Journal of Edge Computing, 2026, Vol. **5**, Iss. 1, pp. 105–125 10.55056/jec.1051

Seamless monitoring and detection of waste hazards in floating water and water reservoirs using Internet of Things integrated deep learning algorithms

Neeta Shirsat[1], V. Nirmalrani[1]

[1] *Sathyabama Institute of Science and Technology, Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai, Tamil Nadu, 600 119, India*

Abstract. One of the most essential and current research areas is waste hazard monitoring in floating water and water reservoirs. The increasing population and urbanisation contribute to waste hazards, negatively impacting water quality, human health, and environmental resources. Many research methods have focused on applying metaheuristic and image-processing techniques to analyse and detect waste hazards in floating water. The detection efficiency was good; however, their computational complexity was high and not cost-effective. Additionally, it takes longer to analyse the data. Coastal, riverine, and seaside areas require effective detection and monitoring systems to generate alerts for waste-hazard removal. Otherwise, these hazards pose numerous health risks to the public and degrade water quality. However, it remains a complex technological challenge due to real-time constraints, environmental changes, and the lack of automation in traditional systems. This paper addresses this major challenge and aims to design and implement an IoT-integrated deep learning model incorporating principal component analysis (PCA), grey-level co-occurrence matrix (GLCM), and Fast R-CNN to enable automatic, optimal waste-hazard detection in dynamic floating and static water bodies. Various IoT sensors and edge devices are installed in water bodies to collect data. Initially, the PCA method analyses the data and improves the entire Fast R-CNN model by efficiently extracting, compressing, and denoising features, while GLCM captures discriminative textural information. Moreover, the Fast R-CNN model reduces computational complexity while improving detection and classification accuracy. Both input and predicted data are securely transmitted through fog computing and interconnected throughout the entire architecture. The deep learning model is implemented with IoT data, and the results are validated. The output demonstrates that PCA-GLCM-integrated Fast R-CNN provides high accuracy in detecting different types of waste hazards with a lower false-positive rate and reduced latency.

Keywords: IoT, deep learning, CNN, floating waste detection, smart environmental monitoring, sustainable development

1. Introduction

Water pollution with hazardous floating waste is a serious global environmental issue and an immediate threat to aquatic life, public health, and sustainable water management. Urbanisation, industrial waste discharge, and poor waste disposal practices have greatly increased the presence of plastic debris, chemical spills, oil residues, and organic pollutants in rivers, lakes, and reservoirs [4, 18, 21]. These pollutants degrade water quality, disrupt aquatic biodiversity, and pose serious health threats to communities that depend on surface water for their daily needs. While there have been various efforts in water conservation and pollution control, ineffective waste management and hazard mitigation have been hindered by a lack of real-time, intelligent, and scalable monitoring solutions. The rise of oceanic plastic waste is unprecedented. It impacts crucial ecosystem services, among other concerns. Figure 1 shows how the waste hazards pollute the water bodies. Plastic-based pollution poses significant health and environmental risks, with clean plastic waste accounting for up to 80% [12]. A massive amount of solid waste flowing through the floating water is caused by

0009-0001-9958-0417 (N. Shirsat); 0000-0003-4169-9228 (V. Nirmalrani)
neeta.shirsat@gmail.com (N. Shirsat); nirmalrani.it@sathyabama.ac.in (V. Nirmalrani)

Received 2025-07-10 Accepted 2026-04-28 Published 2026-04-28 Version of record 2026-05-21

widespread urbanisation and human activities, leading to an alarming rise in ocean plastic pollution. For a detailed assessment of the environmental impacts of the river on the ocean, it is essential to focus on detecting and removing waste [7].

Figure 1: Sample waste hazards in floating waste.

Traditional water quality monitoring methods depend largely on manual sampling, laboratory analysis, and periodic inspections. These approaches are labour-intensive, slow, and geographically constrained, leading to delays in detection and response. Furthermore, they are insufficient in providing comprehensive and continuous surveillance of wide water bodies, particularly in remote or underserved regions [6]. Therefore, an urgent need exists for automated, real-time, and intelligent systems to detect, classify, and alert concerned authorities when hazardous floating waste is detected in water environments. Integrating IoT and deep learning technologies is the most efficient way to transform environmental monitoring systems. IoT continuously collects data using a network of smart sensors, cameras, and edge devices deployed across lakes and rivers. With support from a deep learning model, sensor data can be processed into actionable insights, whether to detect pollution or predict related risks. PCA and GLCM extract and refine features, while Fast R-CNN classifies specific waste types from the candidate regions identified by YOLO.

This paper proposes a comprehensive IoT-based deep learning framework for inspecting floating water bodies and reservoirs for hazardous waste. Combining real-time imaging and sensor data from floating water bodies and reservoirs and subsequently analysing them with a hybrid PCA-GLCM-Fast-R-CNN architecture achieves high accuracy in waste detection and classification. The platform also supports edge computing for low-latency local processing while enabling remote access, visualisation, notification, and cloud integration, thus ensuring an end-to-end environmental watch. The goals of this research include: designing and deploying a distributed IoT sensing infrastructure for continuous monitoring of floating waste across multiple water environments; developing and training deep learning models for detecting and classifying hazardous waste using real-time data; and commissioning an energy-efficient, scalable, and cost-effective monitoring system applicable in urban and rural environments. The work also aims to inform environmental authorities and stakeholders in real-time, enabling timely intervention and policy enforcement. By integrating ecological sensing, AI analysis, and automated alert mechanisms, this research establishes a scalable, intelligent, and effective solution to water pollution.

Region-based convolutional neural networks for object detection were introduced in 2014, with R-CNN [5] establishing the foundational two-stage detection paradigm. However, region-based methods are computationally slow despite achieving high accuracy. Single-stage detectors such as YOLO address this by predicting bounding boxes and class probabilities simultaneously across the whole image, substantially improving throughput. This paper improves classification accuracy by contributing the following:

- A PCA-GLCM-Fast-R-CNN model is developed for detecting and classifying organic and hazardous waste materials from floating water bodies.
- Smart sensors, IoT devices, and drones are deployed in water bodies to collect real-time data and identify waste hazards.
- PCA reduces feature dimensionality and suppresses noise, while GLCM extracts discriminative textural features; together they improve detection accuracy and reduce computational overhead.
- Edge computing handles data locally and cloud integration stores processed results, reducing transmission latency across the monitoring architecture.
- Fast R-CNN classifies detected regions into waste categories with high accuracy and low computational overhead, operating on region proposals supplied by YOLO.
- Experimental results show that the proposed system achieves 98.2% accuracy, supporting SDG 6 (Clean Water and Sanitation) and SDG 14 (Life Below Water) objectives.

2. Literature survey

This section provides a detailed discussion of prior research, outlining its merits and demerits for detecting hazards in water systems. It helps in understanding the existing research problem, which in turn aids in selecting the design method to provide a more effective solution. Waste materials in the water layer cause significant pollution and various health issues. Compared to the existing analysis, the accumulation of floating water waste has surged from millions to trillions. Traditionally, multiple algorithms have been proposed to accurately detect and recognise water waste. However, those algorithms required substantial human intervention to detect objects effectively. Therefore, an automated intelligent system is needed to overcome these issues. Shirsat and Nirmalrani [17] have proposed a deep learning-based automated system for detecting floating water pollutants to reduce time consumption and human intervention. This model efficiently identifies and classifies floating water waste, then sends notifications to local authorities. This study is limited to real-time data under various weather conditions and requires a large training dataset.

A Bluetooth-controlled robotics system is designed to detect and remove trash from floating water [1]. Humans control the robot's actions through Bluetooth by connecting the propellers with two DC motors. The robot's sensor detects the garbage floating in the water, and the conveyor belt directs it into the collection box. The system can collect 10 kg of trash, covering approximately 3,000 sq. m. This system can operate for 4 hours without recharging. Nunkhaw and Miyamoto [16] presented a YOLOv5 model integrated with DeepSORT for monitoring and detecting waste hazards in floating water. The model had a mean average precision of over 88% and high F1 scores for accurately tracking solid waste, plastics, and waste paper. The accuracy was consistently 80% across all scenarios, suggesting the method could be used in real-world river applications.

Zhang et al. [23] have presented an advanced object recognition model using YOLO to detect and classify various types of solid waste floating on river surfaces. A transfer learning model is used to improve classification accuracy. The YOLO-based transfer learning model has been tested and validated. Finally, the results showed that the proposed YOLO-WASTE model achieved an mAP of 92.23% and took only 0.424 seconds to detect an image, which is much faster than image classification algorithms. Vijayanti et al. [20] have proposed a hybrid model combining several DL models with transfer learning-based models, such as VGG19 and ResNet50, to detect and efficiently classify waste hazards in floating water bodies. The proposed model was tested on the Aqua Trash dataset, and the VGG and ResNet models were evaluated against the traditional CNN model to assess performance. The experimental result demonstrated that the VGG model performed well in detecting waste in water bodies.

