UAV-based Coastal Debris Detection: Analyzing the Effectiveness of Enhanced YOLOv8 for Improved Aerial Surveillance

M Manishwaran
School of Computer Science and Engineering.
Vellore Institute of Technology.
Vellore, Tamil Nadu, India

Leona E Joseph
School of Computer Science and Engineering.
Vellore Institute of Technology.
Vellore, Tamil Nadu, India

Meghana Vinod Gaonkar
School of Computer Science and Engineering.
Vellore Institute of Technology.
Vellore, Tamil Nadu, India

Balamurugan R
School of Computer Science and Engineering.
Vellore Institute of Technology.
Vellore, Tamil Nadu, India

***Abstract*—In the face of the pressing issue of pollution, in our oceans, where a staggering 8 million tons of plastic finds its way into the water each year resulting in a mind-boggling 5.25 trillion pieces of waste, our study presents an innovative method for detecting and identifying shoreline and waterborne garbage. This method is crucial for addressing the problem that sees 70% of debris sink into the marine ecosystem while 15% floats and another 15% ends up on beaches. By utilizing machine learning techniques and unmanned aerial vehicle (UAV) technology our approach significantly improves both accuracy and speed in detecting garbage. We conducted an evaluation using configurations of the YOLOv8 model along with a specialized dataset captured by UAVs incorporating custom augmentation techniques. Our results demonstrate how this model excels at identifying types of garbage. This groundbreaking research not only advances object detection technology for conservation but also emphasizes the importance of cutting-edge solutions, in mitigating the profound impact that plastic pollution has on marine ecosystems.**

***Index Terms*—YOLOv8, Wiou Modification, UAV Data Analysis, Dataset Augmentation Techniques, Custom Hyperparameters, Object Detection**

I. Introduction

The widespread problem of plastic debris along shorelines is not only a crisis for the environment, but it also poses serious risks to coastal industries and marine ecosystems. Plastic waste, which can range in size from microplastics to bigger objects, entangles and consumes marine life, causing death, destruction of habitats, and a decrease in biodiversity. The economic effects are just as severe since plastic pollution raises the expense of cleanup and degrades the aesthetic appeal of coastal areas, which hurts local businesses and tourism. Unmanned aerial vehicles (UAVs) fitted with detection algorithms provide an innovative means to effectively identify and map the spread of plastic garbage, thereby addressing this challenge. With the use of cutting-edge imaging and machine learning, these unmanned aerial vehicle (UAV) systems can precisely monitor impacted regions, enabling focused cleanup operations and aiding in the preservation of marine environments. Stakeholders can acquire a better understanding of the scope and details of plastic pollution by integrating UAV technology, which opens the door to more efficient management and reduction tactics in the battle against plastic trash along the shore.

In response to the pressing environmental challenge of coastal debris, this study introduces a novel UAV-based debris detection approach utilizing a Modified YOLOv8 architecture. Wise-IoU(Wise Intersection over Union), a metric that dynamically and non-monotonically focuses on the weighted overlap between areas, aimed at boosting both precision and accuracy in identifying marine debris, thereby offering an innovative solution to the pressing pollution challenge. The paper involves a detailed comparison of various YOLOv8 variants—yolov8n, yolov8m, yolov8s, yolov8l, and yolov8x—both with and without the Wise-IoU modification.

The dataset encompasses UAV data from two distinct sources, creating a comprehensive and diverse environmental representation. Data augmentation strategies are used to boost the performance of the model. Custom hyperparameters for augmentation are strategically tuned for the combined dataset, ensuring optimal model performance in the detection of debris using UAV.

As global pollution increases, we aim to use our comparative analysis of the Wise-IoU integrated Modified YOLOv8 in coastal debris detection. Subsequent sections delve into the methodology, results, and implications, providing valuable insights into potential advancements in UAV-based environmental monitoring.

II. Related Work

By applying WIoU to YOLOv7, it significantly enhances detection performance on the MS-COCO dataset, demonstrating the effectiveness of this dynamic approach over traditional static methods. Applying WIoU to the cutting-edge, real-time detection system YOLOv7 enhances the AP75 metric on the MS-COCO dataset from 53.03% to 54.50%.

