Check for updates

AI-Driven Detection of Waste Hotspots in Surface-Water Bodies: A Comprehensive Review

Nishchay Malhotra[1] , *Sheetal S. Patil [2] , Amit Kumar[3] , Vishal Menia [4], anav Dogra[5] , Nishchal Singh Bishnoi [6]
[1]Central Water and Power Research Station, Pune, Maharashtra, India
[2,3,4,5,6] Department of Computer Engineering BVDUCOE, Pune, Maharashtra, India
malhotranishchay@gmail.com, *sspatil@bvucoep.edu.in, akumar122-comp@bvucoep.edu.in , vmenia22-comp@bvucoep.edu.in , kdogra22-comp@bvucoep.edu.in , nsbishnoi22-comp@bvucoep.edu.in

Abstract: Any form of floating debris in our waterways within rivers, lakes and reservoirs is an increasing environmental concern that has vast ecological and social economic consequences. We have witnessed recent advances in remote sensing and artificial intelligence that offers us the opportunity of automatic detection and mapping of floating waste hence we can accomplish large scale near real time water quality assessment. In the review, we consider articles that were conducted between 2019 and 2025 that employed optical satellite data, UAV data recorded by RGB sensors, hyperspectral sensors, SAR data and deep learning systems to the problem of debris detection and hot spot delineation.In this case, a structured selection process was employed, which took into account sensor selection, algorithmic strategies as well as reported results of the evaluation. According to the literature, high-resolution images of the UAV and multispectral satellite missions such as Sentinel-2 are the most suitable to combine time frequency and spatial characteristics. The deep-learning approaches (CNN classifiers, YOLO series detectors, and transformer systems), tend to perform highly, compared to the classical machine-learning and spectral-index approaches, in the detection of small and fine-scale debris. Even though SAR and hyperspectral data respond well in the clear atmosphere, as well as in turbidity and low light conditions, they have not been used in the study of inland waters. The metrics of precision and recall and F1 and the use of Streamlit were chosen in the selection of the set of data, augmentation, and training to CNN/transfer learning and transformer, basing our experimental pipeline on previous studies. Finally, the literature review makes recommendations of how the operational and scalable systems can be developed to identify the hotspots of debris.
Keywords: Floating litter, riverine plastics, surface-water monitoring, Sentinel-2, SAR, UAV imagery, CNN, YOLO, transformer models, hotspot mapping.

1. Introduction

Anthropogenic rubbish such as plastics are becoming a common sight floating in rivers, lakes, and reservoirs and the trend has brought about considerable concerns as to the implications of such a scenario on the quality of water and welfare of freshwater environments. Nor is it the waters of these waters inland that stand, but flow. Much of the waste that enters the drainage systems in the cities are carried by such waterways and end up in the seas and add up to the overall total quantity of plastic in the world oceans[1], [2].

Normal techniques of inspections of debris, through boat excursions, foot patrol or by hand picking samples are not highly satisfactory. They are both time consuming and man intensive and they barely keep up with the dynamic nature of floating waste [3],[2]. These methods are not necessarily sensitive to the sharp accumulations or changes as a result of wind, currents, or storm water deposits. Among the most used techniques of satellite surveillance are satellites missions such as sentinel-2, which captures images in multispectrals, free to view and the satellites are capable of surveying a large area already and have been proven to observe floating plastics in various light conditions [1], [3]. Applications can also be made to better Hyperspectral sensors

M. Kaushik et al. (eds.), *Proceedings of the International Conference on Responsible, Risk-aware, and Regulated AI (RRRAI 2026)*, Advances in Intelligent Systems Research 210,
https://doi.org/10.2991/978-94-6239-723-1_20

like PRISMA which can penetrate further due to their small spectral bands that can be used to differentiate plastic signature against the surrounding vegetation or water surfaces [2], [4]. The most notable is radar systems- Sentinel-1 SAR, which will offer the added value in the event of optical sensors, which in turn will not be operable, that is, when it is cloudy, turbid water, or at night [5]. UAVs also happen to be the most often favorable when exceedingly minute details are needed because they have the capability of getting such an in-depth picture that they have the ability to detect extremely tiny pieces of debris [5], [6].

It has experienced advancement in the aspect of computation in the same way. These deep-learning networks, like CNN models, YOLO-like detectors, Mask R-CNN and transformer based networks, have now generally outperformed the older machine-learning algorithms and classical spectral-index machines that were previously used to detect the debris [7], [3], [8].

Nevertheless, such tools have a variety of unresolved issues that prevent them to operate in a fully functional mode. The models of coastal or open seas have not been demonstrated to be a significant amount of the transfer to the waters of inland areas that are typified by immense disparities in colour, turbidity, and backgrounds among locations [7], [9]. Objects or debris that is slight below the surface, or that becomes mixed up with natural material, or travel fast in currents, may be terribly difficult to identify by algorithms [3], [6]. Marked datasets, which signify different seasons, water levels, weather conditions and other environmental variation, are not yet quite common [2], [4]. And despite the fact that there is advanced architecture to execute processes like multi-sensor fusion, or tiny-object detection, most of them are not used in this particular area of the research [5], [10], [2].

