International Journal of Scientific Research in Engineering and Management (IJSREM)
Special Edition - Volume 09 NCFT -2025 SJIF Rating: 8.586 ISSN: 2582-3930

A Refined Deep Learning Method Using Yolo for Marine Debris Identification

[1]D. Vivek, [2]G. Supriya, [3]G. Sri Chaitanya, [4] A. Sunitha
[1,2,3]UG Scholars,[4]Assistant Professor
[1,2,3,4]Department of Computer Science and Engineering
[1,2,3,4]Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India
**Corresponding author E-mail: supriyaguntuku661@gmail.com*

ABSTRACT

Environmental ecosystems are seriously threatened by marine trash, hence efficient techniques for locating and detecting it are required. By putting forth a novel strategy that combines the instance detection capabilities of YOLOv11 with different attention mechanisms to improve efficiency and expand application, this study overcomes the constraints currently present in the literature. While YOLOv11 and Coordinate Attention function well in a variety of contexts, the bottleneck transformer is useful in locating debris regions that are missed via manual annotation, despite being somewhat less consistent overall. Furthermore, the bottleneck transformer exhibits improved accuracy in identifying larger debris fragments, suggesting its possible use in particular applications. The adaptability and effectiveness of attention-enhanced YOLOv11 models for maritime trash identification are highlighted in this work, proving their capacity to handle a range of environmental monitoring difficulties. The findings also highlight how crucial it is to match detection models with certain operational needs because each model has special qualities that can be useful in particular situations.

Keywords: YOLOv11, Marine Debris, Coordinate Attention, Bottleneck Transformer, Object Detection, Deep Learning, Environmental Monitoring, Feature Extraction, Instance Segmentation, Drone Surveillance

I.INTRODUCTION:

Marine debris has a widespread effect on both land and marine life, making it a serious danger to environmental ecosystems. The identification and localization of such debris are essential for reducing its detrimental consequences; nevertheless, the accuracy, effectiveness, and applicability of current approaches are sometimes limited. This study offers a novel method by fusing sophisticated attention mechanisms—specifically, coordinate attention and the bottleneck transformer—with the potent instance detection capabilities of YOLOv11. These models' capacity to improve detection performance in a range of situations is assessed. The bottleneck transformer, however marginally less reliable, has special benefits in locating missed debris areas and enhancing accuracy in detecting larger objects, even though YOLOv11 with coordinate attention exhibits strong overall performance. The results highlight the potential of attention-enhanced YOLOv11 models for the detection of maritime trash and stress how crucial it is to choose the right model to satisfy certain operational requirements in environmental monitoring.

This study's main goal is to create an efficient and successful system for locating and identifying maritime debris using cutting-edge deep learning methods, particularly YOLOv11 combined with attention processes. Through the integration of Yolov11's instance detection capabilities with attention models like coordinate attention and the bottleneck transformer, the study seeks to improve debris recognition precision and accuracy in intricate marine environments. In order to evaluate these models' performance, the study looks at their speed of detection, dependability, and capacity to manage a variety of trash sizes. In order to provide a thorough method of marine debris detection that can be used with satellite imagery, drone surveillance, and other real-time monitoring systems, the study also intends to explore how these models can be adjusted to meet particular environmental monitoring needs. The ultimate goal is to aid in the creation of detecting systems that are more resilient and flexible in order to lessen the negative effects of marine trash on the environment.

