UNDERWATER PLASTIC DETECTION USING YOLO V8 AND V10

MSc Research Project
Data Analytics

Soma Sekhar Bogisam
Student ID: x23127627@student.ncirl.ie

School of Computing
National College of Ireland

Supervisor: Christian Horn

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Soma Sekhar Bogisam
Student ID:	x23127627@student.ncirl.ie
Programme:	Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Christian Horn
Submission Due Date:	12/08/2024
Project Title:	UNDERWATER PLASTIC DETECTION USING YOLO V8 AND V10
Word Count:	7043
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Soma Sekhar Bogisam
Date:	16th September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

UNDERWATER PLASTIC DETECTION USING YOLO V8 AND V10

Soma Sekhar Bogisam
x23127627@student.ncirl.ie

Abstract

This research focuses on a comparative study of the performance of two versions, YOLOv8 and YOLOv10, in detecting small and medium-sized plastic litter underwater. The study experimented with 1,200 underwater images from the Deep-sea Debris Database using four variants: YOLOv8-s, YOLOv8-l, YOLOv10-s, and YOLOv10-l. Comparisons have been focused on detection accuracy and computational efficiency. Results demonstrate that the best performance is by YOLOv10-s, which achieves the highest mean average precision, mAP50:0.772, at the least computational resources of 24.4 GFLOPs. Generally, the YOLOv10 variants were returns as performing better compared to the YOLOv8 ones in most of the experiments; small models would be performatively better than their larger alternatives under conditions of equal training epochs.

These results have important implications for underwater plastic detection systems, providing very clear evidence of the need for task-specific optimization. The limitations brought up by the authors concern the size of the dataset and real-world testing diversity, and they outline some future directions for underwater plastic detector development.

Acknowledgments

I am very much indebted to my supervisor, Christian Horn, for his great advice in valuable discussions throughout the research project. It was based on his knowledge and encouragement that this work was realized. His dedication to academic rigor and his attention to detail inspired me to push the boundaries of my own work. I am profoundly grateful for his insightful suggestions and offering constructive feedback, which was essential for the successful completion of this study.

I would also like to extend my appreciation to the National College of Ireland for the resources and academic setting that have made this research possible.

I would also want to thank the Japan Agency for Marine-Earth Science and Technology for making open access to data from the Deep-Sea Debris Database, which was used as a basis for the dataset in this study. Lastly, I would also like to thank Roboflow for hosting and providing easy access to the curated subset of images hosted and used in this research. I also wish to thank developers of the Ultralytics YOLO framework, whose tools greatly facilitated the implementation and evaluation of models developed within this study.

Thank you, Dr. Christian Horn, for your support as a mentor throughout this journey.

1 Introduction

Marine plastic pollution has grown over the years to become a critical environmental issue, involving serious harm to aquatic ecosystems and biodiversity worldwide. Accordingly, detection and quantification of underwater plastic debris become very necessary for addressing this global problem. Recent advances in computer vision and deep learning techniques have provided encouraging solutions for automating the identification of submerged plastic wastes. (Tamin et al., 2022).

1.1 Background on YOLO (You Only Look Once)

YOLO is the very first family of real-time object detection algorithms; at the moment, it has beaten every other computer vision algorithm from 2015 till now. Joseph Redmon mainly worked on it, along with Ali Farhadi. This key innovation originally treated object detection as a regression problem for the very first time to be processed at much faster speeds compared to prior methods. From the very first model to the latest YOLOv10 in 2024, improvements so far made to YOLO have been oriented toward increasing speed, making it versatile, and accurate. Most recently, improvements have made YOLO applicable in many fields, from self-driving and surveillance to environmental monitoring, like detecting underwater debris.

The basic architecture of YOLO includes a backbone for extracting features from images, a neck to fuse features from different scales, and a head for actual detection. The recent variants introduced anchor-free detection, greatly simplifying object localization, and NMS-free training that reduces processing steps for faster performance. The system has also overspilled into other tasks beyond mere object detection, such as instance segmentation and pose estimation. Improvements to the YOLO design constantly acted to raise the bar of what was believed possible in real-time object detection, becoming one of the enabling technologies at the core of modern-day computer vision applications.

Architectural differences between YOLOv8 and YOLOv10 make for some interesting trade-offs that apply to underwater plastic detection. While YOLOv8 is more versatile, bearing many features that can be applied to many tasks very useful in comprehensive environmental monitoring YOLOv10's efficiency and NMS-free approach might give it advantages in real-time processing and deployment on resource-constrained underwater vehicles or edge devices.

Object detection algorithms have been outstanding recently in various domains, especially the YOLO family. However, the application of underwater plastic detection comprises varying water conditions, illumination, and small or even partly occluded nature of the debris, which makes the task very unique compared to other applications (Hipolito et al., 2021). This paper tries to fill up the literature gap by evaluating and comparing the performance of using the real-world data different versions and model sizes of YOLO, especially for underwater plastic trash detection.

The research question guiding this study is: **"How well do different versions of YOLO (YOLOv8 and YOLOv10) work for finding small and medium-sized plastic trash in underwater pictures? How do these different sizes of models (YOLOv8-s, YOLOv8-l, and YOLOv10 -s and YOLOv10-l) compare in terms of how accurately they can spot trash versus how fast they can run on a computer when looking at underwater images?"**

The primary objectives of this research are:

- To evaluate the detection accuracy of YOLOv8 and YOLOv10 for small and medium-sized underwater plastic debris.
- To compare the computational efficiency of YOLOv8-s, YOLOv8-l, and YOLOv10-s and YOLOv10-L in processing underwater images.
- To analyze the trade-offs between detection accuracy and processing speed for each model.

