs with transport access and, in many cases, nearby industrial activity. When such elements are spatially co-located, they form a coherent higher-level structure interpretable as a landfill ecosystem. By aggregating evidence from multiple semantic dimensions (texture, infrastructure, context), the system elevates low-level detections to an interpretable "ecosystem-level" object. This makes the output suitable for environmental monitoring and decision-making. The result of inter-agent clustering is a consolidated detection: each identified LANDFILL ecosystem is represented as a single, contextually verified object. This hierarchical structure mirrors how human experts interpret remote sensing imagery by integrating texture, access, and industrial context into a coherent environmental footprint.

3.2 YOLO architecture

The detection agent used in this study is based on the You Only Look Once (YOLO) family of convolutional networks, which has become a de facto standard for real-time object detection in computer vision. Unlike two-stage detectors (e.g., Faster R-CNN) that first generate region proposals and then classify them, YOLO adopts a single-stage architecture that jointly performs bounding-box regression and class prediction in one forward pass. This design yields substantially faster inference and makes YOLO particularly suitable for large-scale satellite image processing, where millions of patches must be analyzed efficiently. The architecture follows the canonical pipeline of backbone → neck → head, and is illustrated in Fig. 1.

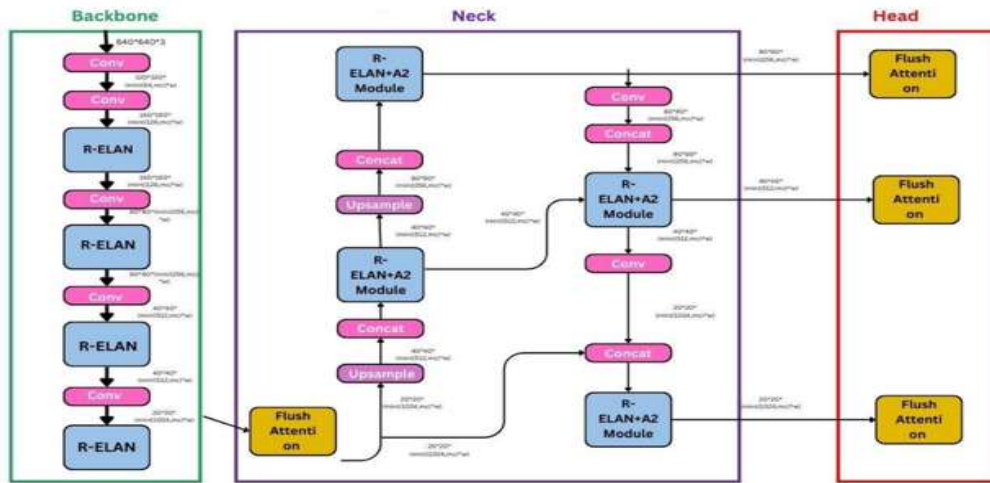


Figure 1. Architecture of major components of YOLOv12. (Chitari et al., 2025)

The backbone (variants of CSPDarknet, ConvNeXt or similar CNN feature extractors) encodes hierarchical image features. The neck (e.g., Feature Pyramid Networks or Path Aggregation Networks) fuses multi-scale representations, ensuring that both large structures (industrial zones) and small objects (access roads) are captured. Finally, the detection head outputs class probabilities and bounding-box coordinates. Recent versions (YOLOv8–v12) have introduced several improvements directly relevant to medium-resolution EO imagery: decoupled heads for classification and localization, deformable convolutions, lightweight efficiency modules, and transformer-inspired attention blocks. These enhancements collectively improve sensitivity to small, irregularly shaped landfills while maintaining high processing throughput.

In our framework, YOLO agents are specialized per feature channel in the array-of-features representation (i.e., WASTE_PILE, ROAD, INDUSTRY). Each detector is trained on task-specific datasets and tuned with class-dependent confidence thresholds so its outputs balance precision and recall according to operational priorities. Detectors emit structured feature entries—bounding geometry, confidence, scale, relevant spectral cues, and contextual distances—that are appended to a per-tile feature array. A lightweight fusion module then aggregates these arrays into a final posterior for each candidate site using calibrated weights and uncertainty estimates. In this way, the system couples YOLO’s efficiency in local extraction with transparent, human-readable feature arrays (rather than grammars), providing both scalable processing and auditable reasoning for reliable landfill monitoring across large territories.

