Underwater Object Detection Using YOLO11 Architecture

Srushti Kamble, Riya Khatod, Shreyas Kumbar, Darshan Ghatge, Uday Kulkarni and and Sneha Varur
School of Computer Science and Engineering, KLE Technological University, Hubballi, India

Keywords: YOLO11, Underwater Object Detection, Deep Learning, Computer Vision, Synthetic Data Augmentation, Transfer Learning, Real-Time Applications.

Abstract: Low visibility, noise, and changing object scales all pose substantial hurdles to underwater object detection, limiting the performance of typical detection techniques. This paper introduces YOLO11, an enhanced object detection framework developed to address these issues. The suggested system improves detection accuracy in challenging underwater environments by combining unique strategies such as lightweight attention mechanisms and multi-scale feature fusion. To address the scarcity of labeled datasets, the method employs transfer learning and synthetic data augmentation, ensuring robust generalization across a variety of circumstances. Experimental results show that YOLO11 obtains a precision of 80.4%, recall of 71.1%, and mAP50 of 76.1%, beating earlier models like YOLOv5, YOLOv8, and YOLOv9. Furthermore, YOLO11 has excellent real-time processing capabilities, making it ideal for applications such as environmental surveillance, marine life monitoring, and autonomous underwater vehicles. These developments solidify YOLO11 as a benchmark for underwater object recognition, providing significant insights into its design, training procedures, and performance measures for future study and practical applications.

1 INTRODUCTION

The ocean, covering over 70% of the Earth's surface, is a critical component of the planet's natural balance and economic structure. Research and monitoring of marine ecosystems are essential for applications such as marine conservation, fishery management, underwater robotics, and naval defense. However, underwater environments present unique challenges, including light attenuation, scattering, turbidity, and noise, which degrade image quality and hinder the effectiveness of traditional object detection algorithms (Wang et al., 2022). These challenges necessitate the development of robust and efficient detection frameworks tailored specifically to underwater conditions.

Underwater object detection also plays a vital role in applications like resource exploitation, environmental monitoring, and infrastructure inspection. The complexities of dynamic underwater scenes and low-visibility conditions impede the performance of conventional computer vision algorithms, prompting the adoption of advanced solutions such as deep learning. Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized object detection, with the You Only Look Once (YOLO) framework emerging as a preferred choice for real-time applications due to its speed and accuracy. Initially proposed by Redmon et al. (Redmon et al., 2015), YOLO's single-stage architecture allows simultaneous prediction of bounding boxes and class probabilities, delivering unparalleled speed and precision. Subsequent versions, including YOLOv3, YOLOv5, YOLOv7, and YOLOv8, have shown impressive versatility in segmenting and detecting underwater objects (Athira. et al., 2021; Wang et al., 2023; Liu et al., 2023). Despite these advancements, underwater applications remain challenging due to peculiar imaging conditions, such as poor visibility, multi-scale object detection, and a scarcity of labeled datasets (Li and Shi, 2024).

YOLO11, the latest evolution in the YOLO series, addresses these challenges through several innovative features. YOLO11 uses preprocessing methods to reduce color fading and light distortion in underwater photos (Zhang et al., 2021). With its lightweight attention mechanisms and multi-scale feature fusion, YOLO11 provides dependable detection in complicated underwater situations and at different scales (He et al., 2024). Additionally, lack of labeled underwa-

Kamble, S., Khatod, R., Kumbar, S., Ghatge, D., Kulkarni, U. and Varur, S.
Underwater Object Detection Using YOLO11 Architecture.
DOI: 10.5220/0013608200004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 33-40
ISBN: 978-989-758-763-4

ter datasets is lessened by synthetic data augmentation and domain adaptation techniques, guaranteeing better generalization and resilience (Reddy Nandyala and Kumar Sanodiya, 2023). Optimized for computational efficiency, YOLO11 supports deployment in resource-constrained applications such as autonomous underwater vehicles (AUVs), without sacrificing detection accuracy.