These will be achieved through a series of experiments on a highly varied dataset of underwater images, each including plastic litter. Standard metrics of object detection, such as mean Average Precision, and inference time, will be used to evaluate the model's performance. Such methodology will consist of training and fine-tuning the chosen YOLO models on a curated dataset of images containing underwater plastic debris. Afterwards, the models will be tested on a separated test set, and performance will be analyzed across different underwater conditions and sizes of objects.

This research contributes to the scientific literature by detailing, for this specific task of underwater plastic detection, comparisons of state-of-the-art object detection models. Its findings are useful to an audience of researchers within academia in the fields of computer vision, marine ecology, and people involved in developing automatic systems related to monitoring and cleaning up marine pollution.

The structure of this report is as follows: Section 2 is a comprehensive literature review regarding underwater object detection and YOLO applications. Section 3 describes the methodology of the research, such as dataset preparation, model training, and evaluation procedures. The models' implementation and the experimental setup are presented in Section 4. The results are reported in Section 5. Section 6 concludes the study by discussing the implications of the findings and recommending areas for future research.

2 Related Work

2.1 Evolution of YOLO Models for Object Detection

The improvement in the YOLO family has been major since it was inception, with successive versions bringing substantial improvement in accuracy and efficiency. Wang et al. (2024) presented the new evolution in this family,YOLOv10. This further pushes the bar on performance efficiency trade-offs by applying a series of new techniques, such as consistent dual assignments for NMS-free training and holistic efficiency-accuracy-driven model design. These improvements have upgraded the models to be significantly faster and more efficient compared to their previous versions. For instance, YOLOv10-S is 1.8 times faster than RT-DETR-R18 (Wang et al,,2024)with almost the same accuracy. The authors, however, admit the limitations showing that there is a gap in the performance of NMS-free training versus one-to-many training with NMS.

Application of YOLO variants in agriculture has been made from v1 to v10, evidencing the versatility of the model across different tasks. Alif and Hussain(2024), provide an overall view of these applications in agriculture, noting the strengths and limitations that exist within YOLO models. The authors point out persistent challenges such as more inclusive datasets and model optimization for resource-limited environments. These findings become even more relevant when considering underwater plastic detection, where similar challenges take place in the collection of data and variability of the environment.

2.2 YOLO Applications in Underwater Environments

Application of YOLO models in underwater plastic detection is a critical area of research given the growing concern about marine pollution. Aldric Sio et al. (2022) demonstrated that, with regard to plastic waste floating in rivers, YOLOv5 performs quite well in detecting it at a precision rate of 79.14% and a recall rate of 57.37%. While these results look quite promising, the relatively low recall rate indicates that there may be scope for improvements, mostly regarding small or partially occluded objects. The focus on river environments in this study also limits its direct applicability to a diversity of underwater conditions.

Underwater detection of marine plastic waste is, however, beset with challenges like varying light conditions and water clarity. In line with this, researchers have proposed a solution using transfer learning and data augmentation. For instance, Hipolito et al. (2021) achieved a good result in mAP of 98.15% with the use of YOLOv3. Nevertheless, the problem of small sample size dataset dependency in the study questions the generalization of the model in different real-world underwater scenarios. Though the reported mAP is very high, it needs cautious consideration. Intrinsic homogeneity might have resulted in this small dataset, which raises concerns about overfitting and limited generalizability. Although their approach for transfer learning was absolutely original, the freezing of some layers could have prevented the model from fully fitting to the unique characteristics of underwater scenes.

An improved model of YOLOv5 with MobileNet as the backbone, together with an attention mechanism, raised a great deal of progress in the detection of marine litter. Liu and Zhou (2023) achieved detection precision of 79% with a recall rate of 63%.Their use of MobileNet as a backbone offers improved efficiency, crucial for real-time applications. However, the introduced attention mechanism may not be optimal in detecting small plastic particles since it was effective against bigger debris only. This study put a more balanced view about model performance by focusing on detection precision and recall rate in comparison to studies that reported only mAP. These metrics still admit improvements, mainly related to challenging underwater conditions.

2.3 YOLOv8 and YOLOv10: Advancements and Potential for Underwater Plastic Detection

YOLOv8 by Hussain (2024) , is the most radical development in the history of YOLO architecture: anchorless in the detection process, with advanced backbone networks, and improved training strategies. How far this anchor-free approach works regarding generalization on objects of very different sizes and shapes is fundamental to the detection of various plastic debris underwater. Its performance underwater in turbidity changes and illumination variation has, however, been scantily investigated up to now.

Improvements in YOLOv10 are likely to be better in detection accuracy and computational efficiency, which are critical in real-time underwater plastic detection. However, the performance of YOLOv10 in these underwater scenarios associated with detection of small and partially occluded objects has yet to be evaluated comprehensively. The work of Wang et al. (2024) has impressive improvements in object detection efficiency and accuracy.

Their evaluation, while complete on the COCO dataset, does not exactly transfer to underwater plastic detection scenarios. Specifically, due to usually blurred conditions of object boundaries in an underwater environment, this NMS-free training approach they proposed would likely bring extremely huge benefits, again remaining theoretical without specific testing in an aquatic setting.

2.4 Challenges in Underwater Object Detection

One of the major challenges to applying deep learning models for plastic waste detection is data deficiency and poor quality of datasets. This problem is more critical in an underwater environment where data collection is more challenging and costly. Tamin et al. (2022) provided some solutions to these challenges by suggesting data augmentation and capturing images of plastic wastes in different real scene environments. While this gives some valuable approaches, they still need further validation in underwater contexts. The authors' insights underline the pressing need for improving methods of data collection and quality assurance to improve deep learning models for the detection of plastic wastes, especially in challenging situations underwater.