3.3 Solution architecture

The general architecture of the landfill detection system is shown in Fig. 2. Aerial or satellite images are archived, divided into tiles, and queued for processing. Each tile is passed to a single, multi-task detector trained to recognize all relevant signals—waste piles, access roads, industrial buildings, vegetation loss, and other contextual cues—in one pass. The detector outputs a structured array of features with per-feature confidences. Because data quality and availability vary across regions, the system maintains robustness by decoupling detection from decision-making. Two lightweight agents operate after the universal detector:

- the confidence agent that computes the posterior likelihood of a landfill for each hypothesis and applies a data-driven confidence threshold to issue the final verdict.

This design retains the benefits of contextual reasoning while avoiding the cost and brittleness of training separate detectors for each feature class.

The resulting decision, along with the underlying feature array and intermediate scores, is written to the results repository for visualization, statistical analysis, and managerial action. Representing a landfill as an ensemble of spatially and functionally organized features provides a transparent basis for downstream tasks—classification by landfill type or status, risk scoring, and temporal monitoring—while improving generalization on lower-quality or incomplete imagery.

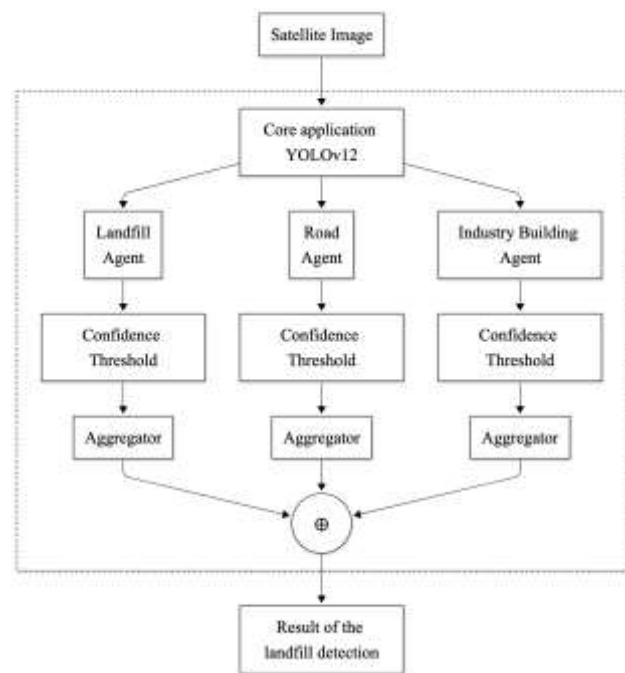


Figure 2. The architecture of landfills detection system

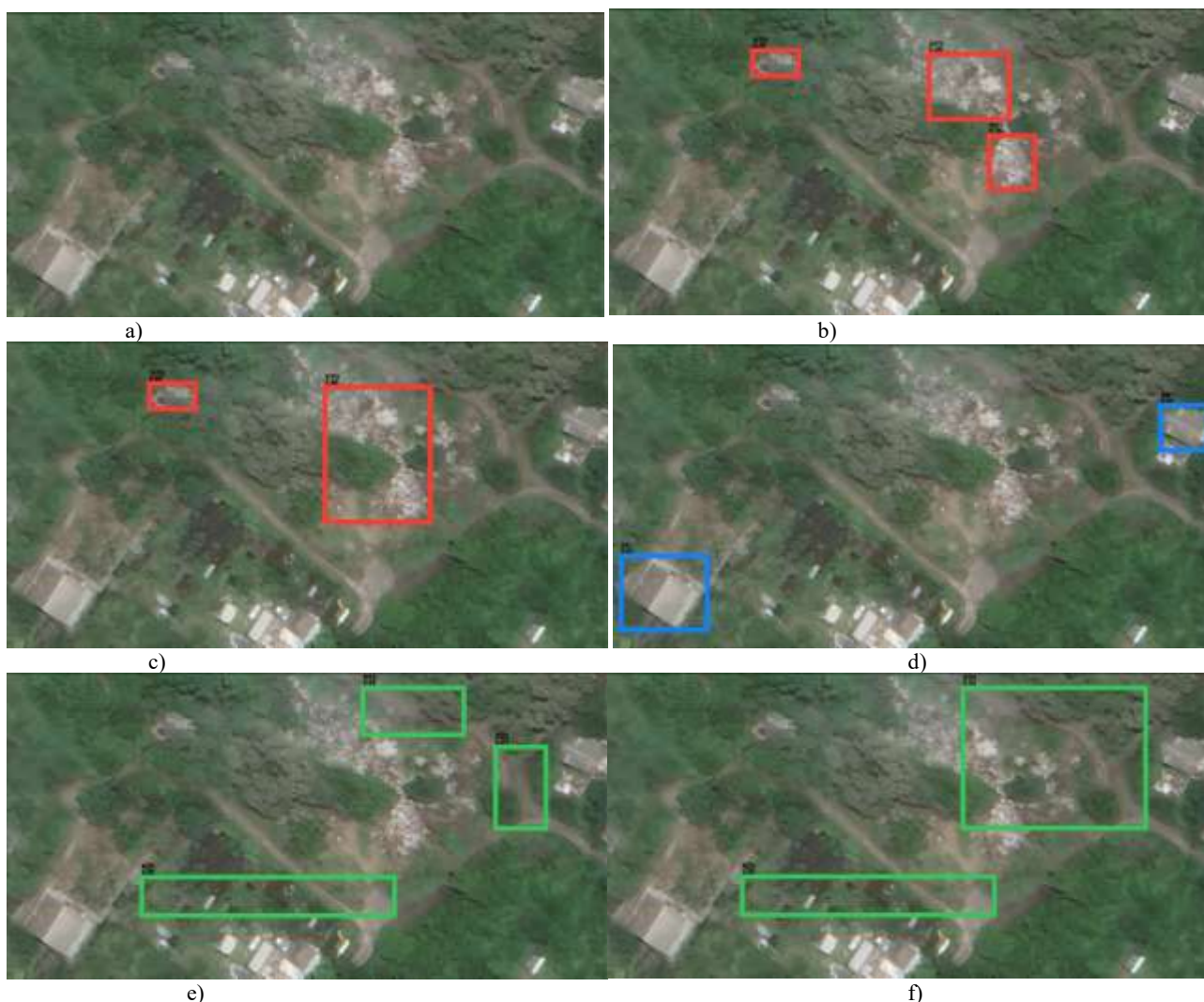


Figure 3. Result for the Taromske (Dnipropetrovsk region, Ukraine). a) Input image b) agent landfill detection, c) agent landfill detection after probability filter and aggregation steps d) agent industrial buildings detection, e) agent road detection, f) agent road detection after probability filter and aggregation steps