The current review is a list of the recent developments in sensing technologies, machine-learning-based detectors and workflow architectures to identify the hotspots of debris-accumulation in the surface waters. The complex of these issues gives a clue on how to build scalable, AI-assisted systems capable of detecting and monitoring floating-waste hotspots.

2. Methodological Background

A wide range of AI methodologies adopted to exploit radiative, textural, and contextual cues:

2.1 Index- and Threshold-Based Approaches

Spectral indices have been used by scientists to imply a mismatch between the plastics and the adjacent water including the Floating Debris Index (FDI) [1], the Plastic Index (PI), and the Normalized Difference Plastic Index (NDPI) [3]. The indices make use of the special light refraction and, in most cases, simple band ratios have effectively been used to isolate rubbish and water surfaces when a clear sky is present, and where there were no major variations in the lighting conditions [9]. That said, threshold-based techniques continue to possess infamous issues. The misclassification may be caused by the following factors: different illumination, sun-glint, turbidity of waters and overlapping between spectral signature of the plastics and the spectral signature of the natural organic or mineral elements [2], [10], [11].

2.2 Classical Machine Learning

The task of differentiating plastics and water has also been tried out using older machine-learning methods, such as Random Forests, Support Vector Machines, Gradient Boosted Trees, and similar types of models [12]. These techniques are usually based on handcrafted inputs which define spectral, spatial, or textural properties [9], [8]. They provide convenient reference points, however when mixed pixels, unpredictable backgrounds or fine and too scattered debris are encountered, they tend to fail in such feature-based models in a clean fashion [2], [13], [14].

2.3 Deep Learning

Deep learning has to a great extent replaced the previous methods due to its capability of learning multiple tiered spatial representation directly out of the visual field and adjusting to a variety of situations [3], [7]. Patch-level separation of debris in rivers has been applied by use of CNN-based classifiers [3], but object-detection classifiers, including YOLO, Faster R-CNN and variants, have demonstrated excellent performance in UAV imagery where small floating objects and larger clustered masses need to be detected [6], [15]. The U-Net, DeepLab and Mask R-CNN semantic-segmentation networks take it a step further where it can map hotspots of debris on a pixel-by-pixel basis [16], [6]. Latest, transformer-based architectures, like ViT and Swin Transformer, have demonstrated improved performance on a scene with cluttered surface-water features since they can better see the global context [10], [17].

2.4 Fusion and Hybrid Approaches

Another is also the multiplication of methods that combine multiple sources of sensing or modelling strategies. Integration of optical imagery and SAR data can be made, e.g., to facilitate more reliable detection when it is cloudy or when there are highly varying light conditions that cannot be used with a traditional optical sensing [18]. The fusion of UAV imagery and satellite data are also targeted at providing solutions to the mismatch of the resolution between the fine and large regional processing that simplifies the mapping of hotspots more precisely [9]. In other works, spectral-unmixing algorithms are combined with deep-learning networks, too, which constrains the effect that the mixed one has on the pixels and offers better generalization of the different seasons and water conditions [2], [7], [19].

3 Literature Review

3.1 Systematic Review

In this review, a protocol was already prepared in advance and therefore, the search and review of the relevant studies was transparent and reproducible. The purpose was to gather the work associated with AI-based detection of floating debris and also to locate hotspots of surface-water.. It focused on the articles released in the period of 2019-2025- the period when both remote sensing and AI-based methods of tracking the debris were growing in size at an alarming speed. This time is the response to the explosion of interest in deep-learning detectors such as YOLO models, Faster R-CNN, and Mask R-CNN, which started to be used increasingly in the studies relying on UAV and satellite data to locate the debris [15], [17]. It is also accompanied by the emergence of transformer-based models like Swin Transformer, ViT, and DETR, which researchers slowly turned to whenever they had to deal with complex or unclean surface-water images [9], [1].

Other notable developments in the same years were major improvements in Earth-observation data, such as the presence of hyperspectral imagery down to PRISMA, and the far more vigorous exploitation of Sentinel-1 SAR and Sentinel-2 multispectral products in work on the topic of debris [2], [3], [9], [11], [18]. The list of more than 120 publications used in the collection of Springer, Elsevier/ScienceDirect, Nature, IEEE Xplore, MDPI, ISPRS Archives/Annals, and Frontiers was narrowed with the help of screening the abstracts and verifying the full-texts.

3.2 Selection Criteria and Scope

The selection was conducted using a set of inclusion criteria [20] that were designed to ensure that it remained in line with the main aims of the review. Research must have been technically relevant, have had a methodologically consistent approach and have been in relation to the floating-waste detection or the hotspots themselves.