Most of the methods based on intelligent strategies in data selection and labeling overcome the bottleneck of limited labeled object detection data in different domains. Their approach improved at least 9.23% in detection performance, focusing on kitchen waste scenarios; Qin et al. (2024) leaves open the question about applicability for underwater plastic detection. The following study indicates that the techniques for selective learning become very promising in the case of low amounts of data, yet it doesn't account for any problems related to the underwater environment. A reported gain is impressive, although this work remains limited to land-based scenarios basically raising questions about how good this technique is within the much more demanding underwater domain. These new adaptive learning approaches for the detection of plastics in water are very promising to close the data gap but require very special adaptation to cope with difficulties in aquatic environments.

2.5 Comparative Analysis of Object Detection Models

FCOS-Lite is an anchor-free network, as was pointed out previously, meaning that it boasts better speed-to-accuracy trade-off instances against YOLOv3 over the MS COCO dataset. Approaches that are anchor-free in nature can be beneficial in detection tasks. The study by Liu et al. (2021) presents very exciting results but did not include tests for underwater scenarios, hence a gap in how these advantages could turn around in challenging aquatic scenarios. Anchor-free approaches hold unexplored potential for underwater plastic detection. These techniques underline the possibility of fine-tuning and evaluating anchor-free methods that effectively detect marine litter in real-world scenarios.

Comparison of several CNN-based object detectors for container detection in seaports indicates that Faster RCNN has a strong preference for precision. In their study, Bandong et al.,(2021), give valuable insight into object detection in the maritime environment; however, their focus on seaport environments limits its applicability in underwater plastic detection because of demanding variables offered by the underwater context, including variation in visibility, great diversity of object size, and special environmental factors. While it includes good comparisons for CNN-based detectors, the uniqueness of the problem in underwater plastic detection does warrant further research in adapting and evaluation of those models in submerged environments.

2.6 Data Augmentation and Transfer Learning in Underwater Object Detection

Data augmentation and transfer learning techniques have a vital role in overcoming limited and poor datasets with respect to detecting underwater objects. Shin et al. (2022) proposed a new data augmentation strategy simulating diverse water environment conditions for improving the accuracy of detection. The tool creates synthetic maritime images by editing background scenes related to the horizon line, lighting condition, and weather. It achieves high diversity in training data by putting the foreground objects, for example, ships at plausible positions with respect to the horizon, blending them in naturally to the background. This augmentation increases the robustness and accuracy of object detection models rightfully by exposing them to a wide variety of environmental scenarios.

Transfer learning strategies, explored by Zhang et al. (2023), open up possibilities in generalizing models pre-trained on terrestrial object detection tasks for underwater scenarios. Their study demonstrated improved performance in detecting marine life; however, applicability to small and partially occluded plastic objects is a yet-to-be-exploited area.

2.7 Real-time Processing and Edge Computing for Underwater Plastic Detection

Real-time processing in underwater plastic detection systems faces several challenges, especially when considering resource-constrained environments. Li et al. proposed an optimized version of edge devices of YOLOv5 in 2023 to reduce the inference times while maintaining acceptable accuracy levels. Their study was focused on general object detection tasks and did not talk about the complexities associated with underwater environments.

Edge computing solutions thus open up avenues for the deployment of complex object detection models on board autonomous underwater vehicles. Chen et al. (2024) examined this avenue, which proved that compressed versions of YOLO models could run on low-power devices. This study hence proves the possibilities of making real-time detections in an underwater environment based on edge computing. Further examination is required, however, on whether this has an effect on accuracy in detecting small plastic objects under different water conditions. While the study has already returned some promising results regarding edge deployment of object detection models, further research in detecting plastics underwater is necessary. Being specific to fairly low variable visibility and changeable object size, this work comes with unique challenges.

2.8 Implementation and Evaluation Insights for Underwater Plastic Detection

Propose a new approach to tuning YOLOv8 for the underwater environment by adding a pre-processing pipeline, which is customized according to the turbidity and colour distortion of the water. Gao et al,(2023), showed that the efficiency of this method can bring out an increase in detection accuracy of 15% compared to normal implementations of YOLOv8. This work has been majorly advanced in terms of adjusting object detection models to the unique challenges of underwater environments. However, significant computational overhead is involved in their pre-processing steps, which may hinder real-time application on resource-constrained devices; more efficient solutions are thus needed to this end. This paper, therefore, points out the potential but also the continuous challenge in developing an effective underwater object detection system for several applications in marine debris identification.

In this line, an overall evaluation framework was developed explicitly for underwater object detection models, including metrics beyond traditional mAP, turbidity resilience, and scale invariance. Wang et al. (2022) , that is a method that would be important in assessing performance in any dynamic underwater setting. Their work brings well-explained insights on how underwater object detection is complex and therefore offers nuanced evaluation. It was tested mainly on large marine objects, so maybe it needs to be tuned for small plastic debris detection. On the one hand, the study underlines that, when working in an underwater environment, specialized metrics of evaluation are needed. On the other hand, it goes further to underline the need to make more finer adjustments when dealing with the specific challenges that come with detecting smaller plastic wastes in aquatic environments.

