International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

Marine Debris Detection System Using Yolov11

Ms. Imreena Ali[1], Faisal Al Farooqui[2]
[1]Assistant Professor; Department Of Computer Science & Engineering ISL Engineering College Hyderabad India
[2]P.G Scholar Department Of Computer Science & Engineering ISL Engineering College Hyderabad India
Mail Id; imreena.cse@gmail.com[1],faisalalfarooqui@gmail.com[2]
Article Received 24-01-2026, Revised 27-2-2026, Accepted 04-04-2026

Abstract
Marine debris has become a major environmental concern that threatens aquatic ecosystems and coastal sustainability. Efficient and accurate detection of marine waste is therefore essential for effective monitoring and cleanup efforts. This study proposes an enhanced object detection approach that integrates attention mechanisms with the YOLOv11 architecture to improve the identification of marine debris in complex ocean environments.The proposed framework combines the strong instance detection capability of YOLOv11 with attention modules, including Coordinate Attention and a Bottleneck Transformer. While YOLOv11 with Coordinate Attention demonstrates consistent performance across diverse environmental conditions, the Bottleneck Transformer contributes to identifying debris regions that may be overlooked during manual annotation. Although slightly less stable in some scenarios, the transformer-based approach shows improved performance in detecting larger debris objects, indicating its potential usefulness for specialized monitoring tasks.Experimental results demonstrate that incorporating attention mechanisms into YOLOv11 enhances detection capability and adaptability in maritime environments. The findings suggest that attention-based object detection models can effectively support environmental monitoring systems by addressing different operational requirements. Furthermore, the study highlights the importance of selecting detection architectures according to specific deployment conditions, as different models may provide advantages depending on the characteristics of the debris and the monitoring environment.

Keywords
Marine Debris Detection; YOLOv11; Attention Mechanisms; Coordinate Attention; Bottleneck Transformer; Environmental Monitoring; Deep Learning; Computer Vision.

Introduction

Marine debris has emerged as a significant environmental issue that threatens both terrestrial and aquatic ecosystems. The accumulation of waste materials such as plastics, fishing gear, and other discarded objects in marine environments negatively affects biodiversity, water quality, and coastal sustainability. Effective monitoring and detection of such debris are essential for minimizing ecological damage and supporting large-scale cleanup initiatives. However, traditional detection methods often struggle with limitations related to accuracy, efficiency, and scalability, particularly in dynamic marine environments where lighting conditions, reflections, and complex backgrounds make object identification challenging.Recent advances in computer vision and deep learning have provided promising solutions for automated marine debris detection. Object detection models based on deep neural networks have demonstrated strong capabilities in identifying objects in complex scenes. Among these approaches, the YOLO (You Only Look Once) family of models has gained considerable attention due to its ability to perform real-time object detection with high accuracy. The latest version, YOLOv11, introduces improvements in feature extraction, detection accuracy, and computational efficiency, making it suitable for environmental monitoring applications.In this study, a novel detection framework is proposed by integrating YOLOv11 with advanced attention mechanisms, specifically Coordinate Attention and the Bottleneck Transformer. Attention mechanisms allow neural networks to focus on the most relevant spatial and contextual features within an image, thereby improving object detection performance. Coordinate Attention enhances the model's ability to capture spatial relationships and channel dependencies, enabling more accurate localization of debris objects within marine scenes. Meanwhile, the Bottleneck Transformer introduces global contextual awareness by combining convolutional operations with transformer-based attention mechanisms.The proposed approach evaluates the effectiveness of these attention-enhanced models in detecting marine debris under varying environmental conditions. Experimental analysis indicates that YOLOv11 integrated with Coordinate Attention delivers stable and reliable detection performance across diverse scenarios. On the other hand, the Bottleneck Transformer demonstrates particular strengths in identifying debris regions that may be overlooked during manual annotation and shows improved accuracy when detecting larger debris objects. Although its performance may exhibit slightly lower consistency in some cases, its capability to capture global contextual information provides valuable advantages in specific monitoring tasks.Overall, the findings of this study highlight the potential of attention-augmented YOLOv11 models for automated marine debris detection. The results also emphasize the importance of selecting detection architectures based on the specific requirements of environmental monitoring systems, as different model configurations may offer distinct advantages depending on operational conditions and debris characteristics.

Literature Survey

A literature survey represents a systematic examination of previously published research, scholarly articles, and technical reports related to a specific topic. The main objective of conducting a literature review is to understand the current state of knowledge in the field, including existing methodologies, theoretical concepts, experimental findings, and technological developments. By critically analyzing earlier studies, researchers can identify strengths, limitations, and research gaps that require further investigation. A comprehensive literature survey also prevents unnecessary duplication of work and provides a strong conceptual framework for the proposed study. In addition, it helps researchers define research objectives, select appropriate methodologies, and develop hypotheses that align with existing scientific knowledge. Therefore, the literature review forms an essential foundation for any research work by contextualizing the problem and justifying the importance of the proposed approach.