II. LITERATURE SURVEY:

H. Liu, C. Li, Q. Wu, and Y. J. Lee's 2023 work Visual Instruction Tuning presents a novel method for enhancing visual models' performance by task-specific assistance. The performance of traditional visual models may be limited in specialized domains because they are usually trained on huge datasets without explicit instructions targeted to specific tasks. By including thorough visual signals and instructions straight into the training process, Visual Instruction Tuning (VIT) fills this gap. By learning from more contextually rich inputs, this technique helps models to better comprehend challenging visual tasks including object detection, image captioning, and semantic segmentation. The authors look into how VIT might be used in multi-modal AI systems that need to improve performance on certain tasks while also integrating textual and visual input. VIT bridges the gap between visual perception and linguistic understanding, enhancing a model's reactivity and adaptability to a variety of task demands. Its usefulness is demonstrated by experimental evaluations on multiple benchmark datasets, which show that VIT significantly improves accuracy and task comprehension. The promise of instruction-based tuning as a foundational strategy for developing more intelligent and adaptable visual AI systems is demonstrated by this study.

III. EXISTING SYSTEM:

The cutting-edge deep learning model YOLOv7, which is well-known for striking a balance between speed and accuracy, was created for real-time object identification and instance segmentation. Expanding upon earlier iterations of the YOLO (You Only Look Once) series, YOLOv7 presents a number of architectural advancements, such as Compound Model Scaling, dynamic label assignment for optimal training, and Extended Efficient Layer Aggregation Networks (E-ELAN). These improvements greatly increase the model's efficiency and learning capabilities, making it very useful for item detection and classification at different scales. Because of its dependability and accuracy, YOLOv7 is a well-liked option for applications where accurate, real-time performance is crucial, such as autonomous driving, security systems, environmental monitoring, and marine trash identification.

Disadvantages of YOLOv7 include its grid-based approach, which limits the model's ability to capture long-range dependencies, reducing its effectiveness in interpreting global context within complex scenes. This shortcoming can impair performance when objects are spaced far apart or exhibit intricate spatial relationships. Additionally, the model's reliance on predefined anchor boxes constrains its adaptability to detect objects with diverse shapes and sizes, often leading to inaccuracies when dealing with irregularly shaped items such as marine debris. YOLOv7 also faces challenges in identifying small objects, particularly in difficult settings like aerial imagery, where limited resolution and receptive field size can cause missed detections or incorrect classifications. Another limitation is its relatively weak instance segmentation performance, which prevents it from generating precise object boundaries—an essential feature for detailed tasks such as mapping marine debris. Furthermore, despite its efficiency during inference, the training phase of YOLOv7 demands significant computational resources, making it less accessible to organizations without advanced hardware capabilities.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Special Edition - Volume 09 NCFT -2025 SJIF Rating: 8.586 ISSN: 2582-3930

IV. PROPOSED SYSTEM:

YOLOv11 expands on its predecessors by combining transformer-based modules for global context capture with convolutional neural networks (CNNs) for targeted feature extraction. Detecting marine waste, which frequently consists of items of various sizes, shapes, and materials, is made easier by this dual capability. Given the existence of both huge objects like fishing nets and buoy fragments and little ones like plastic straws and bottle caps, a model that can recognize many scales is necessary. In order to accomplish this, YOLOv11 uses adaptive spatial attention mechanisms and a sophisticated feature pyramid network (FPN), which guarantees precise identification and localization across a variety of object scales.

The ability of YOLOv11 to handle complex backgrounds, such as reflective water surfaces, waves, and shadows, is one of its defining features. By incorporating self-attention mechanisms and dynamic feature refinement, YOLOv11 can distinguish marine debris from natural elements like seaweed, driftwood, and marine life. This is particularly crucial when detecting items like transparent plastic bags, which can blend into the ocean's surface, or fishing nets that may be partially submerged and entangled with natural objects.

METHODOLOGIES:

Input Data Module

Handles image dataset collected from sources like satellite imagery, drones, or underwater cameras for marine debris monitoring. Accommodates various environmental conditions such as water surface reflections and varying lighting to ensure diverse data sources are effectively managed.

Preprocessing Module

Implements Adaptive Histogram Equalization for contrast enhancement, helping to improve visibility of debris against complex ocean backgrounds.

Standardizes image size and format for consistency across varying image resolutions and input types, ensuring data uniformity for detection models.