The enhanced YOLOv8 method was applied to the UAVVaste Dataset, showing significant improvement over prior benchmarks. Previous studies, utilizing the YOLOv4 model, reported accuracy scores of 62% [1] . The highest mean Average Precision (mAP) recorded for plastic detection in another study reached 93% [2].

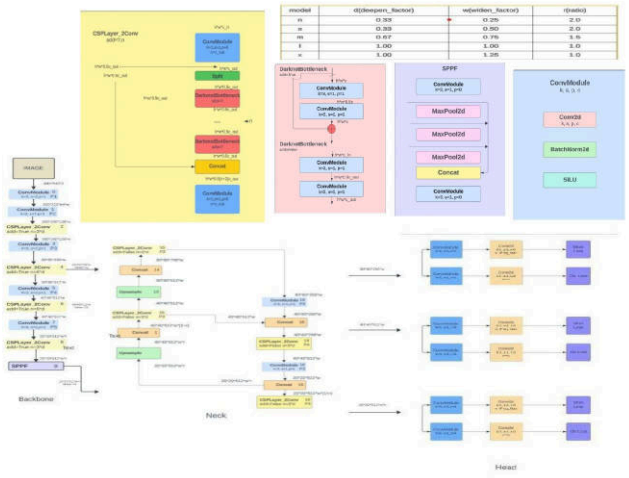


Fig. 1. YOLOv8 architecture diagram based on [5].

III. YOLOv8 Architecture

In January 2023, Ultralytics, the developers behind YOLOv5, launched YOLOv8, presenting five different versions: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra-large). YOLOv8 sets itself apart from earlier object detection systems with a number of significant enhancements. First of all, because of its improved feature extraction capabilities, small and variable debris items—which are sometimes difficult to detect—can be precisely identified. Secondly, by focusing on key image regions and adapting to the scale of objects, YOLOv8's advanced anchoring techniques and integration with attention mechanisms greatly increase the detection accuracy. Furthermore, thorough coverage of vast coastal areas is made possible by the algorithm's effectiveness in processing and analyzing high-resolution aerial imagery from satellites or unmanned aerial vehicles (UAVs).

YOLOv8 utilizes an altered backbone similar to that of YOLOv5, incorporating the C2f module to improve detection accuracy through the combination of high-level features with contextual information. By adopting an anchor-free model with a separate head, it processes tasks related to object presence, classification, and regression independently, leading to enhanced overall accuracy. Sigmoid activation determines objectness probability, while softmax computes class probabilities.By using CIoU and DFL loss functions for bounding-box optimization and binary cross-entropy for classification tasks, YOLOv8 achieves superior performance in the detection of smaller objects. [3] [4]

Based on its improved performance in detecting objects within aerial images, as demonstrated in studies [7] [8], YOLOv8 emerged as the preferred option for identifying coastal debris in aerial images.

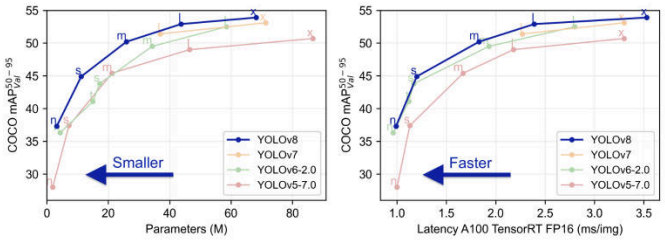


Fig. 2. Comparison between YOLO-v8 and its predecessors. [6]

IV. Hyperparameters

A. Degrees

The process of rotating the input images by a specific angle or degree is known as image angle/degree augmentation. The optimal value for our test case was found to be 45.

B. Fliplr and flipud

Flip-up-down augmentation, achieved by vertically flipping the image, produces a mirrored version where the top and bottom are reversed. This addition aids in the YOLO model's learning process for identifying objects that might appear inverted or upside down in actual situations. The optimal value for our test case was found to be 0.5 On the other hand, flip left-right augmentation involves horizontally flipping the image, transforming the left side into the right and the right into the left, thereby creating a mirror image. The YOLO model can now recognize and learn from objects from various angles thanks to this enhancement. 0.6 was determined to be the optimal value for our test case.