Review of Literature

Recent advancements in computer vision and deep learning have significantly improved object detection systems. One important contribution in this area is the work by H. Liu et al. (2023) on Visual Instruction Tuning (VIT). The authors introduced a training strategy designed to improve the ability of visual models to understand complex instructions associated with images. Traditional vision models are typically trained using large datasets without task-specific guidance, which can limit their effectiveness in specialized tasks. The proposed VIT framework integrates task-oriented visual cues into the training process, allowing the model to better interpret contextual information. Experimental results demonstrate improvements in tasks such as object detection, image captioning, and semantic segmentation. The study also highlights the role of visual instruction tuning in multimodal systems that combine textual and visual information for intelligent decision making.Earlier work by R. Girshick et al. (2014) introduced the concept of rich feature hierarchies through the R-CNN architecture for object detection and semantic segmentation. Traditional detection techniques struggled to achieve high accuracy due to challenges such as background clutter and variations in object scale. The proposed approach used deep convolutional neural networks (CNNs) to extract hierarchical feature representations from images. By applying region proposal techniques such as selective search, R-CNN generated candidate object regions which were subsequently classified using CNN features. This method significantly improved detection accuracy and established deep learning as a powerful tool for visual recognition tasks.Building upon the limitations of R-CNN, R. Girshick (2015) proposed the Fast R-CNN model to improve both computational efficiency and detection performance. The key innovation in Fast R-CNN was the introduction of the Region of Interest (RoI) pooling layer, which allowed feature extraction to be performed once for the entire image instead of repeatedly for each region proposal. This modification reduced redundant computation and significantly accelerated both training and inference processes. Fast R-CNN also improved localization accuracy by jointly optimizing classification and bounding box regression within a unified architecture.Further improvements were achieved with the development of Faster R-CNN by S. Ren et al., which introduced Region Proposal Networks (RPNs). In this architecture, region proposals are generated directly from the convolutional feature maps instead of relying on external algorithms such as selective search. This integration created a unified, end-to-end trainable system that improved both speed and detection accuracy. The model demonstrated strong performance on several benchmark datasets and represented an important milestone in the evolution of deep learning-based object detection frameworks.Another significant advancement in real-time detection was presented by J. Redmon et al. with the introduction of the YOLO (You Only Look Once) model. Unlike two-stage detectors such as R-CNN variants, YOLO reformulated object detection as a single regression problem. The model processes the entire image using a single convolutional network and predicts bounding boxes and class probabilities simultaneously. This unified approach allows YOLO to perform object detection extremely quickly while maintaining competitive accuracy. Due to its speed and efficiency, YOLO became widely used in applications such as autonomous driving, surveillance, and robotics.Similarly, W. Liu et al. (2016) proposed the Single Shot Multibox Detector (SSD), another single-stage object detection framework designed to balance detection speed and accuracy. SSD uses multiple feature maps from different layers of the network to predict bounding boxes at various scales. This multi-scale detection capability enables the model to recognize objects with different sizes and aspect ratios within the same image. By optimizing both classification and localization tasks simultaneously, SSD achieved strong performance in real-time detection scenarios and became widely used in industrial and research applications.

Comparative Analysis of Literature

An analysis of previous studies indicates that object detection techniques have evolved significantly over time, moving from multi-stage detection systems toward unified deep learning architectures. Early approaches such as R-CNN achieved high detection accuracy but suffered from slow processing speeds and high computational costs. Fast R-CNN addressed some of these limitations by introducing RoI pooling and shared feature extraction, while Faster R-CNN further improved efficiency through the integration of Region Proposal Networks. Although these models provided high detection accuracy, they remained computationally intensive for real-time applications.Single-stage detectors such as YOLO and SSD introduced a more efficient approach by predicting object locations and categories in a single network pass. These models offer significantly faster inference speeds, making them suitable for real-time applications. However, early YOLO models struggled with small object detection and localization precision in crowded scenes. Recent advancements, including improved feature pyramids and attention mechanisms, aim to overcome these challenges by enhancing feature representation and contextual understanding.

Project Description And Requirements

This research focuses on the development of an automated system for detecting and locating marine debris using advanced deep learning techniques. Marine pollution caused by floating waste such as plastic bags, bottles, fishing nets, and discarded materials poses a major threat to marine ecosystems and biodiversity. Early identification and monitoring of such debris are essential for effective environmental protection and cleanup operations. To address this issue, the proposed system utilizes the YOLOv11 object detection model combined with attention mechanisms such as Coordinate Attention and Bottleneck Transformers.The integration of attention modules enables the detection model to focus on relevant regions within an image while suppressing background noise. This capability is particularly important in marine environments where visual complexity arises from water reflections, waves, shadows, and changing lighting conditions. High-resolution images obtained from drones, satellites, or underwater cameras are used to train the model to recognize different types of marine debris. The proposed framework is evaluated under challenging conditions including varying object sizes, occlusions, and environmental variability. Special attention is given to detecting small debris particles, which are common in real marine environments but difficult to identify using conventional detection techniques.