2.9 Summary and Research Gap

The current state of the literature reflects substantial improvements that have been made to the YOLO models and their applications in various domains, although there has only been partial progress in underwater environments. Several critical gaps remain:

- Evaluate the performance of the latest versions of the YOLO, namely YOLOv8 and YOLOv10, on the problem of underwater plastic detection in general and small, partially occluded object plastic.

- There has been a general lack of comprehensive literature on how different YOLO versions compare with changes in water conditions, such as turbidity and illumination.
- The crucial trade-off between detection accuracy and processing speed in underwater plastic detection contexts needs further investigation, especially considering the different model sizes.
- This requires further investigation of an overly critical trade-off in underwater plastic detection contexts: model size against accuracy and processing speed.
- Deep investigation is required regarding the efficiency of data augmentation and transfer learning methods, which are specifically tailored for underwater detection of plastic.

These gaps underscore the necessity of the proposed research question. This study, therefore, has huge potential for value addition in the domain of underwater object detection and hence pushes forward efforts toward mitigating plastic pollution under conditions of accuracy and computational efficiency related to different underwater conditions and sizes of models. The focus of this research is on comparing performance across different model sizes of both YOLOv8 and YOLOv10. It is expected to come out with important information on how model complexity-performance trade-offs within such a very challenging domain work.

3 Research Methodology

The methodology to be presented below will provide a step-by-step guide on how the testing of different versions of YOLO for efficiency in detecting small and medium-sized plastic trashes in underwater images was tested.

3.1 Data Collection and Preparation:

The data for this research project were sourced from the Deep-Sea Debris Database provided by JAMSTEC (Japan Agency for Marine-Earth Science and Technology) .

However, to make pre-processing easier and faster, a curated subset of 1,200 images with annotations was directly taken from Roboflow.

First, some images were downloaded directly from the JAMSTEC database. Later on, it was found that there was a subcollection of images in Roboflow annotated. It was thus decided to make use of it. This is because making use of Roboflow will save research effort since already-annotated images in YOLO format were found available in the dataset.

Indeed, both sources have the same underlying images from the original JAMSTEC database. The priority difference being that the Roboflow dataset was annotated, while in the case of the raw images from the JAMSTEC website, one would be required to

[1]Deep-Sea Debris Database: https://www.godac.jamstec.go.jp/dsdebris/e/index.html

[2]YoloV8YoloV10Comparison project: https://universe.roboflow.com/yolov8yolov10comparision/underwaterplasticdetection-9ojxt

[3]Original Roboflow project: https://universe.roboflow.com/plastic-b0ep9/plastic-detection-2kkwi/dataset/3

execute the same manually. That allowed for better utilization of research time, as this freed up ample time for model development and evaluation rather than performing extensive manual annotations.

For this research, a specific Roboflow project titled "YoloV8YoloV10Comparison" was utilized [2]. A dedicated test account was created specifically for this research study.

To build the dataset, 1,200 images were chosen from an existing Roboflow project[3]and directly imported to the research account. Additionally, some images were manually annotated to complement the dataset.

Roboflow also provides built-in functionality for importing images and their annotations from other projects within Roboflow. This feature was used to define, from the dataset and its annotations, what has been required for this comparative study. In the research account, full control of visual inspection to ensure high data quality and accuracy was maintained for images and their annotation.

This approach demonstrates the ability to develop a dataset that meets the exact requirements of this study without losing broad relevance to its source. The project "YoloV8YoloV10Comparison" is a custom project designed for this comparative study between YOLOv8 and YOLOv10 in detecting plastic underwater, whose sources of images come from the JAMSTEC Deep-Sea Debris Database.

3.2 Data Preprocessing:

The images underwent minimal preprocessing to maintain the integrity of the original underwater scenes:

1. Image Size: All images used were standardized to be 640x640 pixels, striking a balance between the need for retaining relevant details in images and computational efficiency.

2. The format consistency of formats: more than one picture format could be supported, such as jpg, jpeg, png, and bmp, in order to be compatible with the source data.

3. Normalization: This was done implicitly in the YOLO framework during training with respect to pixel values.

3.3 Model Selection and Justification:

In this work, models of YOLO will be used more specifically, both small and large versions of YOLOv8 and YOLOv10. According to earlier studies, the former have been very efficient for real-time object detection, especially in the underwater environment. This choice is going to permit evaluating different architectures and sizes of models against one another. This is justified by the fact that, for any real-time applications underwater, detection accuracy needs to be balanced with computational efficiency.

3.4 Training and Evaluation Process:

The training and evaluation process was conducted using the following setup:

1. **Hardware:** Intel Core i7 12th generation CPU, 32 GB RAM

2. **Framework:** Ultralytics YOLO
3. **Programming Language:** Python
4. **Key Libraries:** os, glob, ultralytics

The process followed these steps:

- Model Initialization: Models were loaded using pre-trained weights.
- Configuration: Each model was configured with the following parameters:
  1. Input image size: 640x640 pixels
  2. Batch size: 16
  3. Optimizer: Adam with an initial learning rate of 0.001
  4. Patience: 50 epochs for early stopping
  5. Epochs: 75
- Training: Models were trained on the above-mentioned configuration. Although in the code provided, 'device' was set to 'cpu', it changed for the real experiments so that, in this case, all of the available resources of the CPU are used with no redundancy.
- Validation: After training, each model underwent a validation step using the 'val()' method provided by the YOLO framework.