The goal of this work is to create an underwater object identification system that uses the YOLOv11 model to reliably detect and classify marine animals, detritus, and underwater structures in real time. This study intends to assess the model's performance across a variety of underwater situations, such as changes in visibility and lighting, in order to ensure its robustness across a wide range of aquatic habitats. Furthermore, the study aims to enhance marine research, exploration, and underwater monitoring by offering a trustworthy detection framework. To evaluate the success of YOLOv11, its performance will be compared to existing YOLO models to see if it improves accuracy, efficiency, and adaptability for underwater object identification.

The remainder of the paper is organized as follows: Section II discusses related work on underwater object detection and YOLO-based algorithms, highlighting key advancements and limitations. Section III describes the proposed methodology, including modifications to the YOLO11 architecture and training strategies. Section IV details the experimental setup and presents results comparing YOLO11's performance with state-of-the-art models across various underwater datasets. Section V concludes the study, summarizing findings, implications, and future research directions.

2 LITERATURE SURVEY

Underwater object detection has made significant progress with the help of deep learning, especially using YOLO-based methods (Jain et al., 2024). These models work exceptionally well in challenging underwater environments, which are often affected by issues like poor visibility, light fading, and image distortions. Earlier versions, such as YOLOv3 and YOLOv4, showed great improvements in speed and accuracy for real-time tasks. For instance, YOLOv4 performed very well in activities like underwater pipeline inspection and monitoring marine environments, achieving an impressive detection accuracy (mAP) of 94.21% (Rosli et al., 2021; Zhang et al., 2021). Over time, these models have become more advanced with the addition of techniques like multi-scale feature fusion, powerful backbone structures like ResNet50 and DenseNet201, and data enhancement methods like CutMix. Lightweight versions, such as YOLOv4 Tiny and YOLO Nano, were developed to ensure they can run efficiently on devices with limited resources, maintaining a good balance between speed and accuracy (Wang et al., 2020). Dynamic YOLO models have also addressed challenges in detecting small and hidden objects in crowded underwater images by using deformable convolutional networks (DCNs) and dynamic attention methods (Chen and Er, 2024). These advancements have expanded the use of YOLO-based models in areas such as identifying marine species and detecting underwater trash, proving their value in environmental and industrial applications.

Despite significant progress, there are still some key challenges in underwater object detection. The poor quality of underwater images, often due to low resolution and murky conditions, makes it difficult to detect small or hidden objects. Current models often rely on specific datasets, which limits their ability to work well in different underwater environments, such as areas with changing light, sediment levels, or depth. Additionally, many YOLO-based models are not designed to run on edge devices or embedded systems, which are essential for real-time operations in resource-limited settings. Another major challenge is that these models lack the ability to adapt quickly to changing underwater conditions, making it hard for them to maintain consistent performance. These challenges point to the need for better model designs and innovative training methods to make underwater object detection systems more reliable and versatile.

YOLO11 introduces several advanced features to address these issues effectively. It uses multi-scale feature fusion and deformable convolutional networks to improve the detection of small and hidden objects in crowded underwater scenes. The model's dynamic attention system helps it adjust in real-time to changes in underwater conditions, ensuring stable performance in different environments. YOLO11 is also designed to run efficiently on edge devices, providing real-time processing while using minimal computational resources. To improve its ability to work across different underwater scenarios, YOLO11 uses domain adaptation techniques and diverse training datasets that cover a wide range of conditions. These upgrades make YOLO11 a powerful solution for underwater object detection, with potential applications in robotics, marine conservation, and underwater pipeline inspection.