Methodology

The proposed system is organized into several functional modules that collectively enable the detection and visualization of marine debris. The first component is the input data module, which handles image datasets collected from multiple sources including drones, underwater cameras, and satellite imagery. These images often contain varying lighting conditions, reflections, and environmental disturbances, which must be carefully managed during processing.The preprocessing module improves image quality and prepares the data for analysis. Techniques such as adaptive histogram equalization are applied to enhance contrast and highlight debris objects against complex ocean backgrounds. Images are also resized and standardized to maintain consistency across different datasets and input formats.The segmentation module separates debris objects from the surrounding environment. In this stage, clustering techniques such as Super-Pixel-Based Fast Fuzzy C-Means are applied to group similar pixels and isolate potential debris regions from water surfaces and natural elements. This segmentation step improves the accuracy of subsequent detection stages.Next, the feature extraction module analyzes the segmented regions to identify meaningful visual characteristics such as edges, shapes, and textures. Convolutional neural networks combined with transformer-based attention mechanisms are used to capture both local features and global contextual information. These extracted features enable the model to distinguish marine debris from surrounding objects and natural elements.The YOLOv11 detection module then processes the extracted features to perform object detection and classification. YOLOv11 predicts bounding boxes around detected objects, assigns class labels, and calculates confidence scores indicating the reliability of each detection. Finally, the output visualization module displays the detection results using labeled bounding boxes, enabling users to monitor debris locations in images or video streams.

System Requirements

The proposed system requires both hardware and software resources to support model training, data processing, and real-time detection tasks. Hardware requirements include a system equipped with a dual-core processor, a minimum of 4 GB RAM, and

International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

sufficient storage capacity for dataset management. For improved performance during deep learning training, systems with GPU acceleration are recommended.The software environment includes the Python programming language along with development platforms such as Spyder or Jupyter Notebook. The system is compatible with Windows operating systems and uses deep learning libraries such as PyTorch for model training and implementation.