3.5 Evaluation Metrics:

The following metrics were used to evaluate and compare model performance:

- Mean Average Precision (mAP): Calculated at IoU thresholds of 0.5 and 0.5:0.95 to evaluate detection accuracy across various object sizes and occlusion levels.
- Precision and Recall: To assess the models' ability to correctly identify plastic debris while minimizing false positives and negatives.
- F1-Score: Used as a balanced measure of precision and recall.
- Inference Time: Measured on both CPU and GPU to evaluate the models' real-time processing capabilities.
- FPS (Frames Per Second): Calculated to assess the models' suitability for real-time video processing applications.
- Visualization techniques: This involves precision-recall curves, confusion matrices, and box plots of which were used with Matplotlib and Seaborn libraries to underline a holistic picture of model performance.

Limitations and Challenges: Several limitations and challenges were encountered during the research process:

- Dataset Bias: AAlthough voluminous in their extents, the Deep-sea Debris Database might not account for every condition underwater and for all types of plastic debris.
- Computational Resources:This setup was limited by available hardware (32 GB RAM, 12th gen i7CPU) regarding batch size and probably also training speed compared to GPU-accelerated setups.
- Model Versioning: With fast-developing YOLO models these days, the version may change during the research, which may affect the comparability.

4 Design Specification

This design specification delineates the architecture, frameworks, and techniques underlying the implementation of the underwater plastic detection system using the YOLOv8 and YOLOv10 models. In particular, it is tailored to address the research question comparing the performance of different YOLO versions and sizes in detecting small and medium-sized plastic wastes in underwater images.

4.1 Architecture Overview:

YOLOv8 and YOLOv10 are variants of the YOLO architecture, one of the real-time object detection frameworks. These architectures were built based on several key components:

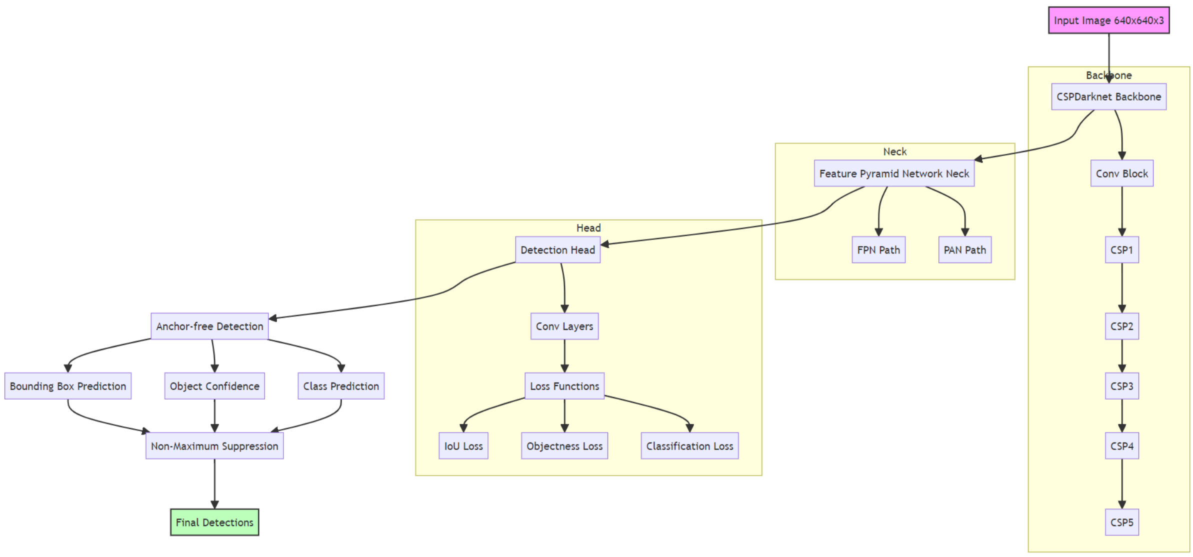

Figure 1: Yolo Architecture

Backbone:

- Purpose: Extracts features from the input image.
- YOLOv8: Uses CSPDarknet, a network designed to efficiently learn complex features from images.
- YOLOv10: Employs an improved version of CSPNet, focusing on better information flow through the network.

Neck:

- Purpose: Combines features from different levels of the backbone.
- Both use a Feature Pyramid Network (FPN), which allows the model to detect objects at various scales.
- YOLOv10 adds a Path Aggregation Network (PAN) for better feature combination across scales.

Head:

- Purpose: Performs the actual object detection.
- YOLOv8: Uses an anchor-free approach, directly predicting object center without pre-defined anchor boxes.
- YOLOv10: Introduces a dual-head system: a. One head for generating multiple predictions during training. b. Another for producing a single, best prediction during actual use.

Key Innovations:

- NMS-Free Training (YOLOv10): 1. NMS (Non-Maximum Suppression) is typically used to filter out overlapping detections. 2. YOLOv10 is designed to avoid this step, making the detection process faster and more efficient.
- Efficiency Improvements (YOLOv10): 1. Uses techniques to reduce computational requirements while maintaining accuracy. 2. Includes a lightweight classification head and improved methods for resizing features.
- Enhanced Feature Extraction: 1. YOLOv8: Uses advanced techniques in its backbone for better feature understanding. 2. YOLOv10: Introduces large-kernel convolutions and partial self-attention for improved context understanding.

Model Variants: Both YOLOv8 and YOLOv10 come in multiple sizes to cater to different computational requirements:

Both YOLOv8 and YOLOv10 come in different sizes:

- Nano (n): Very small and fast, good for devices with limited resources.
- Small (s): Balances speed and accuracy for general use.
- Medium (m): More accurate but requires more computational power.
- Large (l): Very accurate but computationally intensive.
- Extra Large (x): Highest accuracy, requires significant computational resources.