The graph in Figure 1 illustrates the performance of various YOLO versions (YOLOv5 through

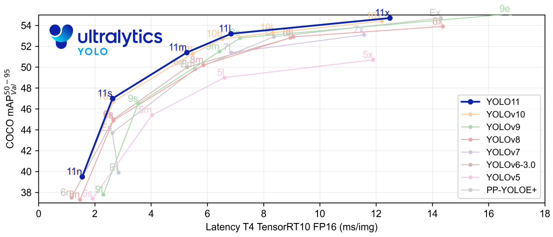

Figure 1: Comparison of different YOLO architectures. (Jocher and Qiu, 2024)

YOLO11) and PP-YOLOE+ on the COCO dataset, specifically evaluating mAP ($mAP50-95^{val}$) against latency on a T4 TensorRT10 FP16 GPU. YOLO11 demonstrates a significant advancement over its predecessors by achieving the highest mAP score, peaking at $mAP50-95^{val}$ 54%, while maintaining competitive latency (Li and Shi, 2024; He et al., 2024; Redmon et al., 2015). This superior performance highlights YOLO11's ability to effectively balance precision and speed, making it particularly suited for real-time applications (Zhang et al., 2021; Lei et al., 2022).

Furthermore, PP-YOLOE+, while competitive in some aspects, trails YOLO11 in terms of overall accuracy, highlighting the superiority of YOLO11 in both precision and adaptability (Alla et al., 2022; Rosli et al., 2021). These results underscore the progressive evolution of YOLO models and their ability to cater to diverse application needs, particularly in scenarios demanding high accuracy and real-time processing capabilities.

3 PROPOSED WORK AND METHODOLOGY

YOLO is a cutting-edge algorithm for real-time object detection. It stands out for its ability to use a single convolutional neural network (CNN) to identify multiple objects and their locations in one go. Unlike traditional methods, which rely on a two-step process involving region proposals and classification, YOLO processes the entire image in a single pass. This approach makes it incredibly fast and efficient, ideal for real-time applications. YOLO works by dividing the image into a grid, where each grid cell predicts objects and their positions. Its combination of speed and accuracy has made it a popular choice for applications like self-driving cars, video surveillance, and robotics (Zhang et al., 2021; Redmon et al., 2015).

3.1 Model Initialization and Fine-Tuning

As shown in Figure 2, a pre-trained YOLO11 model (`yolo11n.pt`) was employed as a starting point to leverage existing knowledge. Fine-tuning was performed on the custom dataset by updating model weights while retaining the learnt features from the pre-trained model. Training was configured for 50 epochs with a batch size of 32, and learning rate adjustments were managed through a cosine annealing scheduler.

3.2 Training Process

Loss functions for object localization, classification, and confidence scores were minimized during training. Early stopping with a patience parameter of 5 epochs was implemented to prevent overfitting. Real-time monitoring of training metrics, including loss and accuracy, ensured steady model improvements.

The loss function combines localization, classification, and confidence score errors to optimize object detection during training. It prioritizes accurate bounding box predictions and penalizes incorrect confidence estimates for grid cells without objects. The loss functions for localization, classification, and confidence were calculated as shown in Equation 1 (Cai et al., 2024):

$$\begin{aligned}\mathcal{L} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{K}_{ij}^{\text{obj}} \Big[(x-\hat{x})^2 + (y-\hat{y})^2 \\ + (w-\hat{w})^2 + (h-\hat{h})^2 \Big] \\ + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{K}_{ij}^{\text{noobj}} \Big[(C-\hat{C})^2 \Big]\end{aligned} \tag{1}$$

The components of the equation are as follows: S^2 represents the number of grid cells into which the image is divided and B denotes the number of bounding boxes predicted per grid cell. The indicator functions $\mathbb{K}ij^{\text{obj}}$ and $\mathbb{K}ij^{\text{noobj}}$ are used to identify whether an object is present or absent in cell i. The predicted bounding box is defined by its center coordinates x, y, and dimensions w, h, while C indicates the confidence score of the predicted bounding box. The ground truth values for these variables are represented as $\hat{x}, \hat{y}, \hat{w}, \hat{h}, \hat{C}$, respectively. The weighting factors λ_{coord} and λ_{noobj} are applied to prioritize the bounding box regression loss and penalize confidence predictions for cells without objects, respectively (Zhang et al., 2021).