DESIGN ENGINEERING

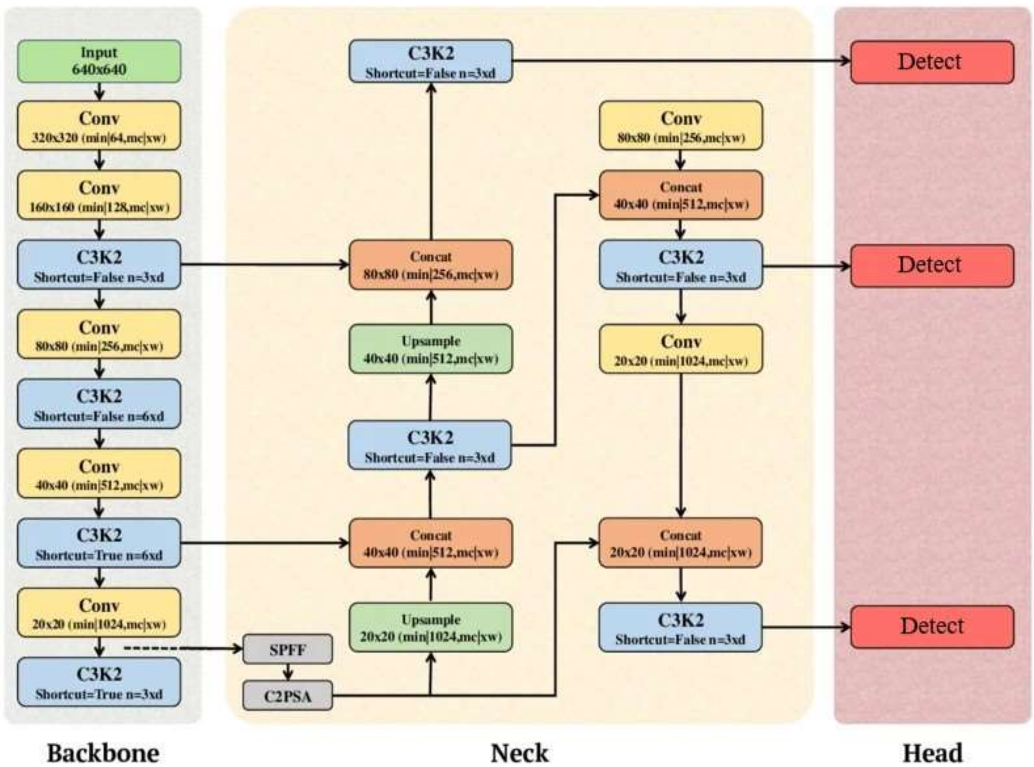

Fig 1 YOLOv11 ARCHITECTURE

Design engineering focuses on transforming system requirements into a structured software architecture that enables efficient implementation. Unified Modeling Language (UML) diagrams are used to visually represent the components, interactions, and workflow of the proposed system. These diagrams provide a conceptual representation of how the system operates and how different modules interact with each other.The system design includes several UML diagrams such as use case diagrams, class diagrams, object diagrams, state diagrams, activity diagrams, and sequence diagrams. The use case diagram illustrates the interactions between the user and the system, highlighting the major functionalities such as image input, debris detection, and result visualization. Class diagrams represent the structural relationships between system components, including data processing modules and detection algorithms.Sequence diagrams illustrate the order of interactions between system components during the detection process, while activity diagrams describe the step-by-step workflow of the system. Additional diagrams such as component diagrams and deployment diagrams provide insight into the software structure and hardware configuration required for system implementation.The YOLOv11 architecture serves as the central component of the system design. The detection process begins with image acquisition and preprocessing, followed by feature extraction using deep neural networks. The image is divided into grid cells, and each cell predicts bounding boxes, class probabilities, and objectness scores. Predictions with low confidence are removed, and Non-Maximum Suppression is applied to eliminate duplicate detections. The final output displays detected objects with labeled bounding boxes and associated confidence scores, enabling accurate identification and monitoring of marine debris.

Development Tools
Python
Python is a high-level interpreted programming language widely adopted in fields such as artificial intelligence, machine learning, computer vision, and scientific computing. Its simple syntax and

readability make it highly suitable for rapid development, experimentation, and prototyping of complex systems. Python supports multiple programming paradigms, including procedural, object-oriented, and functional programming, enabling developers to design scalable and flexible applications.One of the major advantages of Python is its extensive ecosystem of libraries and frameworks that facilitate advanced data processing and deep learning. Libraries such as NumPy support efficient numerical computations, while OpenCV enables image processing and computer vision operations. Visualization libraries like Matplotlib allow researchers to analyze model performance through graphs and plots. Deep learning frameworks such as PyTorch and TensorFlow provide powerful tools for building, training, and deploying neural network models.In this project, Python serves as the core language used to implement the marine debris detection system. It is responsible for managing datasets, performing image preprocessing operations, handling annotations, and converting raw data into formats suitable for training the detection model. Python scripts are also used to train the YOLOv11 model, adjust hyperparameters, and evaluate performance metrics such as accuracy, precision, and recall. In addition, Python enables real-time image processing and visualization of detection results, including the display of bounding boxes and predicted object classes.

Anaconda

Anaconda is a comprehensive Python distribution designed specifically for data science, machine learning, and scientific computing applications. It simplifies the installation and management of packages through the Conda package manager, which allows users to install libraries and maintain project dependencies efficiently. One of the most important advantages of Anaconda is its ability to create isolated virtual environments, ensuring that different projects can run with their required versions of software libraries without causing compatibility conflicts.In the context of this research, Anaconda provides a stable development environment for training and evaluating the YOLOv11 model. It allows the creation of dedicated environments containing specific versions of libraries such as PyTorch, OpenCV, and other deep learning dependencies. This isolation ensures that the marine debris detection system operates reliably and reduces potential issues related to library version mismatches during development and experimentation.

Jupyter Notebook

Jupyter Notebook is an open-source interactive computing platform widely used for data analysis, machine learning experiments, and scientific research. It provides a web-based interface where code execution, explanatory text, and visual outputs can be integrated within a single document-like environment. This interactive structure allows researchers to test algorithms, visualize intermediate results, and document experimental observations simultaneously.In this project, Jupyter Notebook is used as the primary environment for developing and testing the marine debris detection system. It enables visualization of dataset samples, verification of annotations, and experimentation with preprocessing techniques before model training. During training, performance metrics such as loss curves and accuracy values can be monitored in real time, allowing researchers to observe how the model learns over successive training epochs. Additionally, Jupyter Notebook facilitates qualitative evaluation of the model by displaying detection outputs such as predicted bounding boxes and class labels, helping researchers analyze the system's performance on sample images.

Implementation

The implementation phase involves developing the marine debris detection system using the YOLOv11 deep learning architecture. The system is designed to perform both training and testing operations using labeled datasets containing marine debris images. During the training stage, the dataset is first loaded and divided into training, validation, and testing subsets to ensure proper model evaluation. The YOLOv11 model is initialized using the Ultralytics framework and trained iteratively over multiple epochs. In each epoch, batches of images from the training dataset are processed to update the model parameters. After each training cycle, validation is performed to assess model performance and detect potential overfitting. The best-performing model weights are stored for later use during inference.During the testing stage, the trained model is loaded and applied to the test dataset to evaluate its predictive capabilities. For each image in the test set, the model generates predictions that include object locations, class labels, and confidence scores. These predictions are compared with ground truth annotations to calculate evaluation metrics such as precision, recall, F1-score, and mean Average Precision (mAP). These metrics provide quantitative insights into the effectiveness of the trained model.In addition to dataset testing, the system also supports image-based detection through an interactive pipeline. When an input image is provided, it is first preprocessed by resizing and normalizing pixel values to match the model's input requirements. The processed image is then passed through the trained YOLOv11 model, which performs object detection by identifying bounding boxes, class labels, and confidence scores. Finally, the detection results are visualized by displaying the image with labeled bounding boxes around the detected debris objects.