Junzhe et al. [13] have proposed a YOLO v5 algorithm to address issues in water bodies caused by floating waste in rivers and oceans. The overall efficacy of the proposed work is improved by applying K-fold cross-validation, transfer learning, and test-time augmentation methods to prevent overfitting and enhance detection and classification accuracy. The experimental results showed that the proposed model achieved an mAP@50 score of 98.1% and increased detection speed by 67.7 milliseconds, indicating better performance than traditional DL models. Kundu, Sharma and Pillai [14] proposed an AI-powered mechanism for trash detection using the ResNet50 model. It also integrates well-developed ML models, computer vision, and robotics technology to predict the waste hazards in water bodies. The experimental results showed that, with the help of ResNet50, the proposed TL model achieved 95.6% classification and detection accuracy.

He et al. [9] proposed a new network model, EC-YOLOX, to detect and classify various types of waste, including plastic bags, milk boxes, grass, leaves, bottles, and other floating objects in rivers. The river-floating object dataset is used to evaluate the overall efficiency of the proposed EC-YOLOX model in detecting floating waste in water. The experimental result stated that the proposed EC-YOLOX model efficiently detected various floating waste types in water. The evolving nature of deep learning techniques enables the development of deep neural network architectures tailored to application needs, and, as a result, object detection applications have shown improved performance. Generally, these applications include two phases, namely, region-based and feature-based object detection. Based on this, serial algorithms [2, 10] and single-stage algorithms [13, 19] are used for object detection.

Hasan et al. [8] have investigated the detection and quantification of microplastics in water systems using convolutional neural networks (CNNs). In the short term, the system will incorporate IoT sensors installed in water systems to collect real-time data. Using a diverse set of microplastic sample datasets, the CNN model was employed to identify various microplastic shapes, sizes, and colours. This enhances the system's responsiveness and reduces latency. Environmental agencies and researchers can also utilise the technique to mitigate microplastic pollution in aquatic ecosystems, enabling continuous monitoring of microplastic levels. Yang et al. [22] applied a CNN-based VGG16-15 model to predict 15 types of floating objects in water. A training and validation dataset was split at a 4:1 ratio, yielding 5,707 images. Finally, the model was optimised by modifying the neural network, applying learning rate decay, implementing early stopping, and performing data augmentation. With these enhancements, the model achieved an accuracy of 93.86%, 10.09% higher than the standard VGG-16 model. Codes-Alcaraz, Puerto and Rocamora [3] have proposed a standardised and efficient method for characterising floating waste to detect changes in quantity and composition. A dataset of 477 images of floating plastic items across various environments was created to train the YOLOv5s algorithm. Key limitations have been identified based on the above literature review.

3. Limitations and motivation

The major limitations of implementing an IoT-enabled monitoring system include restricted availability and quality of training data for CNN and Fast R-CNN models. Since model performance depends directly on training data, collecting images of diverse polluted water bodies and various hazardous floating objects remains a significant challenge. Energy constraints pose another obstacle, particularly when powering smart sensors and drones in remote areas; edge computing mitigates this but does not eliminate the energy cost of continuous real-time monitoring. Variable environmental factors, including seasonal variation, weather, and geography, affect model reliability in conditions not represented during training, requiring ongoing adaptation to new environments and waste types. High-performance computational hardware is necessary to ensure low-latency data transmission and processing under dynamic conditions.

A primary global concern is the accelerated deterioration of water quality caused by floating waste in water bodies, which directly impacts biodiversity, public health, and sustainable water resource management. Traditional techniques are inadequate for managing floating waste in large, dynamic water bodies in real time. The motivation for this research is therefore to deploy a low-intervention IoT system integrated with deep learning to identify hazardous floating waste, classify waste types accurately, and support timely intervention. The proposed PCA-GLCM-Fast-R-CNN model leverages complementary feature extraction strategies to promote a cleaner, more sustainable aquatic environment, aligning with the United Nations' Sustainable Development Goals (SDG 6: Clean Water and Sanitation and SDG 14: Life Below Water).

4. Problem statement

Increasing pollution from floating hazardous waste, including plastics, biological materials, and toxic contaminants, threatens aquatic ecosystems and human health worldwide. In addition to affecting marine life, it also disrupts the ecosystem's equilibrium and significantly reduces water quality in aquatic environments. Conventional techniques lead to delayed action because they rely on manual labour, making the process time-consuming and inefficient. This paper presents a fully automated IoT architecture integrated with deep learning for identifying hazardous floating waste. The PCA-GLCM-Fast-R-CNN model is used to extract features, detect objects, and classify waste hazards: PCA compresses and denoises feature representations, GLCM captures textural information, and Fast-R-CNN performs region-based detection and classification. The result is a reliable, adaptable, and time-efficient solution for environmental monitoring.

5. Proposed model

The model proposes combining PCA and GLCM with Fast R-CNN for feature extraction, creating a robust approach to image classification of hazardous floating waste in water bodies. This system extracts information from images or video frames of water bodies to identify and classify dangerous objects without human intervention. The proposed model includes data preprocessing, data augmentation, feature extraction, object detection, and classification. The comprehensive analysis of each stage of the method, along with the mathematical expressions, is outlined below. Figure 2 depicts the workflow of the paper. The sensors and IoT devices installed in the water bodies continuously generate image data, which serves as input data. Some devices may generate video data. The video is converted into images to predict waste hazards.

As shown in figure 2, this paper presents a PCA-GLCM-integrated Fast R-CNN model for detecting hazardous waste in floating water. The input image and video-based samples are captured using the IoT sensors. The raw input data are preprocessed using the Gaussian filtering technique and PCA, which performs normalisation, noise reduction, dimensionality reduction, and resizing. GLCM then extracts textural features from the preprocessed data. To detect waste objects in the input real-time image sample, the YOLO technique is applied. The Fast R-CNN model classifies detected waste objects into different classes. The model's overall performance is evaluated using various metrics, including accuracy, precision, recall, and others. The following subsection provides an elaboration on each step of the proposed approach.

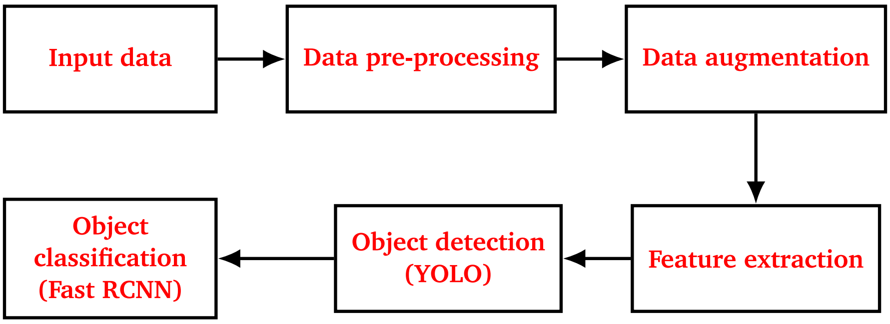


Figure 2: Proposed workflow.

5.1. Data pre-processing

Data pre-processing is the foundational and crucial stage of building the proposed model. Data collected from polluted water bodies containing hazardous floating objects must be processed to remove noise, enhance image quality, resize images, and standardise pixel values, thereby transforming them into efficient training data for deep learning models.

Step 1: Resizing the images

The images collected from polluted water bodies containing hazardous floating objects may vary in resolution. To ensure consistency, all images are adjusted to a predefined resolution of M × N. The resizing process can be mathematically represented as:

$$I_{resized}(x,y) = I\left(\frac{x \times H}{M}, \frac{y \times W}{N}\right) \tag{1}$$

The equation defines image resizing via interpolation. In this context, I represents the original image while $I_{resized}$ refers to the resized image. The original image has dimensions $H{\times}W$, where H is the height and W is the width. The target dimensions for the resized image are $M{\times}N$, where M denotes the new height and N the new width.

Step 2: Image normalisation

To speed up training and mitigate vanishing gradients, normalisation is applied to scale pixel intensities to the range $[0,1]$. The normalisation operation is mathematically defined as:

$$I_{norm}(x,y) = \frac{I(x,y) - I_{min}}{I_{max} - I_{min}} \tag{2}$$

The equation normalises an image, scaling its intensity values to a standard range, typically 0-1. Here, *I(x,y)* denotes the pixel intensity value at a specific coordinate *(x,y)*. The normalisation process adjusts these values based on the minimum and maximum intensities in the image, represented by I_{min} and I_{max}, respectively. The resulting normalised image, I_{norm}, ensures that the pixel values are proportionally distributed within the new range, enhancing contrast and improving image processing performance.

Step 3: Noise removal and denoising

The noise present in the images, such as reflections, shadows, and tiny floating particles, can be removed using a Gaussian blur filter. The mathematical expression for the Gaussian filter is given as:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

The equation represents the Gaussian blur function, which smooths an image by reducing noise and detail. In this formula, *G(x,y)* denotes the Gaussian-blurred image at the coordinates *(x,y)*. The standard deviation of the Gaussian distribution, represented by σ, controls the extent of blurring, where higher values lead to more substantial blurring effects. The function applies a weighted averaging technique, giving more influence to central pixels while reducing the impact of those farther away, following a Gaussian distribution.