3.2.1 AI and Remote-Sensing Methodology

The study must have included one of the methods of AI or remote-sensing that is much important in detecting floating debris or mapping areas of accumulations. The list of techniques under discussion was chosen purposely as extensive as possible because the sphere is still evolving.

The works were considered in the following condition:

- Objects detectors such as the Faster R-CNN or Mask R-CNN models Deep-learning models - are frequently employed to detect the debris at different spatial linear scales [13], [15], [17], [21].
- The classification systems (also called CNN-based) that have been designed to categorize or classify floating debris based on pattern extracted out of optical data [14].
- Transformer-based models, including Swin Transformer, ViT and DETR which rely on self-attention mechanisms to learn busy or strongly varying water environments [17].

- The spectral-index method, which utilizes FDI or PI, or NDPI, or any index, which is also based on the spectral signature of floating debris [2], [3], [9].
- Hyperspectral analysis pipelines with PRISMA or similar apparatuses, often in a narrow spectral band [2], [4], [7], [10].
- SAR-based techniques whereby the floating material can be identified by data on backscatter or surface roughness based on the optical sensing which is difficult when dealing with light conditions and water [11], [18], [22].

3.2.2 Problem Domain

The studies were also required to be eligible meaning that they had to have touched at least one of the thematic areas that were directly related to monitoring floating-waste. The corresponding works were typically of one or more of the following categories:

This was also observed in rivers, lakes, reservoirs or coastal waters; here AI and remote sensing was used to locate or make an approximation of the distribution of surface-litter [1], [15], [23].

Observation of the drift or formation of a hot spot of debris, e.g. the observations of the clusterage of the debris, the hydrodynamic or meteorological conditions which alter the motion thereof, and the general trends of accumulation [19], [22].

Surface-water pollution and debris-transport modeling with the computation of movement pathways, characterization of the pollution levels or combination of different sensor types to control the environment and give early-warning initiatives [14].

3.2.3 Evaluation Requirements

Only the studies that gave quantitative measurements were taken into consideration. This was required to ensure that the results could be compared and the methods were tested in a quantifiable way.

The evaluation measures that were accepted were:

Markers of object-detection such as mAP (0.5-0.95), precision, recall, and F1-score that are common indicators of the accuracy and localization performance [1], [7], [12], [15], [17].

The criteria of classification like ROC-AUC or the overall accuracy were used when the main task was the discrimination of the type of a piece of rubbish or the nature of a material.

Pixel level Pixel-level segmentation scores Pixel-level segmentation scores are used in semantic or instance-segmentation problems to estimate the quality with which models detected debris in highly challenging water scenes [2], [3], [5], [21].

4 Proposed AI Model for Floating Waste Detection and Hotspot Mapping

Certain gaps of recurrence have been detected in literature: first of all, technical ones. There is still limited inland-water data, small objects are still hard to be observed with the model and spectral responses are usually mixed with turbidity or sun- glint. To add to it, an end-to-end system, which can be implemented directly into normal monitoring, is yet to be adopted. According to these arguments, we have come up with a scalable deep-learning pipeline that is designed to enhance the functionality of detecting floating debris in the surface waters as well as tracking down the origin of the concentration of the debris. The design is a combination of various data streams, state of the art detection frameworks and a streamline deployment platform which can operate both in real time or somewhere near real time.

4.1 Dataset Aggregation and Preprocessing

The pictures have been obtained via a number of free-source applications, such as Kaggle, Google pictures, Roboflow, Unsplash and other images of lakes, rivers, and mixed garbage locations with hand-prints to compile a working data set. These changes of conditions find a reflection in the pictures: the transparent or wavering water, the formations of mixed material, the quiet and the stormy sun rays along with the reflections of the inhumans. The annotations were also precise since all the pictures were manually marked with labels in order to identify floating plastics, natural debris, and high-intensity reflections.

The preprocessing phase involved a number of steps:

- input sizes Solution to different input sizes,
- scattering light and damaging noise by elementary enhancement filters is reduced,
- performing image processing to transform them into color space to enable the differentiation between debris and the surrounding water,
- deleting annotations using polygon and bounding-box tools of Roboflow.
- This completely leads to a more diversified training set and this is used to fill in the gap of inland-water imagery mentioned in the paper above.

4.2 Data Augmentation for Realistic Water Conditions

Since the scenes of surface-water fluctuate so rapidly in response to time of the day or night, the augmentation pipeline was also to be varied. It begins with more everyday geometric transformations, i.e. flips, rotation, scaling, changing shapes and orientations. The artificial shadows and brightness and contrast effects were then incorporated so as to follow the actual lighting changes.