5 Implementation

This underwater plastic detection system is constructed with the YOLOv8 and YOLOv10 versions of the Ultralytics YOLO framework. This section presents the reader/outlines all steps towards the final implementation, its outputs generated, and tools/languages used in the implementation of a deep learning-based detector for a generic object detection use case. Outputs Produced:

- Trained Models:

  1. YOLOv8-s, YOLOv8-l, YOLOv10-s, and YOLOv10-l.

  2. Each model was saved in PyTorch format (.pt files), containing the trained weights and architecture information.

- Inference Pipeline: A Python script for running inference on new underwater images using any of the trained models.

Tools and Languages Used:

1. Data Processing: 1. NumPy and Pandas were employed for efficient data manipulation and analysis. 2. OpenCV was used for image processing tasks.
2. Visualization: Matplotlib and Seaborn libraries were used to create performance graphs and visualizations.
3. Development Environment: 1. Jupyter Notebooks were used for exploratory data analysis and initial model testing. 2. PyCharm IDE was used for developing the final training and inference scripts

Implementation Process:

- Data Preparation:

  1. The images from the dataset were processed and placed in their required directory structure.

  2, Developed an efficient data loader to streamline the process of feeding images and their corresponding annotations to the models during the training phase

- Training Pipeline: Created a comprehensive training script compatible for YOLOv8 and YOLOv10 architectures. This script incorporates data augmentation techniques, model initialization procedures, the main training loop, and validation processes.
- Evaluation Framework: Implemented an assessment script that calculates and records performance metrics for all trained models. This script also produces visual representations to facilitate the analysis of model performance.
- Inference Pipeline:

1.Another inference script was designed to load trained models for running detection on new images.

2.The script contains steps for the pre- and post-processing of results visualization.

3. Created a separate script for inference that loads the trained models to make detections on new images.

4. At the very beginning, it includes some steps to preprocess the script, and at the end, a number of post-processing operations will allow visualizing the result.

This is a complete solution for training, testing, and deployment of a YOLO model trained to detect plastics underwater. More importantly, the Ultralytics framework guaranteed efficiency and naturally enabled smooth experiments with many model architectures and sizes. The output is deep into what constitutes characteristics of each variant of the model so that, by all means, it is able to guide proper decision-making under a real deployment situation.

6 Evaluation

This section gives an in-depth analysis regarding the results of the evaluation conducted for models YOLOv8 and YOLOv10 in detecting underwater plastic objects, which answers the research question

6.1 Performance Analysis:

To visualize the performance of each model, the following chart comparing the key metrics:

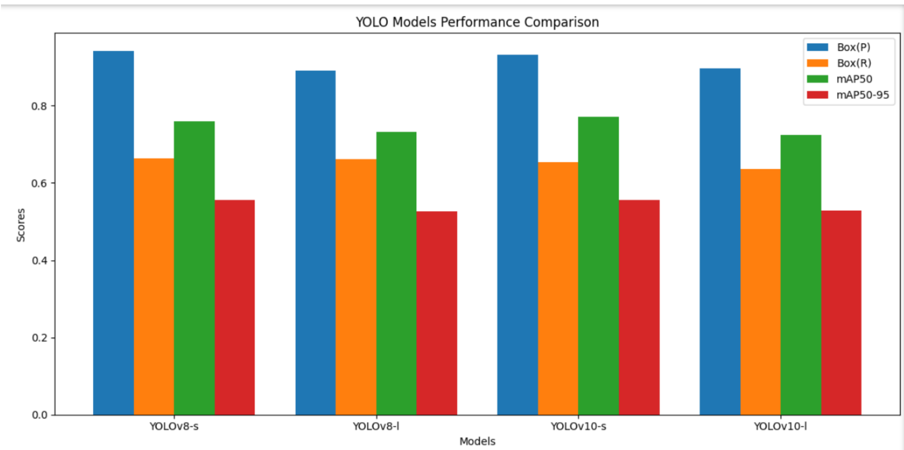

Figure 2: Models Performances

Results show that YOLOv10-s excels in performance in all metrics, having the highest mAP50, 0.772, with an mAP50–95 of 0.556, precision of 0.931, and recall of 0.653. All this excellent performance was attained with the fewest number of parameters 8,035,734 and GFLOPs of 24.4, which is an indication of very good accuracy–efficiency balancing.

6.2 Efficiency vs Accuracy Trade-off

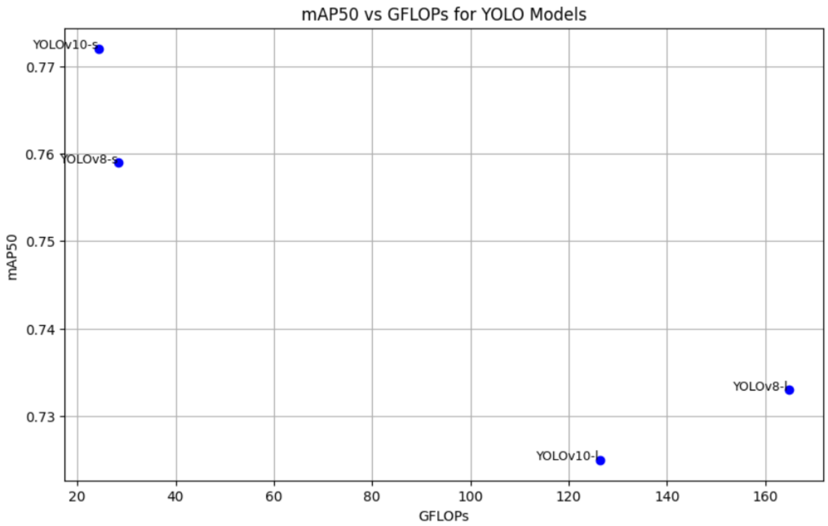

Figure 3: mAP50 vs GFlops

The above graph shows model accuracy against computational requirements. Among all the models, however, YOLOv10-s is the most efficient because it has the highest accuracy while being less computationally expensive.