C. MixUp

MixUp augmentation generates fresh training examples by combining image pairs with appropriate item labels. The YOLO model gains the ability to identify common item features and improve its ability to generalize across other classes by combining photos and their labels. In our specific test scenario, the optimal value was determined to be 0.05.

V. WIOU Math

A. Intersection over Union

Intersection over Union (IoU) is a metric used to assess how much a predicted bounding box and its corresponding ground truth bounding box overlap. It is essential for determining how accurate object localization is. By efficiently reducing the influence of bounding box dimensions, IoU guarantees an equal learning approach for model objects of varying sizes.

B. Wise-Intersection over Union

A dynamic, non-monotonic focusing approach for bounding box regression loss in object detection is provided by WIoU (Wise-Intersection over Union). Its goal is to improve model performance by skillfully handling samples of poor quality. Unlike previous approaches that use static focusing mechanisms, WIoU dynamically adjusts gradient gains based on

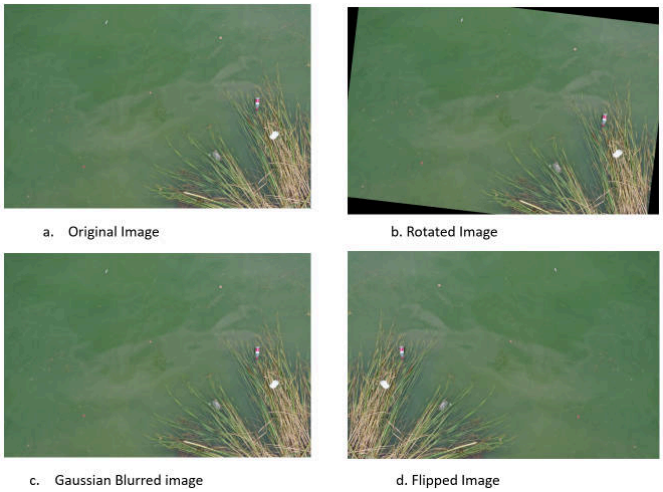


Fig. 3. Augmentation techniques used on the data

the quality of each example, which improves localization and generalization performance. The incorporation of this dynamic methodology into YOLOv7 significantly improves detection performance on the MS-COCO dataset, demonstrating its superiority over traditional static methods. [9]

With a dynamic focussing mechanism, WIoU v1 is a bounding box regression loss. It makes use of two parts:

· RWIoU , which ranges from 1 to e, amplifies LWIoU for ordinary-quality anchor boxes.

· LWIoU, ranging from 0 to 1, diminishes RWIoU for high-quality ones, focusing on the center points' alignment. [9]

The loss is calculated as LWIoU = RWIoU. LWIoU, with RWIoU being a normalized exponential function of the smallest enclosing box's size and the distance between the anchor and ground truth boxes' centers[7.17]. The goal of this approach is to improve model performance and generalization by lessening the influence of geometric elements on the loss function. Likewise, a monotonic focusing coefficient, LWIoUv2, is proposed for LWIoUv1, tackling slow convergence by dynamic normalizing factor updates determined by the average LIoU [7.20]

VI. Dataset

The dataset employed in this research was split into two sections, a beach-side UAV dataset [10] consisting of 772 images and a river-side UAV dataset [11] consisting of 40 images. The dataset utilized for training and validation combines UAV data from UAVVaste and the UAV data monitoring plastic waste dataset, encompassing diverse environmental conditions. We categorized the debris into one class: Garbage. Deep learning networks depend heavily on a vast quantity of labelled data for feature extraction and learning processes. However, the presence of noise and inaccuracies within the dataset can detrimentally impact model performance. The issue intensifies when the dataset contains a significant volume of mislabelled or incorrect data, preventing the model from identifying valid patterns accurately.