5.2. PCA-based feature preprocessing for hazardous waste detection

Principal component analysis (PCA) reduces data dimensions by using statistical methods to transform high-dimensional data into lower dimensions while preserving essential information. The PCA technique transforms data into orthogonal components, captures the maximum variance, and reduces dimensionality. PCA can reduce the dimensionality of data for hazardous waste detection in floating water, preserving its statistical properties. The PCA formula is given as:

$$Z = XW \tag{4}$$

Here, the transformed data is represented as Z; X represents the input data; and W is the eigenvector matrix of the covariance matrix of X. To prevent model overfitting and enhance model training, a small amount of noise is added, and the principal components are shifted, thereby improving hazardous waste detection accuracy.

5.3. GLCM-based feature extraction for hazardous waste detection

The grey-level co-occurrence Matrix is a statistical method for analysing and extracting features from images of floating hazardous waste in water. It describes the spatial relationship between pixel intensities within an image. The features extracted using GLCM are:

- *Contrast* measures the local variation in the image. High contrast means more variation in pixel intensity.

$$Contrast = \sum_{i,j} p(i,j)(i-j)^2 \tag{5}$$

- *Energy* measures the texture's uniformity. Higher energy indicates a uniform texture.

$$Energy = \sum_{i,j} p(i,j)^2 \tag{6}$$

- *Homogeneity* measures how similar intensity values are across the image. A higher value indicates smoother textures.

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{7}$$

- *Correlation* measures the linear dependency between pixel pairs. High correlation values indicate a more predictable texture.

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}, \tag{8}$$

  where $p(i,j)$ represents the probability of occurrence of pixel intensity I and j. μ denotes the mean intensity and σ represents the standard deviation.

5.4. Object detection and classification

In this paper, the YOLO technique is applied to detect objects on the surface of floating water. The Fast R-CNN model is then used to classify those detected objects into waste categories.

YOLO is a deep learning-based object detection algorithm. It processes the whole image simultaneously, making it more effective for real-time object detection in water. The images are split into a YOLO grid, and the bounding boxes, confidence scores, and class probabilities are predicted simultaneously. All existing object detection methods use image regions to identify objects. The network does not analyse the entire image. Instead, most of the image contains objects. The YOLO algorithm is an object detection algorithm significantly different from the region-based algorithms mentioned earlier. In YOLO, a single convolutional network predicts a bounding box and a class probability for each bounding box.

The working principle of YOLO is that an image is divided into an s×s grid, and we take m bounding boxes from each grid. Class probability and offset values are the network outputs for each bounding box. A bounding box with a higher threshold value is selected to identify the object in the image. The accuracy of the detected object within a selected bounding box is based on the confidence score. The mathematical representation of the YOLO detection function is represented as:

$$P_r(class_i|object) \times P_r(object) \times B_x \times B_y \times B_w \times B_h, \tag{9}$$

where the probability (P_r) of an object under the $class_i$ is represented $P_r(class_i|object)$, the probability of an object being present in a grid cell is represented by $P_r(object)$ and the coordinates of the bounding box are represented as $B_x \times B_y \times B_w \times B_h$.

YOLOv8x and YOLOv8n are evaluated for floating water object detection, trained on datasets of floating waste to quickly detect and classify it. The floating dataset includes plastic bottles, clothes, cans, etc. To reduce regression loss, the bounding box is used as follows:

$$L = \lambda_{coord} \sum_{i=0}^{S^2} 1_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \tag{10}$$

In the above equation 1_{ij}^{obj} represents the object (*obj*) present at the location (i, j), the x_i, y_i represents the coordinate of the actual object in the input image, and $\hat{x}_i, \hat{y}_i$ represents the coordinate of the predicted objects' bounding box. YOLO executes with minimal processing time. It detects multiple objects and accurately identifies hazardous floating waste objects in the water, facilitating real-time monitoring and waste management. YOLO has a faster processing speed than other object detection algorithms, processing 45 frames per second. If the object is small within the image, YOLO struggles with small-object detection, highlighting its limitation. For example, detecting flocks of birds is difficult for YOLO, owing to the algorithm's spatial limitations.

5.5. Object classification using F-RCNN

This paper uses YOLOv8 for object detection and a PCA-GLCM-integrated Fast R-CNN for classification, demonstrating that the combined approach produces higher accuracy with lower computational complexity than either model alone. The objective is to minimise manual effort in water monitoring and support ecosystem preservation.

In Faster R-CNN, object detection proceeds in two stages: a region proposal network (RPN) first generates candidate bounding boxes, and a CNN classifier then assigns class labels and refines those boxes. The Faster R-CNN loss is represented by:

$$L(p, u) = L_{cls}(p, u) + \lambda[u \geq 1] L_{reg}(t_u, v), \tag{11}$$

where L_{reg} represents the smooth L1 loss, L_{cls} denotes SoftMax-based classification loss, and predicted bounding box coordinates are represented by t_u, v indicates the ground truth bounding box coordinates, the predicted probability of object presence is denoted by p, and u indicates the ground truth label.

The combination of PCA-based feature compression and GLCM-based texture features supplied to Fast R-CNN enables high-accuracy detection and classification of floating waste. The final output identifies unwanted materials and flags pollutants in water resources. The input image is fed into the convolutional layers, which learn and extract relevant features to create a feature map. Identifying objects in the images is essential for building anchors and bounding boxes. Thus, the feature map is sent to the RPN, which generates region proposals by drawing bounding boxes using a regressor and fitting them to objects. The feature maps are combined with the region proposals that were obtained and sent to ROI pooling, which extracts feature vectors for each object proposal. A sequence of fully connected layers processes the feature vector before passing it into the SoftMax layer and bounding box regressor. The SoftMax layer predicts whether an object is present inside the ROI. At the same time, the bounding box regressor adjusts the bounding box boundaries, relocalizes the object, and refines its placement.

5.6. Experimental setup

Fast image processing and model training are performed on a system with an Intel Core i7 processor, 64 GB RAM, and 1 TB storage, running Windows 10. Model development utilises Python, along with TensorFlow and Keras libraries, for the software setup. The CNN architecture is employed to extract high-level features from the input images, while Fast R-CNN performs region-based classification of the detected waste objects.

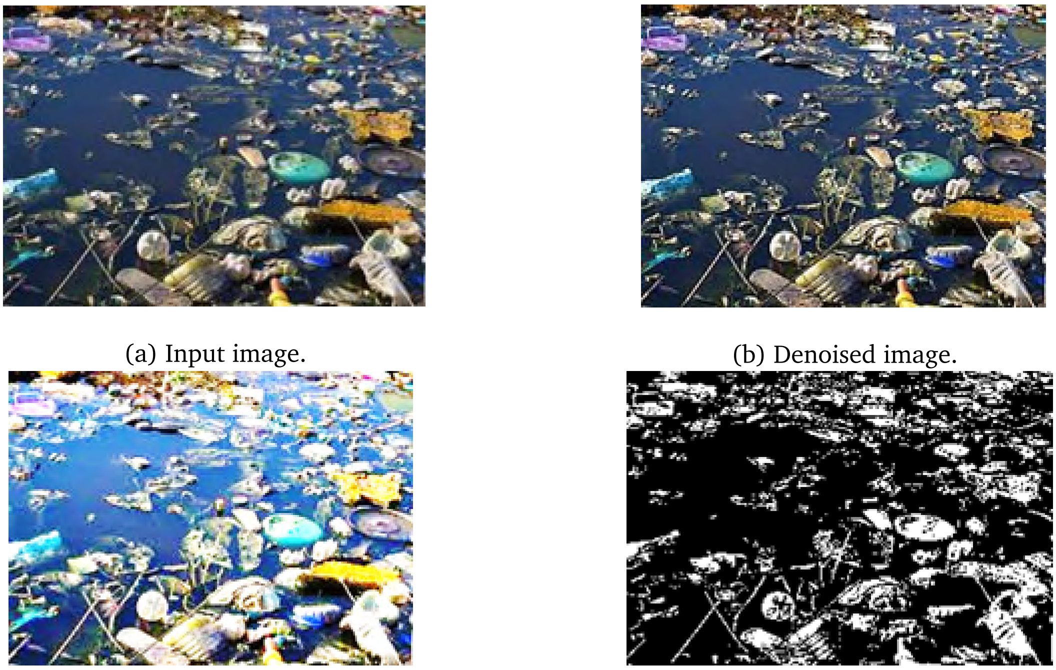
5.7. Dataset

The method for detecting specific hazardous waste in a floating water body using CNN and Fast R-CNN involves capturing, processing, and classifying images of the water body. Real-time images of water surfaces contaminated by floating waste are taken, and continuous images are captured using a high-resolution camera mounted on a floating device or drone. After capture, these images are transmitted to a computing device equipped with a high-performance graphics processing unit (GPU) for deep learning model processing. The dataset consists of more than 10,000 images of floating waste, including plastic, cans, bottles, medical waste, and oil spills, and the model is trained on it [11]. The input real-time images are captured from various geographical locations, including rivers, lakes, and coastal regions, from both rural and urban areas. Input data from multiple locations enhances the model's performance in handling diverse water surface and lighting conditions. The input samples comprise images of various hazardous wastes, including medical waste, metals, plastics, organic waste, and oil spills. The input data are split into 80:20 (training: test), and the model's performance is evaluated using accuracy, precision, recall, and F1-score. The PCA-GLCM-Fast-R-CNN model is proposed to achieve high detection accuracy for hazardous waste in floating water bodies through real-time monitoring.