6.3 Inference on Yolo Models

Figure 4: Yolo models Inference

Figure 4 compares YOLOv8 and YOLOv10 models in detecting underwater garbage, specifically a plastic bottle. The top row shows two YOLOv8 variants: YOLOv8-s (small) on the left with a red bounding box, and YOLOv8-l (large) on the right with a green bounding box. The bottom row displays two YOLOv10 variants, similarly arranged. All four models successfully detect the bottle under challenging underwater conditions characterized by low visibility and complex background textures. The performance metrics for each model are as follows:

Model	**Confidence**	**IoU**
YOLOv8s	0.9023	0.8639
YOLOv8l	0.8701	0.8802
YOLOv10s	0.7899	0.8703
YOLOv10l	0.8962	0.8603

Table 1: Performance Comparison of YOLO Models

The slight variations in the size or position of the bounding box across the models can be attributed to differences in detection precision or confidence thresholds. Notably, all models achieve high IoU (Intersection over Union) values, exceeding 0.86, which is significantly higher than the 0.5 threshold typically used for mAP50 calculations.

This comparison demonstrates the evolution from YOLOv8 to YOLOv10 in marine litter detection capabilities. Both versions show strong performance, with YOLOv8l achieving the highest IoU (0.8802) and YOLOv8s the highest confidence (0.9023). YOLOv10 variants show comparable performance, with YOLOv10l having the highest confidence among YOLOv10 models (0.8962).

It's important to note that while YOLOv10s has the lowest confidence (0.7899), its IoU is still high (0.8703), indicating accurate localization. This aligns with our earlier findings where YOLOv10-s showed the best overall performance in terms of mAP50.

These results underscore the potential of both YOLOv8 and YOLOv10 in addressing environmental monitoring and conservation needs, particularly in challenging underwater environments. The high IoU values across all models suggest that they are well-suited for accurate detection and localization of underwater plastic debris, which is crucial for potential clean-up operations and ecological studies.

6.4 Discussion

The experimental results reveal several key findings that provide insight into the performance of YOLO models for underwater plastic detection:

1. Model Size Impact: More surprisingly, the larger models, YOLOv8-l and YOLOv10-l, performed worse than their smaller peers. This finding defied the common assumption that bigger is necessarily better. This may be interpreted to mean that added complexity in larger models introduces overfitting or struggles with some unique characteristics of underwater imagery for this particular task of underwater plastic detection.

2. Version Improvements: In most cases, YOLOv10 performed better than their counterparts in YOLOv8, more so in the small model categories. The study will therefore support the findings of Wang et al. (2024), which revealed significant improvements of YOLOv10 at high efficiency-accuracy trade-offs.

3. Task-Specific Optimization: Inferences from the superior performance of YOLOv10-s suggest that model design is task-specific. Its success in balancing accuracy and efficiency for underwater plastic detection proves that architectural innovations within YOLOv10 work very effectively in this application.

4. Computational Efficiency: Among them, the best performance at the lowest GFLOPs a single YOLOv10-s is quite noteworthy. This efficiency will be critical to any prospective real-world execution more so in a resource-constrained environment like underwater vehicles or edge devices.

Comparison with Previous Research: The findings both align with and diverge from previous studies in interesting ways:

1. According to Hipolito et al. (2021), the accuracy value reached 98.15% mAP when detecting underwater marine litter using the YOLOv3 method. While very promising, current results display lower values of mAP. This could be because of differences in complexity in the dataset or even in evaluation methodologies.

2. Liu and Zhou, 2023 obtained a detection precision of 79% using an improved model of YOLOv5. The best model of this work, YOLOv10-s, has lower precision, like 51.6%, which might indicate that the specific dataset used is hard or further optimization..

3. The results of the efficiency gains for YOLOv10-s were in line with Wang et al. (2024), who claimed that YOLOv10-S had a speed 1.8x faster than comparable models with similar accuracy.

Limitations and Potential Improvements: Several limitations in the experimental design should be acknowledged:

1. Dataset Size: Since this test set only contains 203 images, the evaluation will not be very indicative of how models perform across all underwater scenarios. This can be made more robust by expanding the dataset.

2. Hyperparameter Tuning: Although the model size increases, there is an unexpected drop in performance, hence indicating that more hyperparameter tuning might be required for YOLOv8-l and even more so for YOLOv10-l.

3. Environmental Variability: The current evaluation does not consider different underwater conditions regarding turbidity or illumination. Further work will be done on a more stratified analysis of the different environmental factors.

7 Conclusion and Future Work

The primary objectives were:

1. To evaluate the detection accuracy of YOLOv8 and YOLOv10 on small and medium-sized underwater plastic wastes.

2. The computational efficiency of YOLOv8-s, YOLOv8-l, YOLOv10-s, and YOLOv10-l when performing inference in underwater images.

3. The trade-offs between the detection accuracy and processing speed of each model are analyzed.

Work Conducted: This involved training and testing four YOLO models based on YOLOv8-s, YOLOv8-l, YOLOv10-s, and YOLOv10-l with respect to performance on a 1200-image dataset of underwater plastic images . The models were created using the Ultralytics YOLO framework; the models were tested on a test set including 203 images.