MODEL	TYPE	mAP-50	mAP50-95	Precision	Recall
YOLOv8n	Original	0.818	0.375	0.832	0.732
	With HP	0.839	0.399	0.84	0.78
	With HP and WIOU	0.829	0.402	0.835	0.763
YOLOV8s	Original	0.808	0.388	0.83	0.742
	With HP	0.835	0.406	0.815	0.776
	With HP and WIOU	0.826	0.414	0.859	0.755
YOLOv8m	Original	0.818	0.372	0.827	0.772
	With HP	0.847	0.423	0.861	0.799
	With HP and WIOU	0.861	0.429	0.854	0.8
YOLOv8l	Original	0.821	0.387	0.794	0.805
	With HP	0.862	0.422	0.836	0.814
	With HP and WIOU	0.645	0.225	0.682	0.574
YOLOv8x	Original	0.84	0.401	0.842	0.786
	With HP	0.844	0.422	0.829	0.791
	With HP and WIOU	0.599	0.192	0.656	0.59

Fig. 4. Comparison of evaluation metrics for YOLOv8 varients

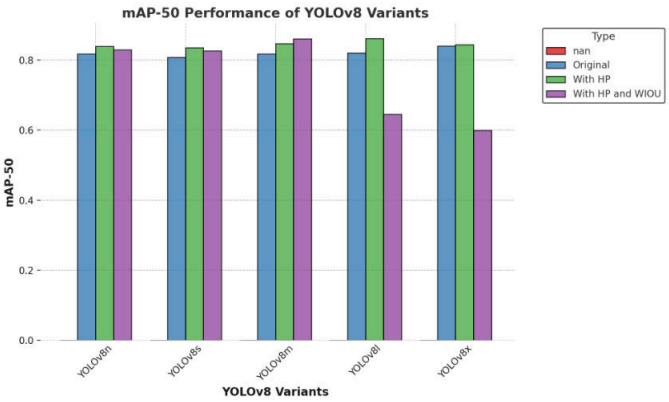


Fig. 5. Comparison of evaluation metrics for YOLOv8 varients

Additionally, a limited dataset can lead to overfitting, where the network overly focuses on the noise and irrelevant details in the image, resulting in decreased accuracy during testing. To address these challenges, this paper introduces offline data enhancement techniques. These techniques involve adding Gaussian blur, applying random rotations, adjusting the scale of images, and manual annotation of images using annotation software. Additionally, drone images capturing sandy beaches and anthropogenic litter along the Saudi Arabian Red Sea were incorporated for testing purposes. These strategies are designed to enhance the model's generalization capabilities and its resilience, which helps reduce the risk of overfitting. In the current research, data augmentation was applied to 40 annotated images, yielding an additional 120 augmented images [12]. The completed dataset ultimately comprised 892 images (772 plus 120), which were allocated into three distinct sets for training, validating, and testing, containing 636, 128,

and 128 images respectively.

VII. Results

A. Evaluation Metrics

Evaluation Metrics: The precision metric calculates the ratio between the total number of predicted positive samples and actual positives in the sample space. The percentage of true positive samples to the total number of positive samples is known as recall. A higher recall suggests that the model minimizes false negatives by better capturing all positive cases. A parameter applied to object detection tasks is mAP50. The average accuracy (AP) is determined for every class at an Intersection over Union (IoU) threshold of 0.5. These values are then averaged for all classes. It offers a general indicator of detection accuracy. The mAP50-95 computation expands the scope of mAP to include IoU thresholds that vary from 0.5 to 0.95. This provides a more thorough assessment of the model's performance at several junction points between the ground truth and anticipated bounding boxes.

For various threshold levels, the precision-recall curve shows recall on the x-axis and precision on the y-axis. Better model performance is shown by a greater precision-recall curve. A curve that is at the top-right corner, where recall and precision are both high, is what we ideally want. A single scalar metric that summarizes the model's performance across all potential thresholds is provided by the areas under the curve (PR AUC). Better model performance is indicated by a higher PR AUC.

B. Hyper-Parameters Optimization (HP) Impact

In most cases, adding hyper-parameters optimization (HP) to the models leads to improvements in performance metrics compared to the original models. This is evident across all model sizes for metrics like Precision, Recall, mAP-50, and mAP-50-95, indicating that HP generally enhances model accuracy and detection quality.

Despite the slight decrease in Precision (0.815 with HP vs. 0.830 original), the YOLOv8s model shows improvements in mAP-50 (0.835 with HP vs. 0.808 original), and mAP-50-95 (0.406 with HP vs. 0.388 original), suggesting a trade-off where overall detection quality is enhanced at the cost of a minor precision drop.