Segmentation Module

Utilizes Super-Pixel-Based Fast Fuzzy C-Means (FCM) clustering to segment marine debris from the water and background effectively. Focuses on segmenting objects like plastic bags, bottles, fishing nets, and masks, even in cluttered or partially occluded environments.

Feature Extraction Module

Incorporates hybrid transformer-based attention mechanisms and convolutional layers to extract key features such as edges, textures, and shapes that differentiate marine debris from surrounding water or natural elements. Enhances the detection model ability to capture subtle details in complex debris, improving detection rates for various types of waste.

YOLOv11 Detection Module

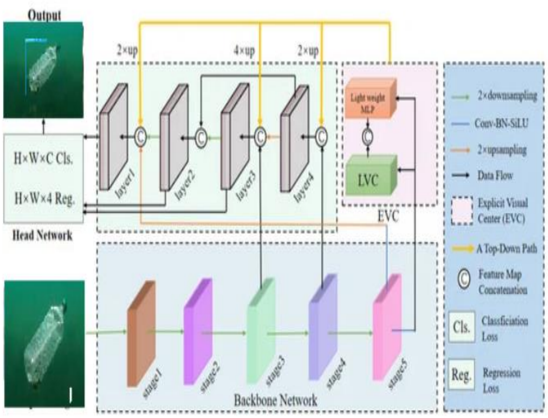
Deploys the cascaded YOLOv11 model for multi-stage detection of marine debris, ensuring higher accuracy in detecting various sizes and types of debris.

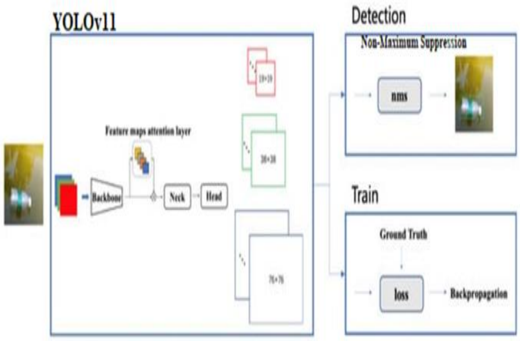
Uses anchor-free detection to reduce computational overhead and improve detection accuracy, allowing for better identification of irregularly shaped debris like tangled fishing nets and dispersed plastic.

Output and Visualization Module

Displays results with bounding boxes and detection confidence scores, allowing for easy identification and tracking of debris locations in images, videos, or real-time monitoring feeds. Provide visual outputs in the form of marked images, video streams, or real-time feeds with overlaid data, assisting in real-time decision-making for marine debris cleanup and monitoring efforts.

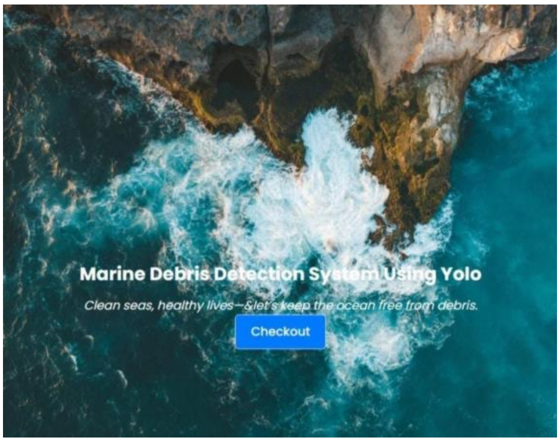
SYSTEM ARCHITECTURE:



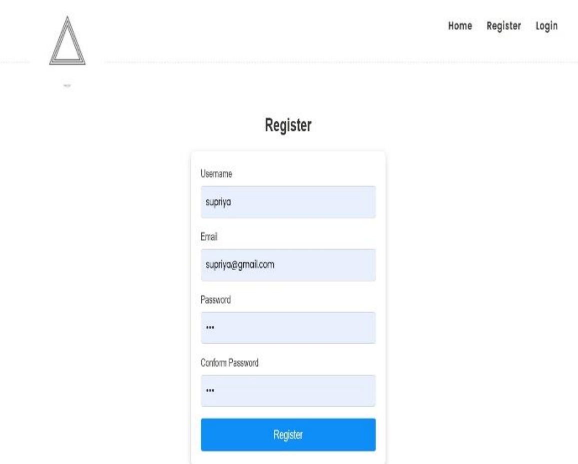


V.RESULT AND IMPLEMENTATION

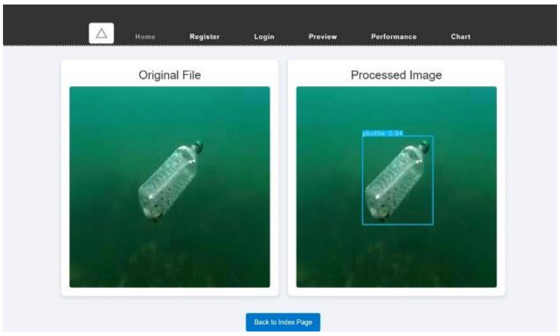
User interface:



Registration:



Result:



VI. FUTURE ENHANCEMENT:

Future enhancements for attention-enhanced YOLOv11 models in marine debris detection include integrating multimodal data sources—such as spectral, thermal, or sonar data—with visual imagery to improve accuracy in challenging environments like turbid water or low light. Optimizing these models for real-time deployment on edge devices (e.g., drones or underwater vehicles) would enable efficient monitoring in remote or resource-constrained areas. Additionally, incorporating advanced training techniques such as self-supervised learning, domain adaptation, and continual learning can enhance model robustness against environmental variability. Increasing dataset diversity by including annotated images from different geographic and ecological regions would also boost model generalizability, ensuring better performance across diverse marine conditions and debris types.

VII. CONCLUSION:

In conclusion, the integration of advanced attention mechanisms, such as coordinate attention and the bottleneck transformer, with YOLOv11 represents a significant step forward in addressing the challenges of marine debris detection. The study highlights the robust performance of YOLOv11 with coordinate attention in delivering consistent and accurate results across diverse scenarios. Additionally, the bottleneck transformer model demonstrates its unique capability in detecting overlooked debris and enhancing precision for larger objects, offering complementary strengths. These findings emphasize the critical role of attention-enhanced YOLOv11 models in environmental monitoring, showcasing their potential to improve detection accuracy and efficiency. Future work should focus on further refining these models and exploring their adaptability to different ecological and operational contexts, ensuring effective deployment for the preservation of environmental ecosystems.

VIII. REFERENCES:

[1] S. Chu, J. Wang, G. Leong, L. Ann Woodward, R. J. Letcher, and Q. X. Li, "Perfluoroalkyl sulfonates and carboxylic acids in liver, muscle and adipose tissues of black-footed albatross (Phoebastria nigripes) from Midway Island, North Pacific Ocean," *Chemosphere*, vol. 138, pp. 60–66, 2015, doi: 10.1016/j.chemosphere.2015.05.043.

[2] H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognit.*, vol. 34, no. 12, pp. 2259–2281, 2001, doi: 10.1016/S0031-3203(00)00149-7.

[3] M. F. Fava, "Ocean plastic pollution an overview: Data and statistics," *Ocean Literacy Portal*, Jun. 9, 2022. [Online]. Available: https://oceanliteracy.unesco.org/plastic-pollution-ocean/

[4] L. Lebreton et al., "Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic," *Sci. Rep.*, vol. 8, 2018, Art. no. 4666, doi: 10.1038/s41598-018-22939-w.

[5] T. Maes, et al., United Nations Environment Programme, "From pollution to solution: A global assessment of marine litter and plastic pollution," Nairobi, 2021. [Online]. Available: https://research.usc.edu.au/esploro/outputs/report/From-Pollution-to-Solution-A-Global/99584903702621