Detector	Training set	No. of annotated images (Train/ Valid/ Test)
Landfill	Aerial-Dumping-Sites (Object Detection of Illegal Dumping Sites,2023)	1385 / 63 / 63
Industrial buildings	roboflow-universe-projects /buildings-instance-segmentation (Roboflow Universe Projects, 2023)	6344 / 891 / 450
Roads	Turf100 Dataset (Andersen, 2025)	144 / 24 / 5

Table 2. Configuration of YOLO-based detection agents: training-set sizes and confidence thresholds used in the multi-agent framework.

4. Experiments

4.1 Statistical results

To implement the objectives of our research, we used satellite images obtained from the Aerial-Dumping-Sites (Object Detection of Illegal Dumping Sites,2023), roboflow-universe-projects (Roboflow Universe Projects, 2023), Turf100 Dataset

(Andersen, 2025) (Table 2). This date was chosen due to the need to analyze the condition of the terrain for the presence of landfills, roads and industrial buildings.

Table 3 reports per-feature detection results, listing the number of training and test images alongside Average Precision (AP). For Landfill, the model was trained on 1385 images and tested on 63, achieving 54% AP. Industrial buildings used 6344 training and 450 test images, reaching 74% AP. Roads was trained on 144 images and tested on 5, got 96% AP.

Testing dataset	Test	Train	Average Precision
Landfill	63	1385	54%
Industrial buildings	450	6344	74%
Roads	5	144	96%

Table 3. Average precision evaluation by class

4.2 A case study

For independent validation, we selected a satellite scene covering the village of Taromske, Dnipropetrovsk Oblast, Ukraine, acquired from SENTINEL-2 in summer 2024 (48.45504°N, 34.79398°E, altitude 626 m, 2192 x 1115 pixels) (Fig. 3a). This scene was deliberately excluded from all training datasets to provide an unbiased evaluation of the proposed framework.