4.3 Model Architecture

The hybrid deep-learning framework has been designed to meet the intricacy of the waste detection issue in inland waters [24] and it comprises of:

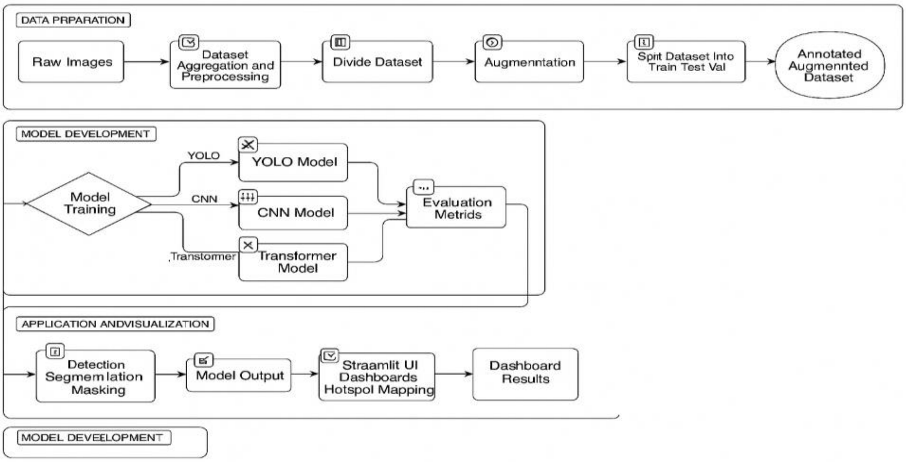

Fig.1 Overview of the proposed model workflow.

4.3.1 Base Detector: YOLO Variant

The chosen model being the core 2 is a YOLO-family detector (YOLOv7 or YOLOv8) because it is quick and efficient in detecting small objects. It has a feature-pyramid design, which makes it be able to process a mixture of object sizes, which can be applied when the observed debris is found to be either in a single or giant clump. Bounding-box labels were used to train the detector to assist the detector in establishing and locating the debris.

4.3.2 CNN and Transformer-Based Feature Enhancements

The architecture entails the combination of two kinds of feature pathways in order to provide a trade off in the local versus more global knowledge of the scene. A CNN base (learns finer-scale textures and edges); an optional transformer branch (learns long-range context) is ViT or Swin-T. Their combination is useful in helping the model to distinguish the debris on the visually identical patterns that comprise foam, glare and surface ripples. The two-track system further performs better when it comes to identifying partially submerged or obscured objects in the intricate water in scenarios.

4.4 Training and Evaluation Protocol

Normal object-detection is the basis that makes up the philosophy of training. The CIoU/DIoU loss is also loss to minimize the location of the bounding-box and the cross-entropy loss is used to track the predictions of the classification. AdamW is used to optimize the parameters, warm-restart cycles are used to stabilize the training and the learning rate is trained using a cosine-annaling schedule to prevent abrupt changes. The precision and recollection are measured together with F1-score and mAP (0.5-0.95) and provide a crude estimate concerning the efficiency of the model when it comes to identifying and locating debris.

4.5 Hotspot Mapping and Visualization Framework

In order to enable deployment of the system in non-research environment, the model is implemented in a lean deployment pipeline. Streamlit application allows a user to serve pictures, on-premise and on-cloud inferences as well as real-time detection output.

It is termed as a combination of detection and visualization that can be applied in filling the gap between the behaviour of the technical model and in the real world [25].

5 Critical Analysis and Findings

The 35 reviewed studies collectively demonstrate strong progress in floating waste detection; however, systematic examination reveals persistent technical, methodological, and operational gaps that limit large-scale inland deployment. This section presents a comparative analysis of study domains, sensor usage, methodological trends, and reported performance metrics, followed by a critical discussion of existing limitations and how the proposed deep-learning-based system addresses them. Fig.2 below gives details of Sensor and platform distribution across the selected studies.

5.1 Overview of Trends Across the Literature: Table 1 provides a consolidated summary of the sensors and platforms utilized across the reviewed studies [26].

Sentinel-2–based observations constitute the largest share of studies (41%), followed by UAV-driven detection approaches (31%). In contrast, SAR and hyperspectral techniques remain comparatively underutilized in the context of inland water monitoring.

Some researchers like [7], [8], [13] utilized supervised machine-learning techniques applied to satellite imagery and achieved high, scene-dependent precision. Wheareas some studies [1], [13], [15], [17], [21], [27], [28] implemented a deep-learning–based detection and counting framework, reporting coefficient of determination values reaching approximately $R^2 \approx 0.80$. Some [3] adopted spectral-index–based thresholding methods and demonstrated that floating debris becomes reliably detectable when occupying more than 10–20 meters of pixel fraction, whereas [2], [4],

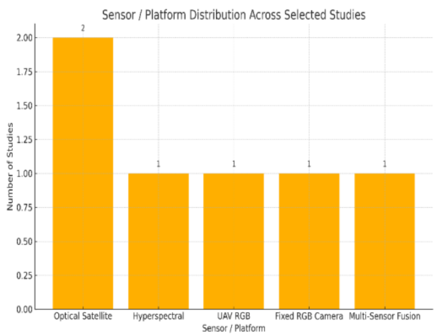

Fig.2 Sensor and platform distribution across the selected studies.