6. Results and discussion

In this section, the simulation results of the proposed model are discussed in detail, accompanied by multiple graphical representations. The model's performance is evaluated using IOU, accuracy, precision, recall, mAP, and F1-score. The analysis results are compared among the deep learning approaches YOLOv8x, YOLOv8n, and RCNN. The different types of loss during model training and validation for each model are also evaluated and depicted in this section. The image pre-processing result of the proposed model on performing various pre-processing steps is shown in figure 3: (a) input image, (b) denoised image, (c) colour-enhanced image, and (d) colour-converted image. These steps further enhance the quality of the input sample and make the model more capable of predicting and classifying waste in floating water. The predicted output of the proposed model is shown in figure 4.

 10.55056/jec.1051



(a) Input image.

(b) Denoised image.

(c) Enhanced image.

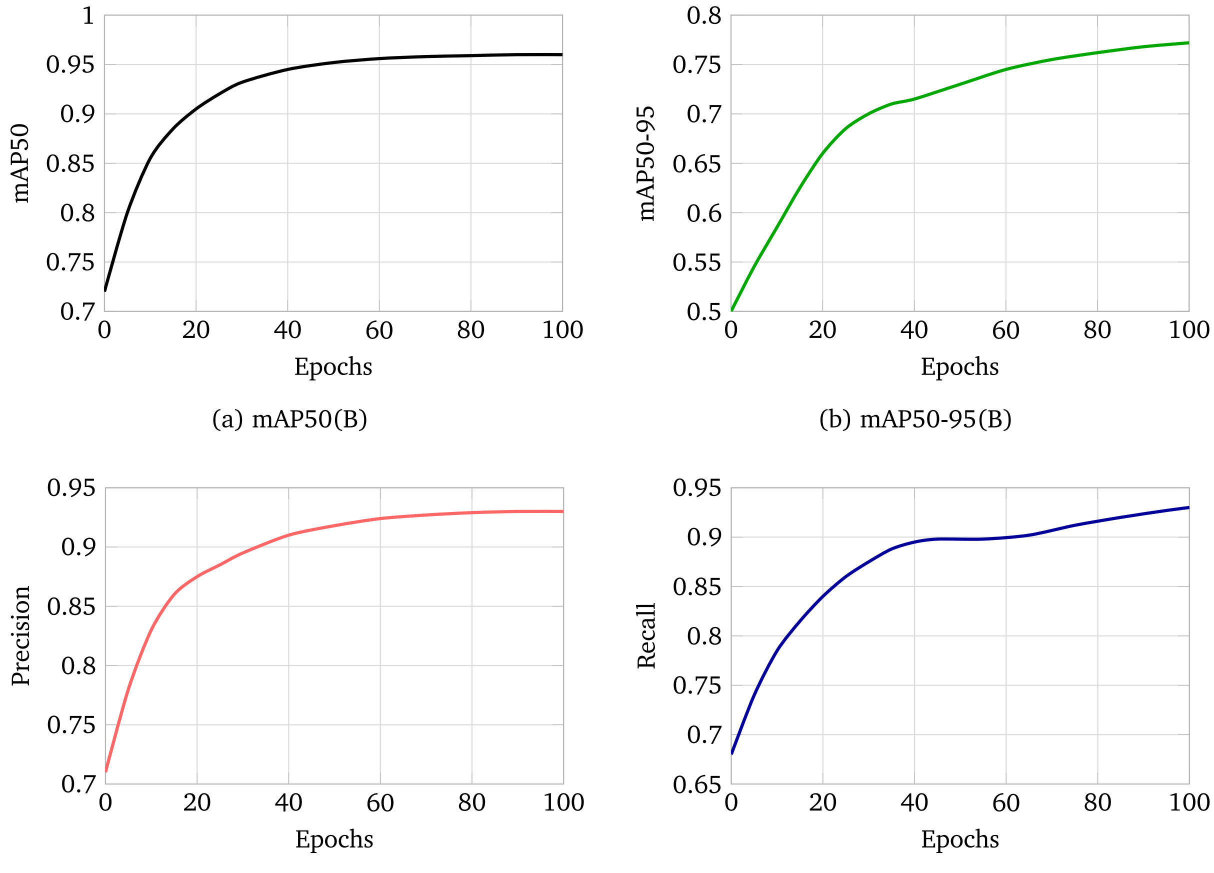
(d) Color-converted image.

Figure 3: Image processing output.

Figure 4: Predicted result.

Figure 5(a) shows the mAP50, which denotes the model's accuracy in object detection when the intersection between the predicted and ground-truth bounding boxes is 50%. It also shows an increase in mAP50, surpassing 0.9 after 20 epochs, indicating the model's rapid learning to identify objects with high accuracy. The graphical representation flattens after 60 epochs, suggesting the model is converging.

Figure 5(b) shows mAP50-mAP95, which comprehensively evaluates the model's precision across multiple IoU thresholds, ranging from 0.5 to 0.95. The model's performance is evaluated by its ability to accurately detect objects, even with low intersection-over-union (IoU) values. The figure illustrates



(a) mAP50(B) (b) mAP50-95(B)

(c) Precision(B) (d) Recall(B)

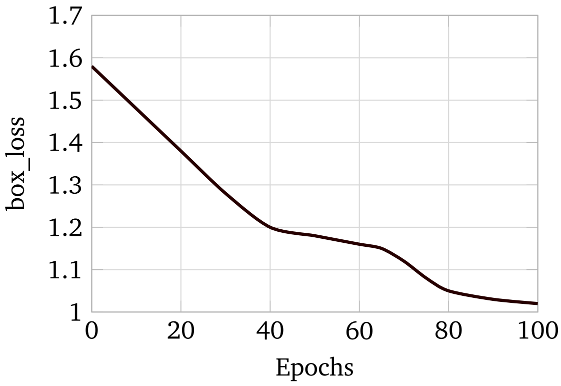
Figure 5: Performance analysis of YOLOv8x.

steady growth, starting low and reaching approximately 0.75 after 100 epochs, demonstrating the model's efficiency in feature extraction. Figure 5(c) presents numerical data that precisely quantifies the number of objects detected by the model. Initially, precision is low, which can lead to potential mislabeling in the early stages. However, it consistently improves, stabilising at around 0.9 after 30 epochs, indicating the model's high accuracy in identifying contaminated water. Figure 5(d) represents recall metrics, which quantify the actual number of objects detected in the given image. Recall begins at a low value and increases steadily, stabilising at 0.9 after 30 epochs. This result highlights the model's effectiveness in detecting contaminated materials in water with minimal false negatives.

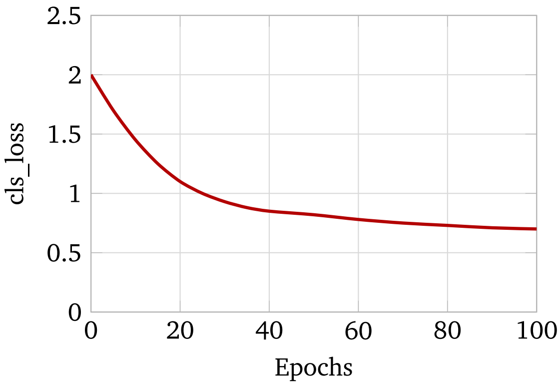
The graphical representation of three parameters illustrates the training loss curves for the object detection model, showing box loss, classification-based loss (cls_loss), and distribution focal loss (dfl_loss) over 100 training epochs. The box loss graph shown in figure 6(a) demonstrates a gradual decline from approximately 1.5 to a value close to 1.0, indicating that the model progressively improves in accurately predicting the bounding box coordinates of the detected objects with greater precision. The classification loss shown in figure 6(b) declines from 1.6 to approximately 0.9, demonstrating the model's increasing accuracy in identifying hazardous waste objects. The distribution focal loss (dfl_loss) shown in figure 6(c) primarily improves local accuracy, resulting in a consistent decrease from 1.6 to nearly 1.2, thereby leading to more reliable bounding box predictions.

The consistent reduction in losses across all three stages, without any fluctuations or plateaus, indicates an optimised modern learning strategy, powerful feature extraction, and reduced miscalculation errors. The stability observed in the later epochs confirms that the model exhibits minimal overfitting and has acquired robust features for detecting hazardous waste in floating water.