International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

Results And Performance Analysis

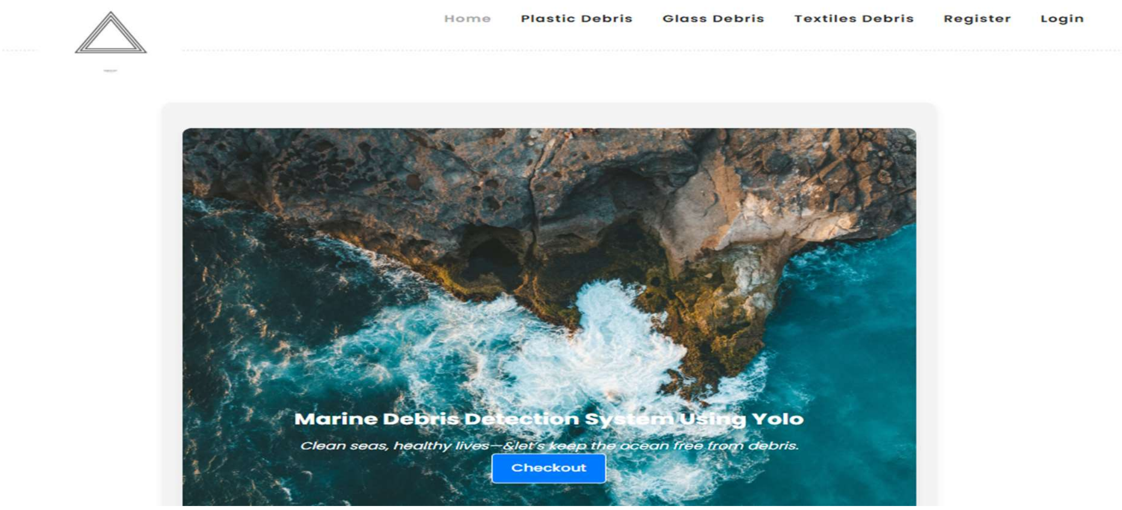

Fig: 2 Output Snapshot

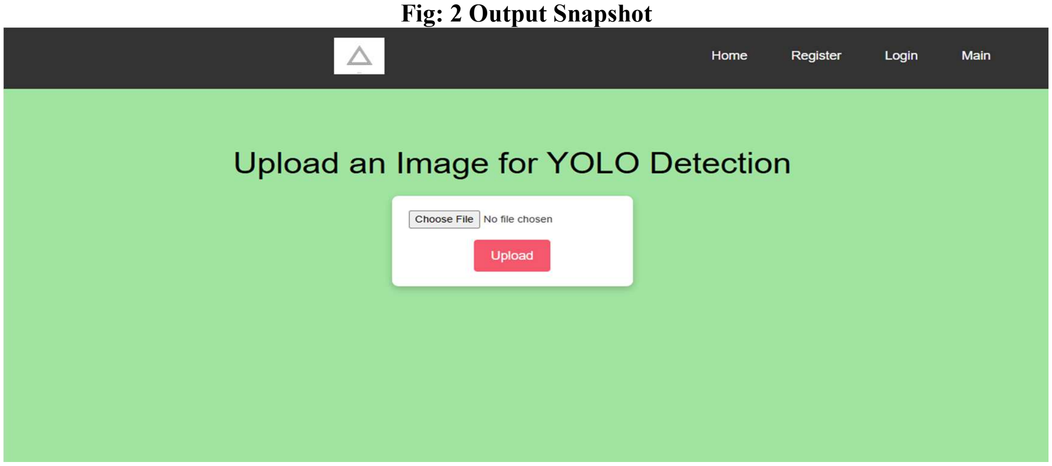

Fig: 3 Output Snapshot

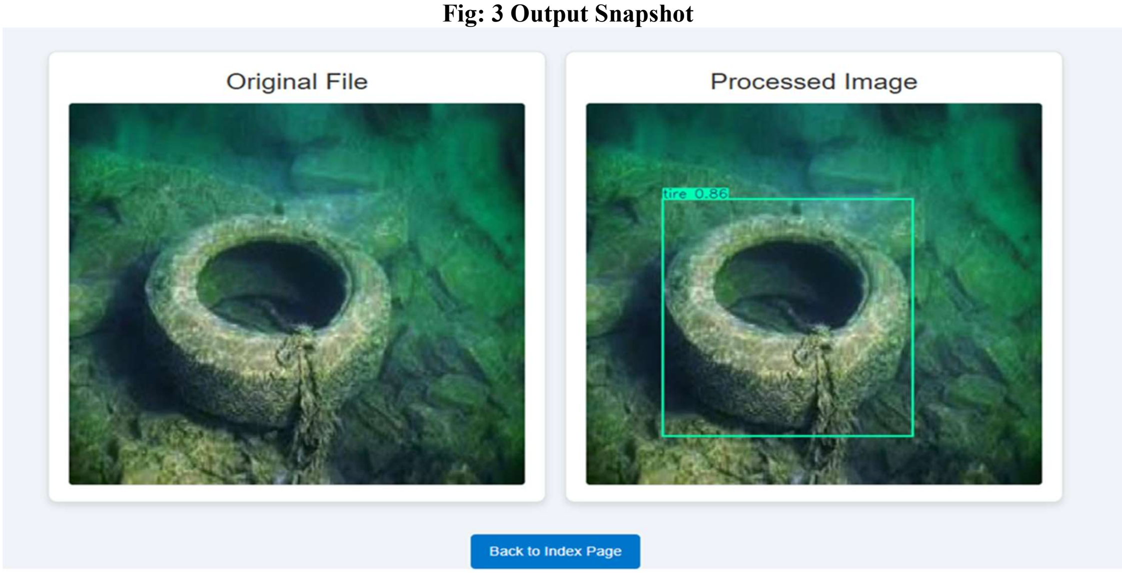

Fig: 4 Output Snapshot

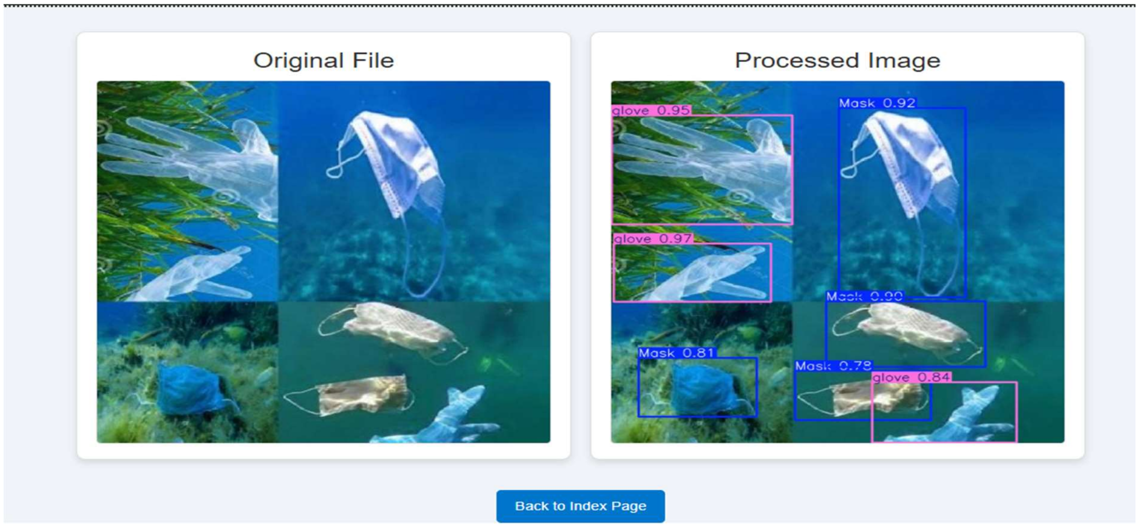

Fig: 5 Output Snapshot

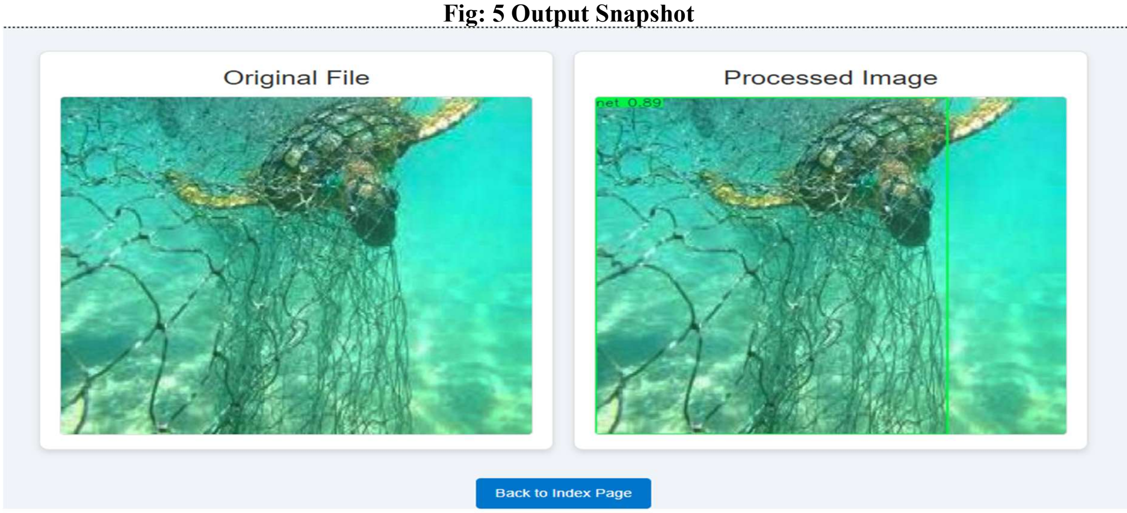

Fig: 6 Output Snapshot

Fig: 7 Output Snapshot

International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

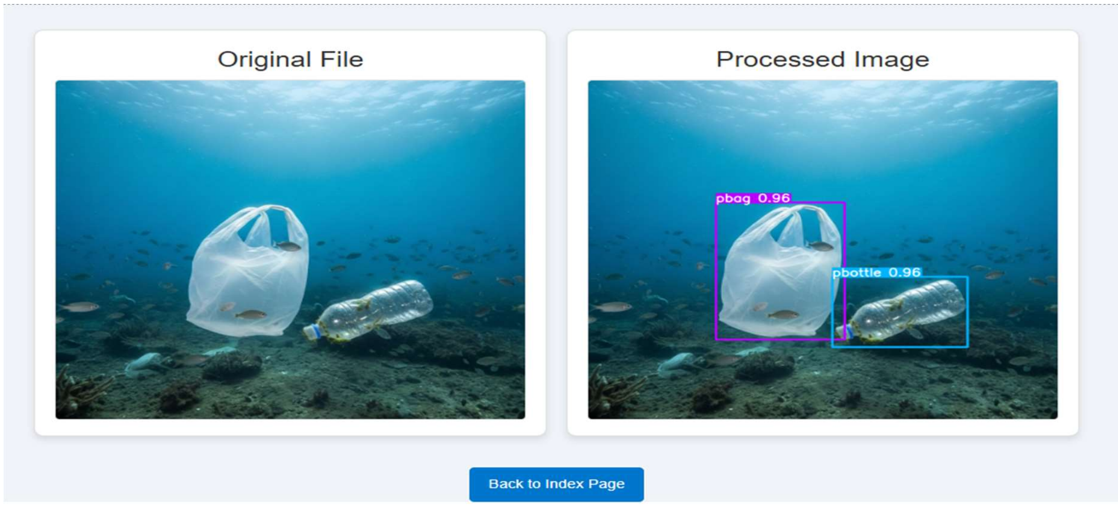

Fig: 8 Output Snapshot

The performance of the proposed marine debris detection system was evaluated using a labeled dataset containing multiple categories of marine waste objects. The dataset included classes such as masks, cans, cellphones, electronic waste, gloves, miscellaneous debris, fishing nets, plastic bags, plastic bottles, tires, and background regions. A confusion matrix was generated to analyze the classification performance of the trained YOLOv11 model across these categories.From the evaluation results, the model achieved strong detection performance for several object categories, particularly for plastic bottles, tires, and fishing nets. These classes demonstrated high precision and recall values, indicating that the model was able to accurately detect and classify these objects with minimal misclassification. For example, the tire class achieved high precision and recall values, demonstrating the model's capability to detect large debris objects effectively. Similarly, the plastic bottle and