To test the multi-agent approach, we trained a YOLO12-based network corresponding to the terminals **WASTE_PILE**, **ROAD**, and **INDUSTRY** using three different datasets from open sources. A summary of their configurations is presented in Table 2. The test scene was then processed independently by each agent. Figure 2b illustrates all candidate landfills detected by the **WASTE_PILE** agent, highlighted with red bounding boxes, while Table 4 lists the corresponding identifiers and confidence scores.

For completeness, inference was executed with a uniform software stack and fixed configuration across agents, with model weights frozen prior to validation and without any ad-hoc threshold tuning on the test scene. Standard non-maximum suppression was applied within each agent to consolidate overlapping boxes, and all predictions were re-projected to the native image grid for comparison. No geographic priors or hand-crafted masks were used. These precautions, while simple, help isolate the contribution of the learned representations and provide a conservative estimate of the framework’s effectiveness under routine deployment settings.

Candidate Id	Confidence
WP1	0,36
WP2	0,29
WP3	0,28
WP4	0,01

Table 4. List of landfill candidates detected in the scene together with their confidence scores.

The first post-processing stage applies a probability filter that removes spurious hypotheses by discarding any detection whose confidence score falls below a predefined threshold (0.25). This cutoff was selected during development and kept fixed for all experiments to avoid test-time tuning and to ensure comparability across scenes. As shown in Table 3, only candidate WP4 lies beneath this threshold; consequently, it is excluded from further analysis and does not enter the subsequent fusion steps. The remaining candidates constitute the working set for downstream processing, including spatial clustering and cross-agent consistency checks designed to consolidate evidence. The next step is intra-agent clustering. Objects WP1 and WP3 lie less than one pixel apart and were therefore merged into a single candidate, WP5. The aggregated confidence score was computed using Bayesian summation:

$$P = 1 - (1 - 0.36)(1 - 0.28) = 0.5392$$

The results after filtering and aggregation are shown in Table 5 and Fig. 3c.

Candidate Id	Confidence
WP2	0,29
WP5	0,54

Table 5. List of landfill candidates detected in the scene together with their confidence scores after filter and aggregation steps.

The **INDUSTRY** agent detected two industrial structures in the vicinity (Fig. 3d). Their identifiers and confidence scores are listed in Table 6.

Element Id	Confidence
IB1	0,81
IB2	0,80

Table 6. List of industrial buildings detected in the scene.

All detections exceeded the confidence threshold, passed the probability filter, and were spatially distant from each other; hence, no further aggregation was applied.

The **ROAD** agent identified three linear features associated with transport access (Fig. 3e). Their confidence scores are presented in Table 7.

Element Id	Confidence
RD1	0,91
RD2	0,88
RD3	0,81

Table 7. List of roads detected in the scene.

All elements exceeded the threshold. Since RD2 and RD3 were located less than one pixel apart, they were merged into a single cluster RD4. The Bayesian aggregation yielded a combined confidence of 0.92. The results are summarized in Table 8 and illustrated in Fig. 3f.

Road Id	Confidence
RD1	0,91
RD4	0,92

Table 8. List of roads detected in the scene after filter and aggregation steps

To estimate the overall likelihood of each landfill candidate, we fused per-agent confidence scores with contextual indicators that reflect the typical structure of a “landfill ecosystem.” We used two simple yet informative contextual rules: proximity to a road and proximity to an industrial facility. Each cue was computed by testing whether the candidate polygon intersected fixed-distance buffers around mapped roads and industrial footprints. We treated the cues as independent pieces of evidence and assigned each an additional weight of 0.40. This design accords with field practice: most illegal dumps have an access track and sit near production sites (Geng et al., 2024) that can act as sources or recipients of waste. Evidence was then combined via Bayesian aggregation. In effect, this reinforces multiple consistent detections while diluting isolated noise: a single weak hit without road/industry context seldom produces a high posterior.

Result = Candidate Confidence \oplus Rule Confidence
 The results are given in Table 9 and visualized in Fig. 3.

Id	Confidence	Road	Industry	Rule Confidence	Total Confidence
WP2	0,29	Yes	Yes	0.8	0.9076
WP5	0,54	Yes	Yes	0.8	0.8588

Table 9. Bayesian-aggregated confidence for landfill candidates, combining detector baseline with contextual cues (roads, industrial sites).