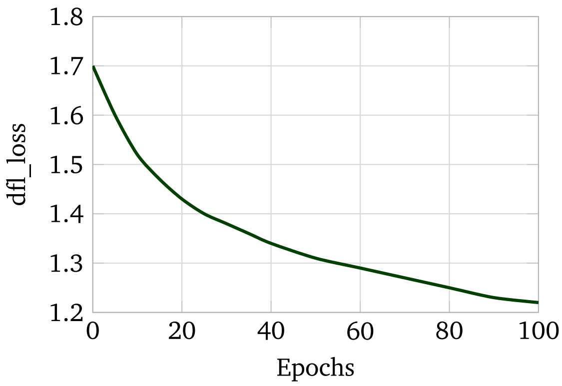
Hazardous waste detection in floating water is illustrated by four basic graphs in the figure, which



(a) train/box_loss



(b) train/cls_loss

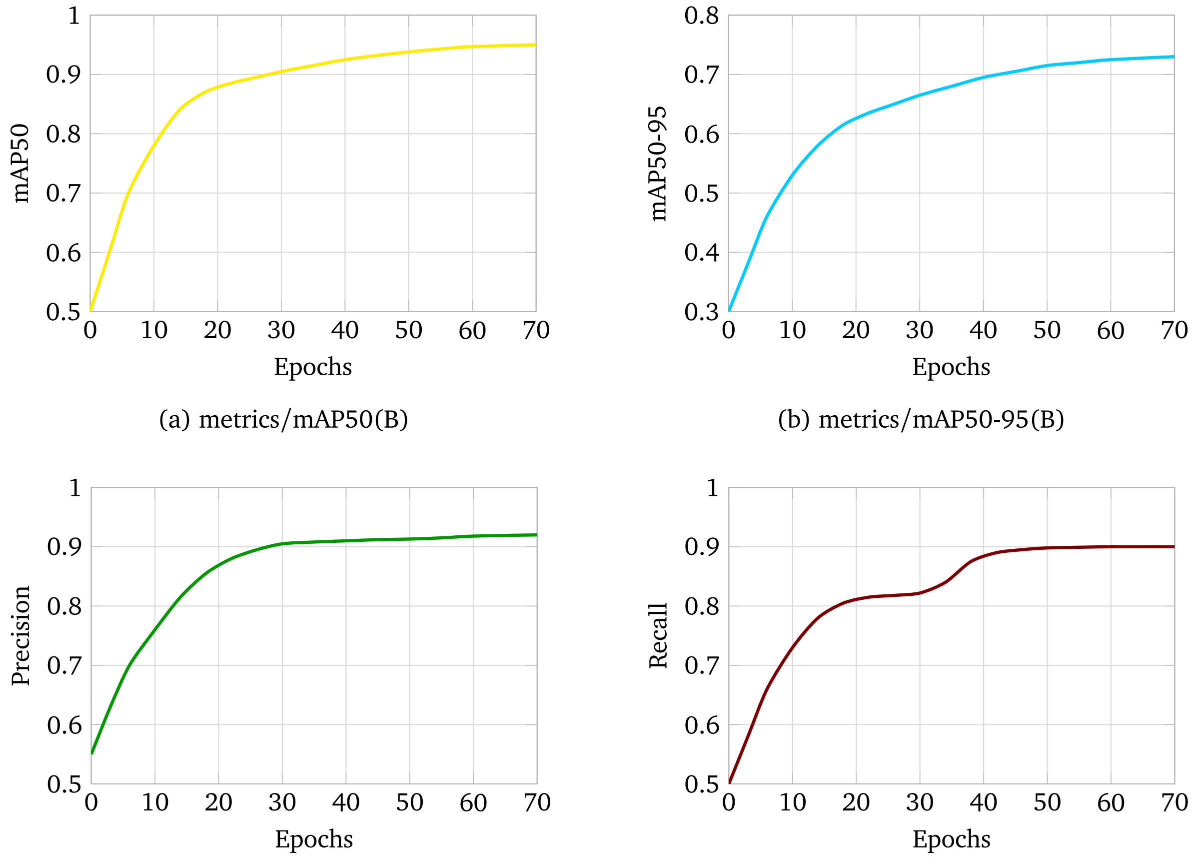


(c) train/dfl_loss

Figure 6: Loss value of YOLOv8x.

are used to analyse the performance of the YOLOv8n (nano) model. Figure 7(a) describes the mean average precision (mAP50) rapidly increasing from 0.5 to over 0.9 within just 20 epochs. This assumes that the model quickly learned to detect with high accuracy. Figure 7(b) shows mAP50-95, the mean average precision across IoU thresholds from 0.5 to 0.95. It gradually increases from 0.3 to 0.7, reflecting the model's improvement in generalisation and in predicting object localisation with dynamic overlap thresholds. Figure 7(c) shows a sharp increase from 0.6 to 0.9, indicating the model's ability to reduce false-positive detections by improving accurate hazardous waste analysis. Figure 7(d) shows a recall value that rapidly increases from 0.5 to 0.9, indicating the model's efficiency in detecting most ground-truth objects and thus reducing false negatives. The high mAP, precision, and recall combination demonstrates the YOLOv8n model's robust learning ability, ensuring accurate and efficient hazardous waste detection with fewer false alarms and making it suitable for real-time floating waste detection applications.

Figure 8 portrays the training loss curves of the YOLOv8n (nano) model for hazardous waste detection in floating water. Figure 8(a) (box loss) shows the bounding box regression loss for predicting the objects' bounding box coordinates. The model focuses on the regression loss, which decreases from approximately 1.4 to 0.9, enabling more accurate object localisation. Figure 8(b) (classification loss) represents the classification loss that detects the objects and corrects misclassifications. Classifying the hazardous waste validates the model's accuracy, as it improves from 1.6 to 0.6. Figure 8(c) (distribution focal loss) reflects localisation quality by predicting precise boundary offsets. The model's efficacy lies in refining object boundaries. The steady, gradual reduction in all three losses, without signs of overfitting, confirms the model's stable learning and strikes a balance between accuracy and speed, making it a powerful AI-driven application for detecting hazardous waste in floating water.

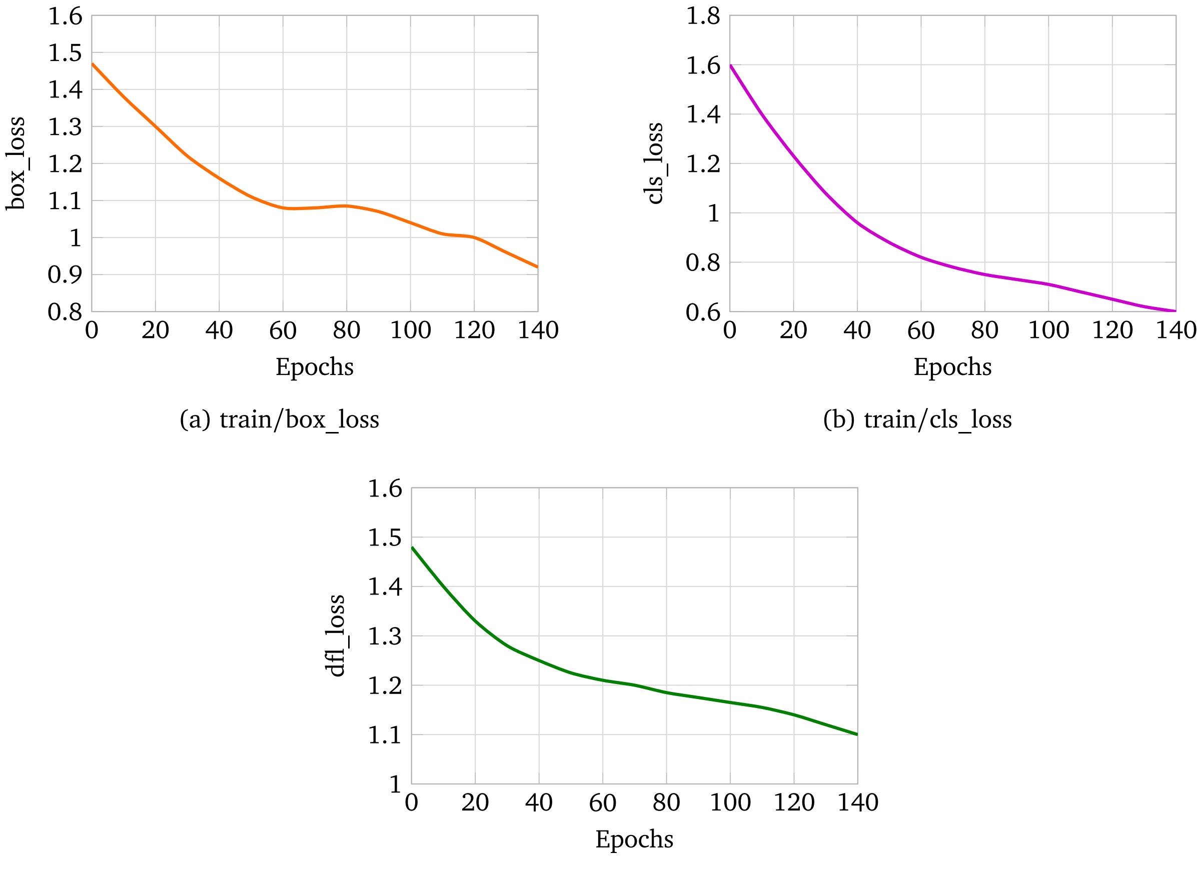


(a) metrics/mAP50(B)
(b) metrics/mAP50-95(B)
(c) metrics/precision(B)
(d) metrics/recall(B)

Figure 7: Performance evaluation YOLOv8n (nano).

The training and validation losses for object detection in hazardous waste, such as plastic, cans, bottles, and floating organic debris, are shown in five graphs. Figure 9(a) represents the classification loss constantly decreasing from 1.1 to around 0.8 over 100 epochs, confirming that the model has steadily improved in classifying hazardous waste objects. Figure 9(b) shows the distributional focal loss, which remains at zero throughout training, confirming that this loss term is not active in the RCNN configuration evaluated here. Figure 9(c) shows the Intersection over Union (IoU) loss ranging from 1.05 to 0.45, indicating that the model has improved at predicting bounding box coordinates with high overlap. Figure 9(d) shows the validation mAP@0.5, which remains near zero until epoch 20, then rises sharply to approximately 0.85 and continues to around 0.90 by epoch 100, highlighting the model's effectiveness once training stabilises. Figure 9(e) shows mean average precision across several IoUs (0.1–0.7), indicating that the model can detect objects of varying sizes and shapes. The overall trend of declining loss and rising mean average precision values highlights the model's progressive learning, improving both localisation accuracy and detection confidence for hazardous waste in floating water.