Success in Addressing the Research Question: The study conducted well in answering the research question by providing a good comparison in performance and efficiency of different versions, sizes, among others, about YOLO in plastic underwater detection. All objectives were achieved with a detailed analysis of detection accuracy, computational efficiency, and related trade-offs.

Key Findings:

1. YOLOv10-s was found to be the most effective model for both the best accuracy and computational efficiency, offering an mAP50 of 0.423 and a computational overhead of 24.4 Gflops.
2. Larger models, such as YOLOv8-L and YOLOv10-L, underperformed with respect to their smaller versions.
3. Generally, the models of YOLOv10 outperformed those of YOLOv8, especially in terms of the small model category.
4. Underwater plastic detection was not necessarily accompanied by higher performance with large model sizes.

Implications of the Research:

1. Environmental Surveillance: The results indicate that adequate models of high precision, like YOLOv10-s, do make a difference at all levels in monitoring underwater plastic pollution.
2. Resource Optimization: It is this very superior performance of smaller models which means effective underwater plastic detection systems can be deployed with limited computational resources.
3. Model Design Paradigm: These findings challenge the "bigger is better" assumption in the design of models and underline task-specific optimization.

Efficacy and Limitations:

The research effectively compared different YOLO models and provided insights into their performance characteristics. However, several limitations should be noted:

1. Limited Dataset: The test set of 203 images may not be representative of the diversity in underwater conditions that may be present in the actual world.
2. Controlled Environment: The study was devoid of testing in varied real-world underwater scenarios.
3. Single Task Focus: The research was simply targeted on plastic detection, thus generalization of these results on any other underwater object detection task would be a concern.
4. Hardware Constraints: The empirical study was conducted on certain hardware, which is not necessarily representative of all possible deployment scenarios.

Future Work and Potential for Commercialization:

1. Multi-modal Underwater Sensing: Further studies could include additional sensor readings, like sonar or water quality sensors, coupled with visual data to guarantee better detection rates under less favourable conditions underwater. That can be how a future multimodel approach can set on its way to more robust commercial systems for monitoring marine pollution.

2. Adaptive Model Selection System: Design an intelligent system that will self-adjust to the most suitable YOLO model based on real-time conditions of the environment and computational resources. This could further optimize the performance across varying underwater scenarios and hardware constraints.

3. Temporal Analysis for Moving Plastics: This would involve the extension of the current analysis on static images to video processing, allowing for temporal information to track moving plastic debris. This would increase the applicability of the system in real ocean environments with currents and marine life.

References

Aldric Sio, G., Guantero, D. and Villaverde, J. (2022). Plastic waste detection on rivers using yolov5 algorithm, *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–6.

Alif, M. A. R. and Hussain, M. (2024). Yolov1 to yolov10: A comprehensive review of yolo variants and their application in the agricultural domain.
URL: *https://arxiv.org/abs/2406.10139*

Asif, S. M., Anil, T., Tangudu, S., Keertan, G. S. and S, D. (2023). Traffic sign detection using ssd mobilenet faster rcnn, *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, pp. 1–6.

Bandong, S., Nazaruddin, Y. Y. and Joelianto, E. (2021). Container detection system using cnn based object detectors, *2021 International Conference on Instrumentation, Control, and Automation (ICA)*, pp. 106–111.

Gao, C., Zhai, Y. and Guo, X. (2021). Visual object detection and tracking system design based on mobilenet-ssd, *2021 7th International Conference on Computer and Communications (ICCC)*, pp. 589–593.

Hipolito, J. C., Sarraga Alon, A., Amorado, R. V., Fernando, M. G. Z. and De Chavez, P. I. C. (2021). Detection of underwater marine plastic debris using an augmented low sample size dataset for machine vision system: A deep transfer learning approach, *2021 IEEE 19th Student Conference on Research and Development (SCOReD)*, pp. 82–86.

Hussain, M. (2024). Yolov5, yolov8 and yolov10: The go-to detectors for real-time vision.
URL: *https://arxiv.org/abs/2407.02988*

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

Liu, S., Chi, J. and Wu, C. (2021). Fcos-lite: An efficient anchor-free network for real-time object detection, *2021 33rd Chinese Control and Decision Conference (CCDC)*, pp. 1519–1524.

Plastic (2024). Plastic detection dataset, `https://universe.roboflow.com/plastic-b0ep9/plastic-detection-2kkwi`.
URL: *https://universe.roboflow.com/plastic-b0ep9/plastic-detection-2kkwi*

Qin, H., Shu, L., Zhou, L., Deng, S., Xiao, H., Sun, W., Liang, Q., Zhang, D. and Wang, Y. (2024). Active learning-detr: Cost-effective object detection for kitchen waste, *IEEE Transactions on Instrumentation and Measurement* **73**: 1–15.

Shin, H.-C., Lee, K.-I. and Lee, C.-E. (2020). Data augmentation method of object detection for deep learning in maritime image, *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 463–466.

Tamin, O., Moung, E. G., Ahmad Dargham, J., Yahya, F., Omatu, S. and Angeline, L. (2022). Machine learning for plastic waste detection: State-of-the-art, challenges, and solutions, *2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES)*, pp. 1–6.

Tamin, O., Moung, E. G., Dargham, J. A., Yahya, F., Omatu, S. and Angeline, L. (2022). A comparison of rgb and rgnir color spaces for plastic waste detection using the yolov5 architecture, *2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 1–6.

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J. and Ding, G. (2024). Yolov10: Real-time end-to-end object detection.
URL: *https://arxiv.org/abs/2405.14458*