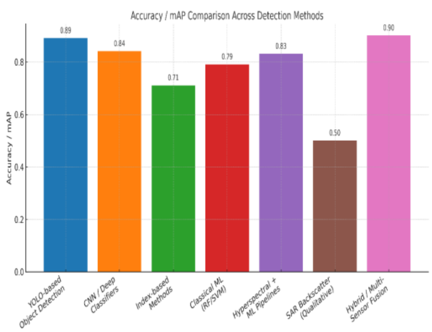

Fig. 3. Accuracy variation across different Detection methodologies evaluated int the reviewed studies.

[7], [10], [29] applied hyperspectral spectral-analysis and band-selection techniques, effectively identifying wavelength regions most discriminative for debris detection.

5.2 Comparative Findings from Key Studies

1. Optical Multispectral Approaches.

The optical multispectral applications continue to comprise a significant share of the research on wide-area floating-debris monitoring. As an example, it was reported in [7] that the pixel-level accuracies with the help of the Sentinel-2 data were more than 90 percent, though the results of the brew deteriorated significantly in presence of sun-glint or an increased turbidity. Similarly in [3], Floating Debris Index (FDI) was proposed and it was demonstrated that debris has to occupy about 20-30 per cent of a pixel in order to be detected. The optical properties of ocean and other water bodies debris were outlined in [10], [19], [29], who indicated that there is an overlap in the spectral properties of plastics, foams and organic matter and thus class separation cannot be assured. Another study by [2] observed that SWIR bands can be used to a certain degree in discrimination but still the case is not very successful when the debris is under water. Goddijn-Murphy et al.[22], [29] continued to note that reflectance-based methods are still affected by the atmosphere, variations in illumination and mixed-pixels.

2. UAV / Drone-Based Vision Studies.

The type of resolution that is provided by UAV-based approaches can typically detect considerably smaller debris. In a case study, [3], [5], [15], [17] has attained 92% accuracy with a bridge-mounted camera with RF/SVM classification. Kaizouji [5] discovered that YOLOv5 was more effective in detecting river debris compared to Faster R-CNN particularly small or odd shaped objects. In another research, [6], [15], the precision was more than 90 percent with UAV-collected RGB images. Multispectral UAV sensors have been also employed- [11], [18] demonstrated that these sensors are more consistent in detecting fine debris and even when the lighting conditions are changed.

3. Hyperspectral and Laboratory Studies

The key feature of hyperspectral imaging is the capability to distinguish materials in accordance with their spectral characteristics. [4] and [2] both showed that hyperspectral data could effectively be used to identify the various types of polymers. It has been shown by [30] that similar benefits were realized in freshwater settings. The disadvantage is that these systems are costly and they have to be carefully calibrated as well as the generated data may be challenging to routinely process. Dierssen and Garaba [4] stressed the fact that until these methods can be transferred to larger scale use, there is still a need to have standard atmospheric corrections and more fully developed spectral libraries.

4. SAR-Based Approaches.

SAR imaging is now more noticeable as a supplemental instrument, especially where optical sensors are not doing well. In studies by [10], [22], the floating debris was found to change the short-wave surface roughness in a manner that is visible in C-band imagery of SAR. Other studies, such as [11] and [22], [29], discovered that SAR is reasonably effective with large top of the debris slicks but struggle to detect small or discrete particles. In spite of the fact that SAR provides credible cloud-penetration and night-time functionality, its functionality remains affected by sea-state circumstances and background clutter.

5.3 Dataset Limitations Observed Across Studies

There is a considerable percentage of the considered studies, 26 of 35, that cite data limitations as one of the primary issues. The scarcity of annotated datasets of inland waters is one of the most prevalent ones, as stated in [4], [5], [7], [15], [17], [21], [27] and [22]. It was also reported by many studies that models trained in marine environments do not transfer to rivers well due to domain-shift effects [2], [3], [22], [29]. Another problem that was mentioned frequently was spectral confusion of plastics, foams, and organic materials, particularly in multispectral data sets[1], [5]. Also, some of the authors noted that there were not enough samples of small debris (less than 10 pixels), and it is difficult to train the models that can identify the objects of fine scale [10], [13], [23], [23], [28], [30]. The available data are also biased toward clear-water imagery, which makes it difficult to generalize to the more muddy conditions [30]. In general, riverine debris data sets are still incomplete and significantly underdeveloped compared to those that can be used to study the marine environment [22][30].

5.4 Object Detection Challenges in Dynamic Water Surfaces

In the literature, some of the problems are recurring in the process of identifying debris on the moving water surfaces. The most frequent sources of error are sun-glint and bright reflections [11], [13], [17], [29] and can easily lead to a false positive due to the unpredictable reflective qualities of floating materials. Detection is also complicated by wave motion and say variation of surface-textures [2], [9], [10], [19] as very fast changes may lead to changes in bounding boxes or can even disappear between frames. Spectral separability is further diminished by turbidity, suspended sediments and partial submergence which complicates the identification of objects [17], [18]. Mixed garbage piles: so when plastics, vegetation, and driftwood are on the same clog, this is a different complication, since the model often misclassifies similar-looking things, as also seen in [22], [29].