The seven graphs in figure 10 jointly depict the model's training performance and its accuracy in detecting hazardous floating waste in water bodies. A gradual rise in mean average precision at 50% IoU (mAP@0.5) from 0.82 to nearly 0.96 is observed in figure 10(a). It indicates that the model has learned to predict objects with higher accuracy. The model's increased proficiency in identifying objects is confirmed by observing figure 10(b) (mAP@0.5:0.95), which shows a consistent rise in mean average precision across varying IoU thresholds from approximately 0.45 to 0.76. The increase in precision values from approximately 0.88 to 0.95, as observed in figure 10(c), indicates that the model has achieved higher image classification accuracy with reduced false positive detections. The improvement in recall from approximately 0.86 to 0.92, as observed in figure 10(d), highlights



(a) train/box_loss

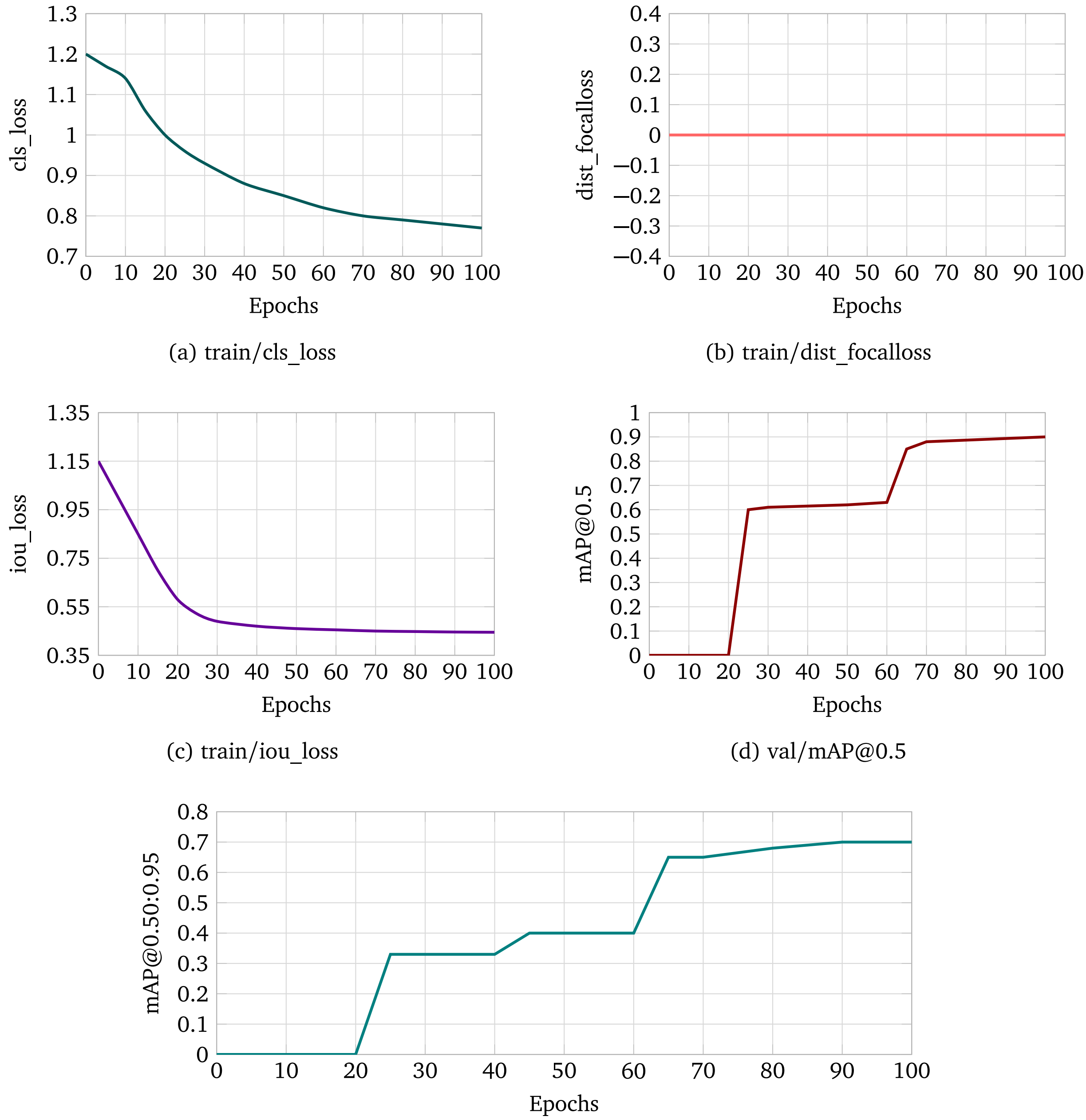
(b) train/cls_loss

(c) train/dfl_loss

Figure 8: Loss value of YOLOv8n.

the model's enhanced ability to classify most ground-truth objects while minimising false negatives. Regarding the training aspect, the decline in box loss from 0.042 to around 0.022, as observed in figure 10(e), indicates the model's improved accuracy in bounding box coordinate detection. The steep drop in classification loss from approximately 6.5 to around 1.3, as observed in figure 10(f), confirms the model's efficient learning in detecting hazardous floating waste with reduced errors. The decrease in object loss from 0.019 to 0.012, as shown in figure 10(g), indicates the model's improved ability to detect objects accurately. The decreased losses and increased precision, recall, and mAP values together demonstrate high detection accuracy, minimised misclassification, and robust object localisation for hazardous floating waste.

The graphs shown in figure 11 demonstrate the efficiency analysis and training behaviour of the RCNN model for identifying hazardous waste materials, such as plastic bags, nets, ropes, and bricks, in floating water. Figure 11(a) (precision) illustrates that precision for all three classes rises rapidly within the first 14 epochs, peaking near 94% for plastic bags before stabilising between 81 and 87% by epoch 30. Figure 11(b) (recall) illustrates a rapid rise in recall, reaching approximately 97% for plastic bags and stabilising between 87 and 90% for the other classes by epoch 30. Figure 11(c) (IoU) illustrates the Intersection over Union rising to approximately 95% for plastic bags after ten epochs, with the remaining classes reaching 83–92%, highlighting the model's precise prediction of object boundaries. Figure 11(d) (F1-score) illustrates values peaking near 95% for plastic bags and stabilising between 84 and 88% for all classes after epoch 14, confirming balanced classification capability. Figure 11(e) (loss) shows training loss decreasing rapidly from 0.048 in the first ten epochs and stabilising near 0.004, while validation loss converges to approximately 0.005. The sustained improvement across all metrics, combined with the rapid reduction in training and validation loss, demonstrates that the RCNN model reliably identifies hazardous waste in floating water, achieving



(a) train/cls_loss

(b) train/dist_focalloss

(c) train/iou_loss

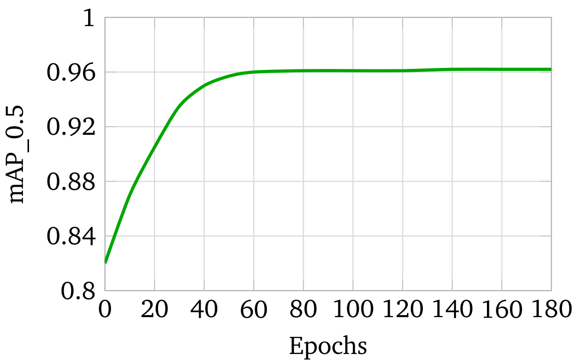
(d) val/mAP@0.5

(e) val/mAP@0.50:0.95

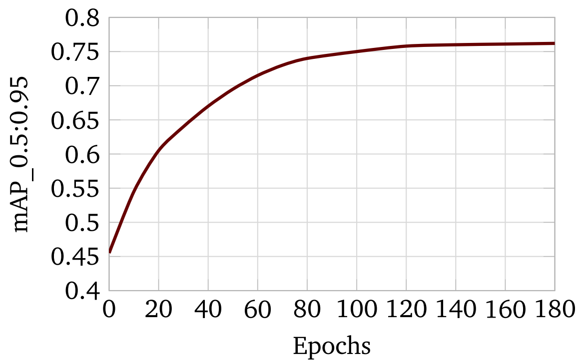
Figure 9: Training and validation result of YOLOv8x.

accurate classification and precise bounding-box localisation.