The result calculation is shown in Figure 4.



Figure 4. Result image with found landfills

A field visit was carried out to verify the highest-ranked site (WP2). The location was confirmed to be an active landfill (Fig. 5), providing empirical evidence for the validity of the proposed detection framework and demonstrating the practical utility of combining YOLO-based detectors with an array of features representation that fuses contextual evidence.



Figure 5. Photos of the detected landfill from different angles (48.45552°E, 34.79359°N)

5. Conclusion

The multi-agent framework that fuses textural, spectral, and contextual cues proved highly effective for detecting illegal landfills in medium-resolution satellite imagery (Sentinel-2). Validation on an independent test site near Taromske, Dnipropetrovsk Oblast, followed by ground truthing, confirmed localization accuracy: joint aggregation of partial results from three specialized detectors - landfill, road, and industrial-building models - increased the posterior probability of the principal cluster from 0.27 to 0.91. A structured array-of-features representation of the "landfill ecosystem" with per-feature scores and provenance, makes the decision process transparent and enables regulators to trace exactly which evidence combination led to the final verdict.

By providing explainability, transferability to 10 m resolution, and seamless integration of heterogeneous evidence, the proposed algorithm fills gaps left by conventional CNN-based solutions and can serve as a backbone for rapid waste-monitoring systems at regional and national scales. The proposed framework is cost-effective because it operates on freely available medium-resolution satellite imagery and avoids dependence on expensive very-high-resolution data. Its multi-agent architecture enables scalable, real-time processing over large territories while maintaining high detection accuracy through probabilistic evidence fusion.

Future work should focus on

- near-real-time processing of Sentinel-2 time series with automated cloud filtering;
- incorporation of additional indicators—hyperspectral data and thermal methane anomalies—to assess reclamation stages and environmental risks;
- extending the feature-array schema to other anthropogenic disturbances such as open-pit mining and construction debris, with automatic adaptation of feature weights to regional and seasonal conditions.

These developments will enhance the timeliness of environmental monitoring, reduce field-survey costs, and provide a transparent decision-support tool for waste-management authorities.

Beyond these immediate findings, several system-level considerations reinforce the practical value of the approach. In operational deployments, agencies often face heterogeneous acquisition conditions, evolving annotation protocols, and strong class imbalance. The array-of-features representation helps handle this variability by encoding stable structural signals (e.g., road proximity and industrial footprints as contextual anchors), while posterior aggregation reduces the influence of any single noisy cue. Routine ablation and sensitivity studies, leave-one-agent-out tests, threshold perturbations, and controlled degradations of spatial resolution can be layered on top of the pipeline to map regions of reliability and to calibrate decision policies for conservative versus aggressive screening. Coupled with uncertainty calibration (temperature scaling or isotonic regression) and explicit open-set handling for out-of-distribution textures, this yields a robust, auditable detector suitable for integration into existing environmental workflows.

References

- Andersen, J., n.d.: *Turf100 [Dataset]*. Roboflow Universe. <https://universe.roboflow.com/jesper-andersen-za3iu/turf100> (15 Sept 2025).
- Bazrafkan, A., et al., 2025: *Automated Detection of Center-Pivot Irrigation Systems Using YOLOv11 with Landsat-8, Sentinel-2,*