5.5 How the Proposed Model Overcomes Current Research Limitations

The model presented in this paper is an answer to some of the weaknesses present in the literature. First, the scarcity of labeled inland-water data is offset through the creation of a multi-source dataset that integrates materials on Kaggle, Google images, Roboflow, and Unsplash, to provide the researcher with a more diverse set of conditions than those datasets that were employed in previous research like [7]. The traditional issue of small objects detection is approached with a hybrid design that combines the use of YOLO with CNN and Transformer elements, which enhance the detection of small pieces of debris and act upon the concerns expressed in [15], [17], [31]. Issues related to glare, turbidity and surface variability due to waves are addressed using a hydrological augmentation strategy that is a simulated representation of such environmental conditions that directly resolves the limitations mentioned as in [6], [23], [23]. Lastly, Transformer-based global-context modeling will increase robustness in scene visually cluttered scenes, which can benefit generalization and can directly react to the noted constraints, [1], [17].

6. Mathematical Analysis:

6.1 Feature Space Formulation

Where H, W and C are the spatial height, spatial width, and spectral channels, respectively, let an input surface-water image be represented by $\mathcal{I} \in \mathbb{R}^{H \times W \times C}$. The image is converted to a latent feature tensor with the aid of the backbone feature extractor $\Phi(\cdot)$,

$$\mathbf{F} = \Phi(\mathcal{I}), \mathbf{F} \in \mathbb{R}^{h \times w \times d} \quad (1)$$

where d denotes learned representations. It is an extractor that combines convolutional filters to encode local texture and optional attention based modules to encode long range spatial dependencies across water surfaces. In order to minimize sensitivity to change in illumination and scale variations, the normalization of features is used as:

$$\hat{F}_{i,j,k} = \frac{F_{i,j,k} - \mu_k}{\sigma_k + \epsilon} \quad (2)$$

where $\mu_k$and $\sigma_k$represent the mean and standard deviation across channel k, and $\epsilon$is a numerical stabilizer.

6.2 Detection and Classification Model

This involves the identification and categorization of the way to address the identified problem. Allow the detection head to produce a list of candidate regions:

$$\mathcal{B} = \{(b_m, s_m) \mid m = 1,2, \dots, M\} \quad (3)$$

and $b_m = (x_m, y_m, w_m, h_m)$ represents the parameters representing the bounding-box and $s_m \in [0,1]$ the score representing the confidence of the floating debris being present.

Each candidate has a classification function $g(\cdot)$ which attaches to it a semantic label:

$$\hat{y}_m = \arg\max_{c \in \mathcal{C}} g_c(b_m) \tag{4}$$

where C={"waste","background"}. One achieves the final prediction set on the process of non-maximum suppression to remove spatial redundancies.

6.3 Pixel Level consistent Segmentation.

In the case of hotspots delineation, the framework also generates a pixel-wise probability map $P \in [0,1]^{H\times W}$. $\mathcal{S}$is is binary segmentation mask calculated based on the value threshold τ:

$$\mathcal{S}(i,j) = \begin{cases} 1, & P(i,j) \geq \tau \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

The Intersection over Union (IoU) is used to measure the spatial overlap between the predicted and reference masks:

$$\text{IoU} = \frac{|\mathcal{S} \cap \mathcal{G}|}{|\mathcal{S} \cup \mathcal{G}|} \tag{6}$$

and Dice similarity Coefficient (DSC):

$$\text{DSC} = \frac{2|\mathcal{S} \cap \mathcal{G}|}{|\mathcal{S}| + |\mathcal{G}|} \tag{7}$$

These measures are especially applicable to fragmented and irregular distributions of debris which are usually found in inland waters.

6.4 Performance Metrics

Where TP, FP, TN, and FN represent true positive, false positive, true negative and false negative, respectively. The measures of standard evaluation are determined to be:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

The given values are experimentally reported and reported values of accuracy accuracy ≈ 0.91, F1 ≈ 0.91) are balanced (detection sensitivity and false-alarm control) and this is essential to operational monitoring systems. It should be noted that the model handles a wide range of computers with varying memory speeds and latencies.

6.5 Computational Efficiency and Latency

It is necessary to note that the model supports a large variety of computers with diverse memory speeds and latencies. The overall inference time / frame of the image is:

$$T_{\text{inf}} = T_{\text{feat}} + T_{\text{det}} + T_{\text{post}} \tag{8}$$

In which T_{feat} is feature extraction T_{det} detection and classification and T_{post} post-processing and visualization.