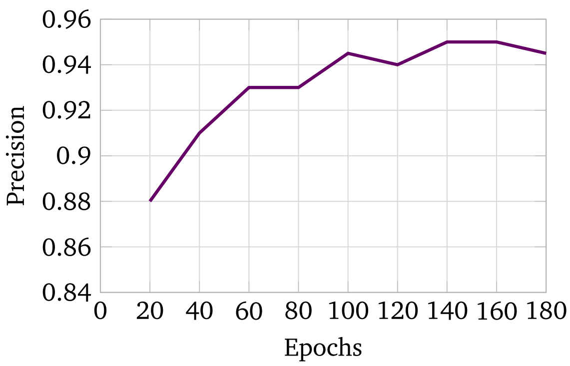
To represent the individual contribution of the proposed PCA and GLCM with FRCNN, which is evaluated in terms of accuracy, precision, recall, and F1-score. The analysis results in table 1 show that the simple FRCNN model for detecting hazardous waste in river water achieves an accuracy of 94.6%. When implementing the GLCM technique with FRCNN, the accuracy range increases to 96.8%. It increases the accuracy by extracting more essential texture features from the input data. Similarly, the PCA + FRCNN model further increases the accuracy, precision, recall, and F1-score to 97.2%, 97.0%, 97.1%, and 97.1%, respectively. The proposed combination of PCA, GLCM, and the FRCNN model achieves the highest accuracy and performance metrics, with results of 98.2% accuracy, 97.8% precision, 98.0% recall, and 97.9% F1-score. From the results, it is clear that GLCM is better suited for extracting textural features, while PCA is more effective for denoising and reducing the dimensionality of the input data. The synergy between the two preprocessing stages makes the



(a) metrics/mAP_0.5

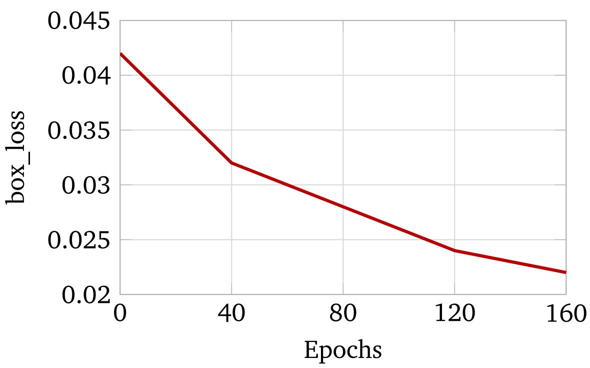


(b) metrics/mAP_0.5:0.95

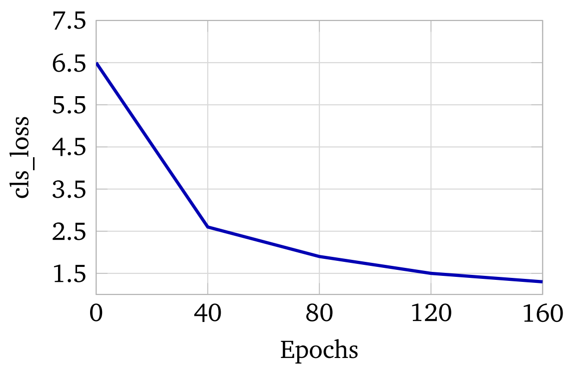


(c) metrics/precision

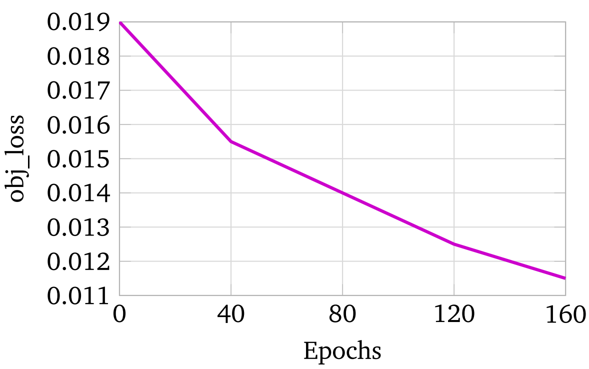
(d) metrics/recall



(e) train/box_loss



(f) train/cls_loss



(g) train/obj_loss

Figure 10: Training and validation result of YOLOv8n.

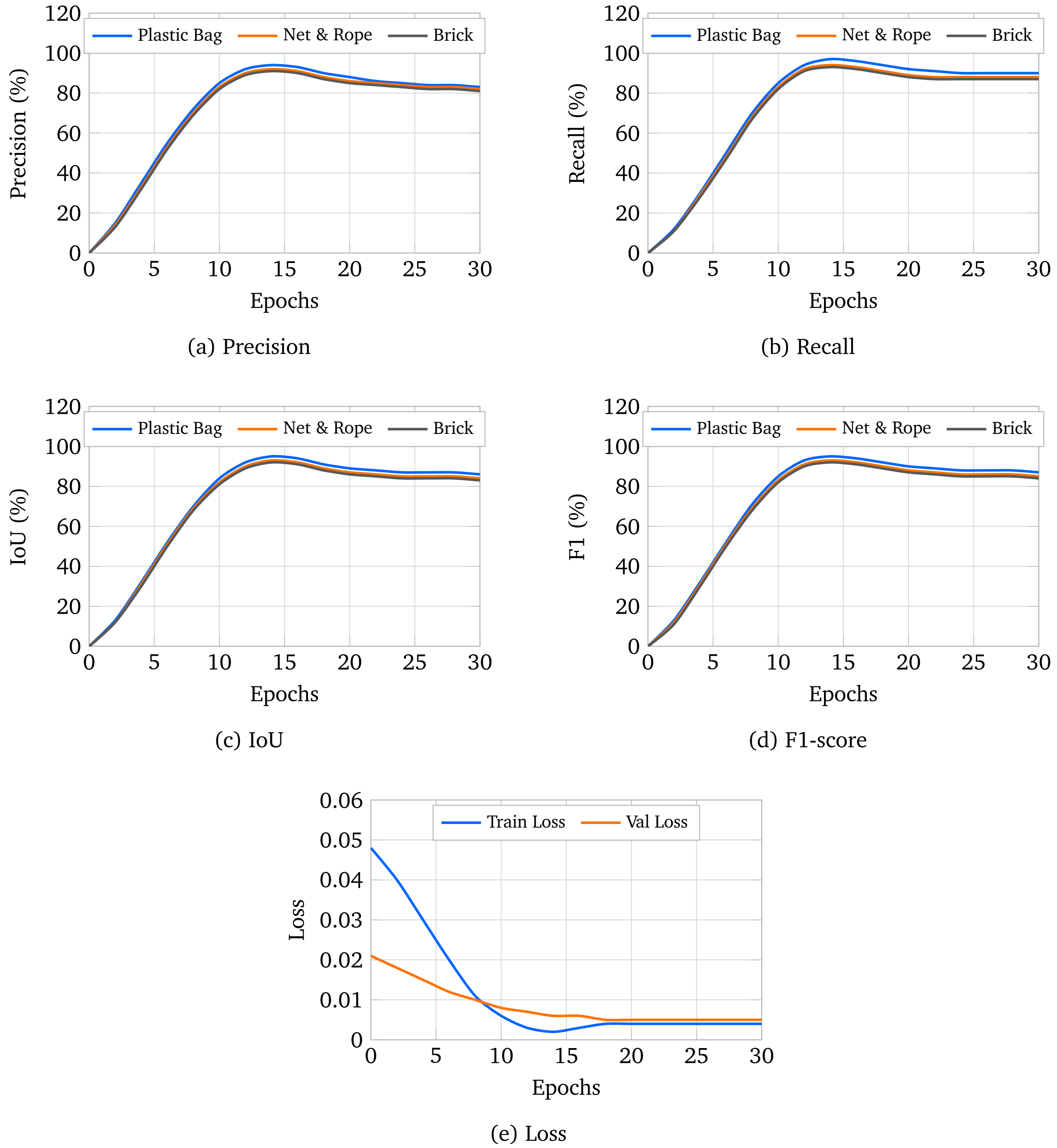


Figure 11: Performance evaluation of RCNN.

proposed PCA-GLCM-Fast-R-CNN model superior to its individual variants.

The performance comparison results in table 2 illustrate that, compared to other models, the proposed PCA-GLCM-Fast-R-CNN model performs better in detecting and classifying different classes of input data. Unlike standalone YOLO and other deep learning models, combining YOLO-based detection with a PCA-GLCM-Fast-R-CNN classifier yields superior results on the evaluation dataset. The proposed YOLO model accurately detects candidate objects, and the PCA-GLCM-Fast-R-CNN model precisely classifies the detected regions with 98.2% accuracy, 97.8% precision, 98.0% recall, and an F1-score of 97.9%. This result demonstrates that analysing real-time data from multiple sensors is more effective for detecting both normal and hazardous waste on any water surface.

Table 1
Statistical analysis.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
FRCNN	94.6	94.1	93.8	93.9
GLCM+FRCNN	96.8	96.2	96.5	96.3
PCA+FRCNN	97.2	97.0	97.1	97.1
PCA+GLCM+FRCNN	98.2	97.8	98.0	97.9

Table 2
Performance comparison.