C. Adding Wise Intersection Over Union (WIOU)

Incorporating WIOU after HP optimization shows mixed effects across different model sizes. For YOLOv8n (mAP-50 Original-0.818 and WIoU -0.829), YOLOv8s(mAP 50 Original-0.808 and WIoU -0.826), YOLOv8m(mAP-50 Original-0.818 and WIoU -0.861), it further enhances some metrics, particularly mAP-50-95, indicating better average precision across a range of IoU thresholds. The enhancement in mAP metrics showcases its capability to boost the precision of object detection. This is achieved by controlling the impact of sample quality on the training of the model and its performance in localizing objects. However, it slightly reduces Precision and Recall in certain models, suggesting a trade-off between detecting more objects at the expense of precision in some scenarios.

Significant Drop in Larger Models: Notably, the largest models (YOLOv8l and YOLOv8x) experience a significant drop in all metrics with the addition of WIOU after HP. This drastic reduction, especially in the YOLOv8l and YOLOv8x models, suggests that the WIOU adjustments may not scale well with larger models or could be counterproductive in certain configurations.

D. Model Size and Performance

Size vs. Performance: Larger models do not necessarily guarantee better performance across all metrics. While YOLOv8m shows consistent improvement further with HP and WIOU(Original- 0.817 and mAP score with HP-0.847 and with HP+WIOU - 0.861), the largest models YOLOvl - Precision drops significantly to 0.682 from 0.836 with HP, and mAP-50-95 plummets to 0.225 from 0.422 with HP land YOLOv8x- Precision sees a drastic reduction to 0.656 from 0.829 with HP, and mAP-50-95 drops to 0.192 from 0.422 with HP when Wiou is added highlighting that model optimizations may need to be tailored to the specific architecture and size of the model.

The addition of HP generally benefits all models, but the subsequent addition of WIOU is more beneficial for smaller models (e.g., YOLOv8n, YOLOv8s, YOLOv8m) in terms of certain metrics like mAP-50-95. For larger models, particularly YOLOv8l and YOLOv8x, this adjustment leads to a notable decrease in performance, indicating the complexity of optimising larger models.

VIII. Conclusion

This paper successfully exhibits the improvement of garbage detection capabilities in aquatic and coastal areas through the integration of a modified YOLOv8 with Wise Intersection Over Union (WIOU). Significant improvements in detection accuracy were achieved by completing a comprehensive analysis across several YOLOv8 setups, as well as employing data augmentation and hyperparameter optimization on a specialized UAV dataset, particularly with smaller models. These findings underscore the potential of advanced object detection techniques in environmental monitoring, delivering a promising avenue for improving the precision and efficiency of detecting and handling waste in critical ecosystems, and thus contributing to greater initiatives in environmental preservation and sustainability.

IX. Future Works

Future investigations will primarily address the scalability challenges of WIoU with larger models, as our findings show a significant reduction in performance metrics for the YOLOv8l and YOLOv8x variants. Exploring adaptive or model-specific optimization techniques may yield promising solutions. We also intend to improve WIoU's application, possibly using advanced machine learning algorithms or alternative optimization methods, to assure scalability and efficiency. Delving

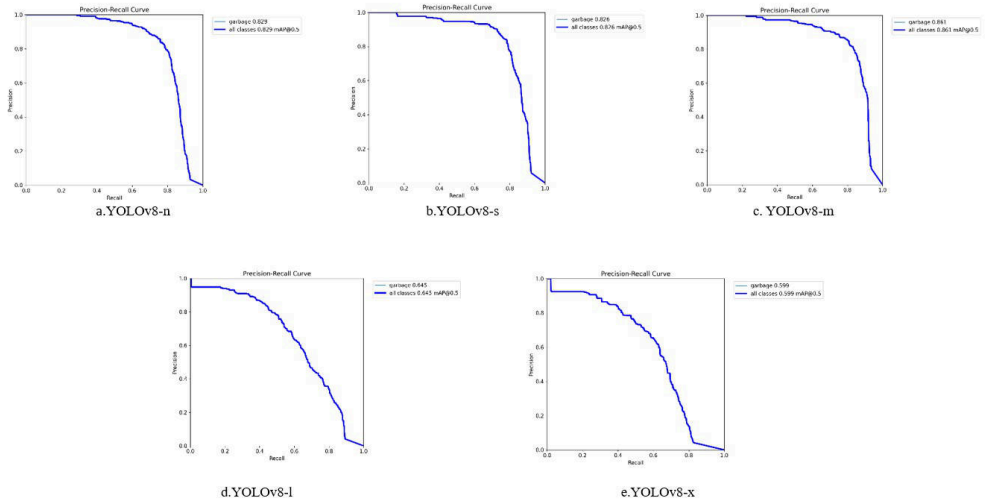


Fig. 6. Precision and Recall curves for YOLOv8 varients enhanced with hyperparameters and WIOU