The rate of processing possible is estimated as:

$$\text{FPS} = \frac{1000}{T_{\text{inf}}} \tag{9}$$

Practically, the performance of the framework that can be provided in real time, surpassing the conventional limits (≥25 FPS) proves the importance of implementing the framework to conduct an uninterrupted monitoring of rivers or reservoirs.

6.6 Model Robustness and error Behavior.

The probability $\mathcal{E}$ of the system of ensemble errors as a dependence upon individual detector error e -and inter-model correlation ρ:

$$\mathcal{E} \leq \bar{e}(1 - \rho) \tag{10}$$

The reduction of systematic misclassification due to glare, turbidity, and texture noise generated by the waves is through the low correlation among feature representations, which are obtained through a variety of augmentations and multi-scale learning.

Conclusion and Future Research Directions:

In the current paper, the reviews of new tendencies in the utilization of AI to floating waste recognition were examined on the basis of the research that encompasses multispectral and hyperspectral imaging, SAR data, UAV measurements, and different computer-vision techniques. Despite an improvement that is apparent, the review has demonstrated that majority of the available methods still fail- especially where applied on in-land water like rivers, lakes and reservoirs. The shifting hydrology, strong sun-glint, turbidity, and mixed nature of debris are problematic on a chronic basis and most systems have not yet been able to detect only small or partially submerged objects. Worse still, the quality of datasets available is low and the reporting standards are skewed towards the low side and it is harder to trust the models generalization.

To fill these gaps, the existing work suggests a modular deep-learning framework created around some of these building blocks: a multi-source dataset, the choice of hydrological improvements, a hybrid YOLO-CNN-Transformer model, a hotspot-mapping pipeline facilitated by the use of GIS tools. The idea here is to have a system designed that will be capable of handling more environment conditions and can be deployed with ease through a lightweight simple interface. The results suggest that it can make the field move in the direction of scalable working solutions to the monitoring of debris in the inland waters.

References:

[1] C. van Lieshout, K. van Oeveren, T. van Emmerik, and E. Postma, “Automated River Plastic Monitoring Using Deep Learning and Cameras,” *Earth Sp. Sci.*, vol. 7, no. 8, 2020, doi: 10.1029/2019EA000960.

[2] P. Tasseron, T. van Emmerik, J. Peller, L. Schreyers, and L. Biermann, “Advancing floating macroplastic detection from space using experimental hyperspectral imagery,” *Remote Sens.*, vol. 13, no. 12, 2021, doi: 10.3390/rs13122335.

[3] L. Biermann, D. Clewley, V. Martinez-Vicente, and K. Topouzelis, “Finding Plastic Patches in Coastal Waters using Optical Satellite Data,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020, doi: 10.1038/s41598-020-62298-z.

[4] S. P. Garaba and H. M. Dierssen, “Hyperspectral ultraviolet to shortwave infrared characteristics of marine-harvested, washed-ashore and virgin plastics,” *Earth Syst. Sci. Data*, vol. 12, no. 1, pp. 77–86, 2020, doi: 10.5194/essd-12-77-2020.

[5] M. M. Duarte and L. Azevedo, “Automatic Detection and Identification of Floating Marine Debris Using Multispectral Satellite Imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–16, 2023, doi: 10.1109/TGRS.2023.3283607.

[6] M. Geraeds, T. van Emmerik, R. de Vries, and M. S. bin Ab Razak, “Riverine plastic litter monitoring using Unmanned Aerial Vehicles (UAVs),” *Remote Sens.*, vol. 11, no. 17, pp. 6–8, 2019, doi: 10.3390/rs11172045.

[7] À. Solé Gómez, L. Scandolo, and E. Eisemann, “A learning approach for river debris detection,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 107, no. December 2021, 2022, doi: 10.1016/j.jag.2022.102682.

[8] A. Using and S. Images, “Areas Using Satellite Images and Machine Learning,” 2024.

[9] K. Themistocleous, C. Papoutsa, S. Michaelides, and D. Hadjimitsis, “Investigating detection of floating plastic litter from space using sentinel-2 imagery,” *Remote Sens.*, vol. 12, no. 16, 2020, doi: 10.3390/RS12162648.

[10] S. P. Garaba and H. M. Dierssen, “An airborne remote sensing case study of synthetic hydrocarbon detection using short wave infrared absorption features identified from marine-harvested macro- and microplastics,” *Remote Sens. Environ.*, vol. 205, no. April 2017, pp. 224–235, 2018, doi: 10.1016/j.rse.2017.11.023.

[11] S. Lee *et al.*, “Quantitative Detection of Floating Debris in Inland Reservoirs Using Sentinel-1 SAR Imagery: A Case Study of Daecheong Reservoir,” *Water (Switzerland)*, vol. 17, no. 13, pp. 1–17, 2025, doi: 10.3390/w17131941.