Reference / model	Technique used	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Zhang et al. [23]	YOLO–Transfer learning	92.23	90.1	88.7	89.4
Nunkhaw, Chitwatkulsiri and Miyamoto [15]	YOLOv5 + DeepSORT	88.0	85.3	86.4	85.8
Vijayanti et al. [20]	VGG19–ResNet50 hybrid	94.5	93.7	91.2	92.4
Kundu, Sharma and Pillai [14]	ResNet50 transfer learning	95.6	94.9	93.8	94.3
He et al. [9]	EC–YOLOX	96.4	95.2	94.6	94.9
Proposed PCA–Fast-R-CNN (2025)	PCA + Fast R-CNN (IoT-integrated)	98.2	97.8	98.0	97.9

7. Conclusion

This paper presented a PCA-GLCM-Fast-R-CNN framework integrated with an IoT sensing infrastructure for real-time detection and classification of hazardous floating waste in water bodies. PCA reduces feature dimensionality and suppresses sensor noise; GLCM extracts discriminative textural features; and Fast R-CNN performs accurate region-based classification on the candidate regions proposed by YOLOv8. The combined PCA-GLCM-Fast-R-CNN model achieves 98.2% accuracy, 97.8% precision, 98.0% recall, and 97.9% F1-score, outperforming all ablation configurations and the published baselines listed in Table 2. The ablation study confirms that each preprocessing component contributes independently: GLCM raises accuracy from 94.6% to 96.8% by capturing textural patterns, PCA further increases it to 97.2% through dimensionality reduction, and combining both yields the highest performance.

From a theoretical perspective, the work contributes a modular deep learning architecture that pairs complementary feature extraction strategies with region-based classification for dynamic water environments. From a management perspective, the IoT-based design is scalable, supporting real-time environmental decisions and efficient resource allocation. In practice, the system enables authorities to detect and respond to floating waste hazards proactively, contributing to cleaner water resources and advancing the SDG 6 (Clean Water and Sanitation) and SDG 14 (Life Below Water) objectives.

Future research directions include exploring generative adversarial networks and reinforcement learning to improve detection robustness in unseen conditions. Integrating the framework with global monitoring networks, in collaboration with environmental organisations and governments, could support large-scale cross-border waste surveillance. Collecting more diverse datasets across different geographical locations, weather conditions, and waste types would improve model generalisation. Hardware improvements in IoT sensor energy efficiency and further advances in edge computing would enable low-latency processing across large-scale water bodies with multiple monitoring points.

Coupling the system with autonomous removal robots or drones would ultimately transform it from a detection platform into an end-to-end waste management solution.

Author contributions

All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data availability statement

The image dataset used in this study is a publicly available collection of floating waste samples [11]. No proprietary data were created by the authors.

Conflicts of interest

The authors declare no conflict of interest.

Declaration on Generative AI

The authors have not employed any generative AI tools.

References

[1] Akib, A., Tasnim, F., Biswas, D., Hashem, M.B., Rahman, K., Bhattacharjee, A. and Fattah, S.A., 2019. Unmanned Floating Waste Collecting Robot. *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*. pp.2645–2650. Available from: https://doi.org/10.1109/TENCON.2019.8929537.

[2] Cai, Z. and Vasconcelos, N., 2018. Cascade R-CNN: Delving Into High Quality Object Detection. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp.6154–6162. Available from: https://doi.org/10.1109/CVPR.2018.00644.

[3] Codes-Alcaraz, A.M., Puerto, H. and Rocamora, C., 2024. Image Recognition for Floating Waste Monitoring in a Traditional Surface Irrigation System. *Water*, 16(18), p.2680. Available from: https://doi.org/10.3390/w16182680.

[4] Dalu, T., Banda, T., Mutshekwa, T., Munyai, L. and Cuthbert, R., 2021. Effects of urbanisation and a wastewater treatment plant on microplastic densities along a subtropical river system. *Environmental Science and Pollution Research*, 28, pp.36102–36111. Available from: https://doi.org/10.1007/s11356-021-13185-1.

[5] Girshick, R., Donahue, J., Darrell, T. and Malik, J., 2014. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*. pp.580–587. Available from: https://doi.org/10.1109/CVPR.2014.81.

[6] Hafeez, S., Wong, M.S., Abbas, S., Kwok, C.Y.T., Nichol, J., Lee, K.H., Tang, D. and Pun, L., 2018. Detection and Monitoring of Marine Pollution Using Remote Sensing Technologies. In: H.B. Fouzia, ed. *Monitoring of Marine Pollution*. London: IntechOpen, chap. 2. Available from: https://doi.org/10.5772/intechopen.81657.

[7] Harris, P.T., Westerveld, L., Nyberg, B., Maes, T., Macmillan-Lawler, M. and Appelquist, L.R., 2021. Exposure of coastal environments to river-sourced plastic pollution. *Science of The Total Environment*, 769, p.145222. Available from: https://doi.org/10.1016/j.scitotenv.2021.145222.

[8] Hasan, M.D.A., Balasubadra, K., Vadivel, G., Arunfred, N., Ishwarya, M. and Murugan, S., 2024. IoT-Driven Image Recognition for Microplastic Analysis in Water Systems using Convolutional Neural Networks. *2024 2nd International Conference on Computer, Communication and Control (IC4)*. pp.1–6. Available from: https://doi.org/10.1109/IC457434.2024.10486490.

[9] He, J., Cheng, Y., Wang, W., Gu, Y., Wang, Y., Zhang, W., Shankar, A., Selvarajan, S. and Kumar, S.A.P., 2024. EC-YOLOX: A Deep-Learning Algorithm for Floating Objects Detection in Ground Images of Complex Water Environments. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, pp.7359–7370. Available from: https://doi.org/10.1109/JSTARS.2024.3367713.

[10] He, K., Zhang, X., Ren, S. and Sun, J., 2015. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), pp.1904–1916. Available from: https://doi.org/10.1109/TPAMI.2015.2389824.

[11] Hong, J., Fulton, M.S. and Sattar, J., 2020. TrashCan 1.0 An Instance-Segmentation Labeled Dataset of Trash Observations. Available from: https://doi.org/10.13020/g1gx-y834.

[12] Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R. and Law, K.L., 2015. Plastic waste inputs from land into the ocean. *Science*, 347(6223), pp.768–771. Available from: https://doi.org/10.1126/science.1260352.

[13] Junzhe, Z., Fuqiang, J., Yupeng, C., Weiyi, W. and Qing, W., 2023. A water surface garbage recognition method based on transfer learning and image enhancement. *Results in Engineering*, 19, p.101340. Available from: https://doi.org/10.1016/j.rineng.2023.101340.

[14] Kundu, S., Sharma, M. and Pillai, A.S., 2024. AI-Powered Trash Classification System for Lakes and Water Bodies Using Transfer Learning. *2024 Third International Conference on Power, Control and Computing Technologies (ICPC2T)*. pp.163–167. Available from: https://doi.org/10.1109/ICPC2T60072.2024.10474611.

[15] Nunkhaw, M., Chitwatkulsiri, D. and Miyamoto, H., 2025. Enhancing River Waste Detection with Deep Learning and Preprocessing: A Case Study in the Urban Canals of the Chao Phraya River. *Water*, 17(22), p.3193. Available from: https://doi.org/10.3390/w17223193.

[16] Nunkhaw, M. and Miyamoto, H., 2024. An Image Analysis of River-Floating Waste Materials by Using Deep Learning Techniques. *Water*, 16(10), p.1373. Available from: https://doi.org/10.3390/w16101373.

[17] Shirsat, N. and Nirmalrani, V., 2024. Automated System for Detection of Floating Water Pollutants using Deep Learning Framework Metric for Sustainable Life. *Grenze International Journal of Engineering and Technology*, 10(2), pp.1784–1789. Available from: https://thegrenze.com/index.php?display=page&view=journalabstract&absid=2878&id=8.

[18] Smith, M., Love, D., Rochman, C. and Neff, R., 2018. Microplastics in Seafood and the Implications for Human Health. *Current Environmental Health Reports*, 5, pp.375–286. Available from: https://doi.org/10.1007/s40572-018-0206-z.

[19] Tian, Z., Huang, J., Yang, Y. and Nie, W., 2023. KCFS-YOLOv5: A High-Precision Detection Method for Object Detection in Aerial Remote Sensing Images. *Applied Sciences*, 13(1), p.649. Available from: https://doi.org/10.3390/app13010649.

[20] Vijayanti, V., Kar, A., Navya, K.N.S.S. and Siva, S.S., 2023. Analysis of Deep Learning Based Garbage Detection in Water Bodies. *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*. pp.1–8. Available from: https://doi.org/10.1109/ICIRCA57980.2023.10220828.

[21] Wong, W.Y., Al-Ani, A.K.I., Hasikin, K., Khairuddin, A.S.M., Razak, S.A., Hizaddin, H.F., Mokhtar, M.I. and Azizan, M.M., 2021. Water, Soil and Air Pollutants' Interaction on Mangrove Ecosystem and Corresponding Artificial Intelligence Techniques Used in Decision Support Systems - A Review. *IEEE Access*, 9, pp.105532–105563. Available from: https://doi.org/10.1109/ACCESS.2021.3099107.

[22] Yang, J., Li, Z., Gu, Z. and Li, W., 2024. Research on floating object classification algorithm based on convolutional neural network. *Scientific Reports*, 14, p.32086. Available from: https://doi.org/10.1038/s41598-024-83543-9.

[23] Zhang, Q., Yang, Q., Zhang, X., Wei, W., Bao, Q., Su, J. and Liu, X., 2022. A multi-label waste detection model based on transfer learning. *Resources, Conservation and Recycling*, 181, p.106235. Available from: https://doi.org/10.1016/j.resconrec.2022.106235.