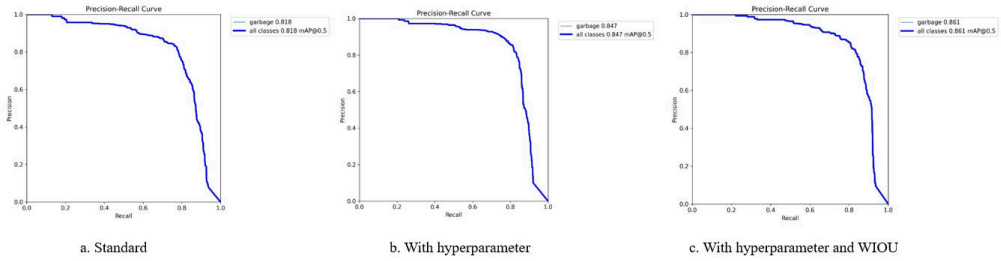


Fig. 7. Precision and Recall curves for YOLOv8m

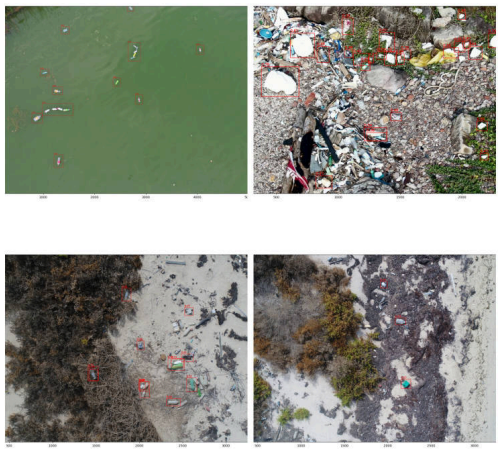


Fig. 8. Debris Detection using YOLOv8m with Custom Hyperparameters and WIOU

deeper into scalable as well as efficient data-augmentation strategies offers the potential to enhance model resilience in several kinds of environmental conditions. Furthermore, expanding the dataset to cover a broader range of waste and environmental conditions might improve model generalizability. Investigating the integration of real-time detection capabilities and automated waste collection systems could also extend the practical applications of this research in combating environmental pollution.

REFERENCES

[1] A. J. M, S. Nandini, and A. T, "Real-time litter detection system for moving vehicles using yolo," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, (Tirunelveli, India), pp. 1311–1315, IEEE, 2022.

[2] W. Han, W. Luo, Y. Jin, and M. Zhu, "A deep learning model for automatic plastic waste monitoring using unmanned aerial vehicle (uav) data," *Authorea Preprints*, 2022.

[3] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, 2023.

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE*

conference on computer vision and pattern recognition, pp. 779–788, 2016.

[5] M. Contributors, "Yolov8 by mmyolo." https://github.com/open-mmlab/mmyolo/tree/main/configs/yolov8, 2023. Accessed on 13 May 2023.

[6] G. Jocher, A. Chaurasia, and J. Qiu, "Yolo by ultralytics." https://github.com/ultralytics/ultralytics, 2023. Accessed on 12 January 2023.

[7] I. P. Sary, S. Andromeda, and E. U. Armin, "Performance comparison of yolov5 and yolov8 architectures in human detection using aerial images," *Ultima Computing: Jurnal Sistem Komputer*, vol. 15, no. 1, pp. 8–13, 2023.

[8] J. Farooq, M. Muaz, K. Khan Jadoon, N. Aafaq, and M. K. A. Khan, "An improved yolov8 for foreign object debris detection with optimized architecture for small objects," *Multimedia Tools and Applications*, 2023.

[9] Z. Tong, Y. Chen, Z. Xu, and R. Yu, "Wise-iou: Bounding box regression loss with dynamic focusing mechanism," *arXiv preprint arXiv:2301.10051*, 2023.