plastic bag categories showed strong detection performance due to their distinctive visual features.However, some classes such as cans and miscellaneous debris exhibited lower detection accuracy. These categories often include objects with irregular shapes or low visual contrast against the surrounding environment, which makes them more difficult to detect. Small objects also present challenges in aerial or underwater imagery due to their limited pixel representation. Despite these challenges, the overall model performance remained stable across most categories.

Based on the confusion matrix calculations, the system achieved an overall accuracy of approximately 72.4%, with precision and recall values close to 0.72 and 0.73, respectively. The resulting F1-score of approximately 0.726 indicates a balanced performance between precision and recall. These results demonstrate that the proposed YOLOv11-based detection framework can effectively identify marine debris in complex environmental conditions.

Software Testing

Testing Methodology

The testing methodology adopted in this project involves evaluating both individual modules and the integrated system to ensure compliance with the specified requirements. A structured testing strategy was developed to assess system performance across different operating conditions and datasets. Quality assurance procedures were applied to verify that the system operates without errors and satisfies the functional specifications described in the system requirements document. The testing framework focuses on validating the correctness, efficiency, and reliability of the detection system.

Types of Testing

Several types of testing were conducted to ensure the robustness of the system. Unit testing was performed to verify the functionality of individual modules within the application. Each component was tested independently to ensure that it produced correct outputs for given inputs and followed the intended program logic.Functional testing was carried out to confirm that all system functions operate according to the defined requirements. This included verifying that valid inputs are accepted by the system while invalid inputs are appropriately rejected. The testing also ensured that detection outputs such as bounding boxes and class labels were generated correctly.**System** testing evaluated the performance of the complete integrated system. This testing stage confirmed that all modules interact correctly and

International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

that the overall workflow—from image input to detection output—operates smoothly without errors.**Performance testing** was conducted to measure the system's response time and processing speed. This evaluation ensured that the model could generate detection results within an acceptable time frame, which is particularly important for real-time monitoring applications.Integration testing was performed to verify that different system components function together without compatibility issues. This testing focused on the interaction between preprocessing modules, the YOLOv11 detection model, and the visualization module.