- and NAIP. *Remote Sens.*, 17(13), 2276. <https://www.mdpi.com/2072-4292/17/13/2276>
- Cui, B., Liu, Z., and Yang, Q., 2025. UAV-YOLO12: A Multi-Scale Road Segmentation Model for UAV Remote Sensing Imagery, *Drones*, vol. 9, no. 8, p. 533. <https://doi.org/10.3390/drones9080533>.
- Fraternali, P., Morandini, L., Herrera González, S.L., 2024: *Solid Waste Detection, Monitoring and Mapping in Remote Sensing Images: A Survey*. *Waste Manag.*, 189, 88–102. <https://doi.org/10.1016/j.wasman.2024.08.003>
- Gao, J., Chen, Y., Wei, Y., and Li, J., 2021. Detection of Specific Building in Remote Sensing Images Using a Novel YOLO-S-CIOU Model: Case—Gas Station Identification., *Sensors*, vol. 21, no. 4, p. 1375. <https://doi.org/10.3390/s21041375>.
- Geng, J., Ding, Y., Xie, W., Fang, W., Liu, M., Ma, Z., Yang, J., Bi, J., 2024: *An Ensemble Machine Learning Model to Uncover Potential Sites of Hazardous Waste Illegal Dumping Based on Limited Supervision Experience*. *Fundam. Res.*, 4(4), 972–978. <https://doi.org/10.1016/j.fmre.2023.06.010>
- Gevaert, C.M., et al., 2022: *Explainable AI for Earth Observation: A Review Including Societal and Regulatory Perspectives*. *Int. J. Appl. Earth Obs. Geoinf.*, 112, 102869. <https://doi.org/10.1016/j.jag.2022.102869>
- Hnatushenko, Vik., Honcharov, O., 2024: *Land Cover Mapping with Sentinel-2 Imagery Using Deep Learning Semantic Segmentation Models*. Proc. X Int. Sci. Conf. "Information Technology and Implementation" (IT&I 2024), Kyiv, Ukraine, 20–21 Nov 2024, 1–18. CEUR-WS.org/Vol-3909.
- Object Detection of Illegal Dumping Sites, 2023: *Aerial-Dumping-Sites: Dataset and Pre-Trained Model [Dataset]*. Roboflow Universe. <https://universe.roboflow.com/object-detection-of-illegal-dumping-sites/aerial-dumping-sites> (15 Sept 2025).
- Rajkumar, A., Sziranyi, T., Majdik, A., 2022: *Detecting Landfills Using Multi-Spectral Satellite Images and Deep Learning Methods*. PML4DC Workshop @ ICLR 2022. https://pml4dc.github.io/iclr2022/pdf/PML4DC_ICLR2022_9.pdf
- Roboflow Universe Projects, 2023: *Buildings Instance Segmentation [Dataset and Model]*. Roboflow Universe. <https://universe.roboflow.com/roboflow-universe-projects/buildings-instance-segmentation> (15 Sept 2025).
- Ultralytics, 2025: *Models Supported by Ultralytics [Documentation]*. Ultralytics YOLO Docs. <https://docs.ultralytics.com/models/> (12 May 2026).
- UNDP. *Illegal Dumpsite Detection Guatemala: Technical Paper (Contributing Paper)*. 2025. Available at: https://www.undp.org/sites/g/files/zskgke326/files/2025-04/guatemala_design_april_vf.pdf
- United Nations Environment Programme (UNEP), 2024: *Global Waste Management Outlook 2024: Beyond an Age of Waste – Turning Rubbish into a Resource*. UNEP, Nairobi. <https://www.unep.org/resources/global-waste-management-outlook-2024>
- Wan, Y., Zhan, Z., Ren, P., Fan, L., Liu, Y., Li, L., Dai, Y., 2024: *Storage Tank Target Detection for Large-Scale Remote Sensing Images Based on YOLOv7-OT*. *Remote Sens.*, 16(23), 4510. <https://doi.org/10.3390/rs16234510>
- Wu, L., Jiang, X., Zhu, W., Huang, Y., Liu, K., 2024: *Bayesian Decision-Based Fusion Algorithm for Remote Sensing Images*. *Sci. Rep.*, 14, 11558. <https://doi.org/10.1038/s41598-024-60394-y>
- Zhang, S., Ma, Y., 2024: *CascadeDumpNet: Enhancing Open Dumpsite Detection Using AutoML and Hybrid Blocks*. *Remote Sens. Environ.*, Article 113856. <https://doi.org/10.1016/j.rse.2024.113856>
- Zhao, Y., Sun, H., Wang, S., 2024: *Small Object Detection in Medium-Low-Resolution Remote Sensing Images Based on Degradation Reconstruction*. *Remote Sens.*, 16(14), 2645. <https://doi.org/10.3390/rs16142645>
- Zhu, Q., Lei, Y., Sun, X., Guan, Q., Zhong, Y., Zhang, L., Li, D., 2022: *Knowledge-Guided Land Pattern Depiction for Urban Land-Use Mapping: A Case Study of Chinese Cities*. *Remote Sens. Environ.*, 272, 112916. <https://doi.org/10.1016/j.rse.2022.112916>
- Zhu, S.-C., Mumford, D., 2006: *A Stochastic Grammar of Images*. *Found. Trends Comput. Graph. Vis.*, 2(4), 259–362. <https://doi.org/10.1561/06000000018>