[10] W. Han, "Uav beach dataset." https://www.scidb.cn/en/detail?dataSetId=0bdfaa4da799474eba39ca9ce939083b, 2023. Accessed: insert the date here.

[11] C. Martin, Q. Zhang, D. Zhai, Z. X., and C. M. Duarte, "Uav river side dataset." https://data.mendeley.com/datasets/gpdsntb3y6/1, 2023. Accessed: insert the date here.

[12] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, "A lightweight yolov8 tomato detection algorithm combining feature enhancement and attention," *Agronomy*, vol. 13, no. 7, p. 1824, 2023.

[13] T. Wu and Y. Dong, "Yolo-se: Improved yolov8 for remote sensing object detection and recognition," *Applied Sciences*, vol. 13, no. 24, p. 12977, 2023.

[14] S. G. CS, "Intelligent debris mass estimation model for autonomous underwater vehicle," *arXiv preprint arXiv:2309.10617*, 2023.

[15] Y. Zhang, Y. Yin, and Z. Shao, "An enhanced target detection algorithm for maritime search and rescue based on aerial images," *Remote Sensing*, vol. 15, no. 19, p. 4818, 2023.

[16] H. Zheng, X. Guo, G. Guo, Y. Cao, X. Hu, and P. Yue, "Full stage networks with auxiliary focal loss and multi-attention module for submarine garbage object detection," *Scientific Reports*, vol. 13, no. 1, p. 16115, 2023.

[17] V. D. Matta, K. V. R. R. Mudunuri, B. C. S. Sai Baba, K. B. Kiran, C. L. Veenadhari, and B. V. Prasanthi, "Single use plastic bottle recognition and classification using yolo v5 and v8 architectures," in *International Conference on Cognitive Computing and Cyber Physical Systems*, pp. 99–106, Springer Nature Switzerland, December 2023.

[18] W. Zhou, F. Zheng, G. Yin, Y. Pang, and J. Yi, "Yolotrashcan: A deep learning marine debris detection network," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, no. Art no. 5002012, pp. 1–12, 2023.

[19] N. A. Zailan, A. S. Mohd Khairuddin, K. Hasikin, and et al., "An automatic garbage detection using optimized yolo model," *SIViP*, vol. 18, pp. 315–323, 2024.

[20] D. Ma, J. Wei, Y. Li, F. Zhao, X. Chen, Y. Hu, S. Yu, T. He, R. Jin, Z. Li, and M. Liu, "Mldet: Towards efficient and accurate deep learning method for marine litter detection," *Ocean & Coastal Management*, vol. 243, p. 106765, 2023.

[21] N. A. Zailan, A. S. Mohd Khairuddin, U. Khairuddin, and A. Taguchi, "Yolo-based network fusion for riverine floating debris monitoring system," in *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–5, 2021.

[22] A. Sánchez-Ferrer, J. J. Valero-Mas, A. J. Gallego, and J. Calvo-Zaragoza, "An experimental study on marine debris location and recognition using object detection," *Pattern Recognition Letters*, vol. 168, pp. 154–161, 2023.

[23] K. Sasaki, T. Sekine, L. J. Burtz, and W. J. Emery, "Coastal marine debris detection and density mapping with very high resolution satellite imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 6391–6401, 2022.

[24] Y. Taddia, C. Corbau, J. Buoninsegni, U. Simeoni, and A. Pellegrinelli, "Uav approach for detecting plastic marine debris on the beach: A case study in the po river delta (italy)," *Drones*, vol. 5, no. 4, p. 140, 2021.

[25] T. Tran, Z.-C. Huang, K.-H. Tseng, and P.-H. Chou, "Detection of bottle marine debris using unmanned aerial vehicles and machine learning techniques," *Drones*, vol. 6, no. 12, p. 401, 2022.

[26] J. Liu and Y. Zhou, "Marine debris detection model based on the improved yolov5," in *2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, pp. 725–728, 2023.

[27] F. Zocco, T. C. Lin, C. I. Huang, H. C. Wang, M. O. Khyam, and M. Van, "Towards more efficient efficientdets and real-time marine debris detection," *IEEE Robotics and Automation Letters*, vol. 8, no. 4, pp. 2134–2141, 2023.