Future Enhancements

Although the proposed attention-enhanced YOLOv11 framework demonstrates promising performance in marine debris detection, several opportunities exist to further enhance its effectiveness and applicability. One important direction for future research involves the integration of multimodal sensing technologies. By combining visual imagery with additional data sources such as multispectral, thermal, or sonar information, detection models could achieve greater robustness under challenging environmental conditions. Marine environments often include scenarios with poor visibility, turbid water, or complex backgrounds where traditional RGB imagery may not provide sufficient information. Incorporating complementary sensing modalities could enable the model to capture additional features and improve detection reliability across diverse environmental conditions.Another promising research direction involves optimizing the detection model for deployment on edge computing platforms. Implementing lightweight versions of the YOLOv11 architecture on embedded systems, drones, or autonomous underwater vehicles would enable real-time monitoring of marine debris in remote locations. Such deployment would reduce the dependence on centralized computing infrastructure and enable continuous environmental monitoring in large oceanic regions. Edge-based processing could also facilitate rapid response during cleanup operations by providing immediate detection results.Future work may also explore advanced training strategies to improve the adaptability of detection models. Techniques such as self-supervised learning, domain adaptation, and continual learning can allow models to adapt to new environmental conditions without requiring extensive manual labeling. These methods could improve model performance under changing lighting conditions, seasonal variations in debris distribution, and different marine ecosystems.Furthermore, expanding the training dataset to include images collected from diverse geographical locations and ecological environments would significantly improve model generalization. A larger and more diverse dataset would allow the model to learn a wider range of visual patterns associated with marine debris, thereby improving its ability to detect objects across different ocean regions and environmental scenarios. Collectively, these future enhancements could significantly strengthen the performance, scalability, and real-world applicability of marine debris detection systems.

Conclusion

This research presents a deep learning–based framework for automated marine debris detection using the YOLOv11 object detection architecture enhanced with attention mechanisms. Marine pollution caused by floating debris poses serious environmental risks to aquatic ecosystems, making efficient detection systems essential for monitoring and mitigation efforts. By integrating advanced attention modules such as Coordinate Attention and the Bottleneck Transformer, the proposed system improves the ability of the YOLOv11 model to identify debris objects in complex ocean environments.Experimental results demonstrate that the attention-enhanced YOLOv11 framework is capable of detecting multiple categories of marine debris with reliable accuracy. The integration of Coordinate Attention improves spatial feature representation, enabling the model to achieve stable performance across different environmental conditions. Meanwhile, the Bottleneck Transformer contributes to capturing broader contextual information within images, which helps identify larger debris objects and regions that may be overlooked during manual annotation. These complementary capabilities highlight the advantages of incorporating attention mechanisms into modern object detection frameworks.The evaluation results indicate that the proposed system can effectively identify various types of marine debris, including plastic bottles, fishing nets, and tires, even in visually complex environments characterized by reflections, waves, and background clutter. The obtained performance metrics demonstrate the potential of deep learning–based detection systems for supporting automated environmental monitoring and marine conservation initiatives.Overall, this study highlights the importance of attention-based object detection models in addressing real-world environmental challenges. Future research can further enhance system performance through improved datasets, multimodal sensing integration, and deployment on real-time monitoring platforms. Such advancements could play a significant role in large-scale marine pollution monitoring and contribute to the protection of global ocean ecosystems.

International Journal of Artificial Intelligence and Computer Electronics
Volume: 02, Issue: 02, April-June 2026, pp:1-10
P-ISSN 3117-8529 | E-ISSN 3117-8537

References

1. National Oceanic and Atmospheric Administration (NOAA), "Understanding Marine Debris and Its Environmental Impact." Available: https://marinedebris.noaa.gov
2. T. Maes et al., *From Pollution to Solution: A Global Assessment of Marine Litter and Plastic Pollution*, United Nations Environment Programme (UNEP), Nairobi, 2021.
3. L. Lebreton et al., "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic," *Scientific Reports*, vol. 8, 2018.
4. K. Song, J. Y. Jung, S. H. Lee, S. Park, and Y. Yang, "Assessment of marine debris in remote regions using unmanned aerial vehicles and deep learning segmentation models," *Sustainability*, vol. 14, no. 14, 2022.
5. A. Jamali and M. Mahdianpari, "A cloud-based framework for large-scale monitoring of ocean plastics using multispectral satellite imagery," *Water*, vol. 13, no. 18, 2021.
6. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
7. W. Liu et al., "SSD: Single Shot MultiBox Detector," *European Conference on Computer Vision (ECCV)*, 2016.
8. R. Girshick, "Fast R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015.
9. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.
10. A. Shah, L. Thomas, and M. Maskey, "Marine debris dataset for object detection in PlanetScope imagery," Radiant MLHub Dataset Repository, 2021.