[12] T. Fian and G. Hauger, “Identifying High-Risk Patterns in Single-Vehicle, Single-Occupant Road Traffic Accidents: A Novel Pattern Recognition Approach,” *Appl. Sci.*, vol. 14, no. 19, 2024, doi:

10.3390/app14198902.
[13] A. P. Periyasamy and R. Perumalsamy, “Machine learning for microplastic quantification: Techniques, challenges, and future directions,” *Clean. Water*, vol. 4, no. July, p. 100158, 2025, doi: 10.1016/j.clwat.2025.100158.
[14] I. A. Arturo and P. L. Corcoran, “Categorization of plastic debris on sixty-six beaches of the Laurentian Great Lakes, North America,” *Environ. Res. Lett.*, vol. 17, no. 4, 2022, doi: 10.1088/1748-9326/ac5714.
[15] N. Maharjan, H. Miyazaki, B. M. Pati, M. N. Dailey, S. Shrestha, and T. Nakamura, “Detection of River Plastic Using UAV Sensor Data and Deep Learning,” *Remote Sens.*, vol. 14, no. 13, 2022, doi: 10.3390/rs14133049.
[16] T. D. Leder, N. Leder, M. Baučić, and S. Bačić, “Optical Remote Sensing Methods for Floating Marine Debris Detection – Review and Bibliography Analysis,” *Trans. Marit. Sci.*, vol. 13, no. 2, pp. 1–31, 2024, doi: 10.7225/toms.v13.n02.005.
[17] S. Goharinezhad, A. Hadi, and S. Mirzaei, “Automated defect classification and localization in sewer pipelines using hybrid ResNet50–Swin transformer and modified YOLOv8 on CCTV inspection images,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–26, 2025, doi: 10.1038/s41598-025-27765-5.
[18] M. D. Simpson *et al.*, “Monitoring of Plastic Islands in River Environment Using Sentinel-1 SAR Data,” *Remote Sens.*, vol. 14, no. 18, 2022, doi: 10.3390/rs14184473.
[19] A. de Fockert *et al.*, “Assessing the detection of floating plastic litter with advanced remote sensing technologies in a hydrodynamic test facility,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–15, 2024, doi: 10.1038/s41598-024-74332-5.
[20] D. Version, “Blue whiting (Micromesistius poutassou): behaviour and distribution in Greenland waters,” 2021.
[21] M. Rußwurm, S. J. Venkatesa, and D. Tuia, “Large-scale detection of marine debris in coastal areas with Sentinel-2,” *iScience*, vol. 26, no. 12, 2023, doi: 10.1016/j.isci.2023.108402.
[22] N. Maximenko *et al.*, “Towards the integrated marine debris observing system,” *Front. Mar. Sci.*, vol. 6, no. JUL, 2019, doi: 10.3389/fmars.2019.00447.
[23] L. C. M. Lebreton, J. Van Der Zwet, J. W. Damsteeg, B. Slat, A. Andrady, and J. Reisser, “River plastic emissions to the world’s oceans,” *Nat. Commun.*, vol. 8, pp. 1–10, 2017, doi: 10.1038/ncomms15611.
[24] K. et al 2023, *No Title 済無 No Title No Title No Title*, vol. 32, no. 3. 2021.
[25] A. K. Y. Gin and M. Supervisor, “Modelling the Transport and Fate of Emerging Contaminants in Singapore Fresh and Coastal,” 2023.
[26] T. Gouin, R. Ellis-Hutchings, L. M. Thornton Hampton, C. L. Lemieux, and S. L. Wright, “Screening and prioritization of nano- and microplastic particle toxicity studies for evaluating human health risks – development and application of a toxicity study assessment tool,” *Microplastics and Nanoplastics*, vol. 2, no. 1, 2022, doi: 10.1186/s43591-021-00023-x.
[27] M. Rußwurm, S. J. Venkatesa, and D. Tuia, “Large-scale detection of marine debris in coastal areas with Sentinel-2,” *iScience*, vol. 26, no. 12, p. 108402, 2023, doi: 10.1016/j.isci.2023.108402.
[28] T. P *et al.*, “Spatial Quantification of Marine Litter Using Satellite and Drone Data through Empirical and Deep Learning Techniques – A Case Study from India,” 2024, [Online]. Available: https://www.researchsquare.com/article/rs-4255724/v1
[29] V. Martínez-Vicente *et al.*, “Measuring marine plastic debris from space: Initial assessment of observation requirements,” *Remote Sens.*, vol. 11, no. 20, pp. 8–14, 2019, doi: 10.3390/rs11202443.
[30] D. González-Fernández and G. Hanke, “Toward a harmonized approach for monitoring of riverine floating macro litter inputs to the marine environment,” *Front. Mar. Sci.*, vol. 4, no. MAR, pp. 1–7, 2017, doi: 10.3389/fmars.2017.00086.
[31] S. Shukri, “YOLO-Based Real-Time Floating Debris Detection and Counting in Rivers for Early Flood Warning and Long-Term Water Resource Management,” pp. 1–34.