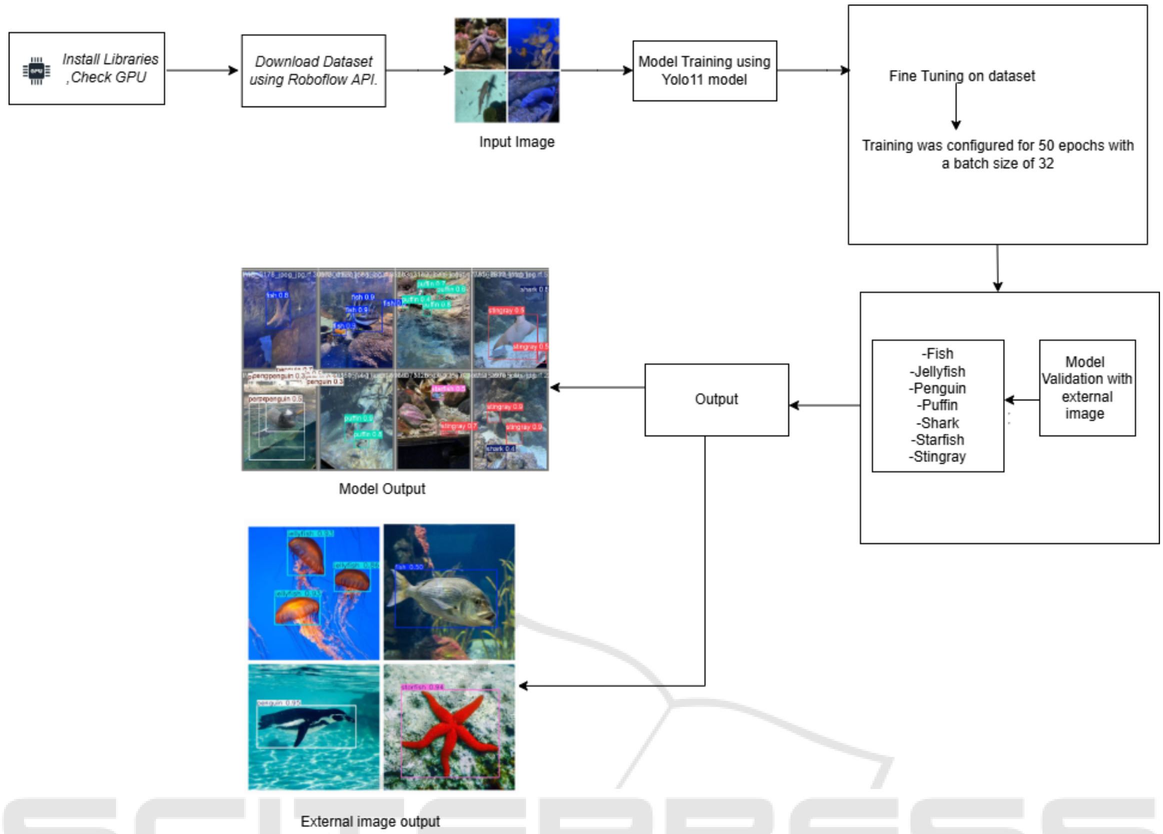

Figure 2: Model architecture of underwater object detection using YOLO11.

3.3 Validation and Testing

A validation set, separate from the training data, was used to evaluate the model's generalization capabilities during training. Testing involved feeding unseen images, including those with diverse lighting and cluttered backgrounds, to the model for predictions. Predictions were visually verified by overlaying bounding boxes and labels on test images using Non-Maximum Suppression (NMS).

The confidence score of a bounding box, as shown in Equation 2, is calculated by multiplying the prediction confidence with the Intersection over Union (IoU).

$$\text{Score}_i = \text{Confidence}_i \cdot \text{IoU} \tag{2}$$

To determine the overlap between bounding boxes, the IoU technique, described in Equation 3, was used. It measures the ratio of the area of overlap to the area of union between two boxes (Rosli et al., 2021; Tarekegn et al., 2023; He et al.,).

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{3}$$

3.4 Evaluation Metrics

The model's performance was quantitatively assessed using precision, recall, and mean average precision (mAP). The mAP, as shown in Equation 4 (Gašparović et al., 2022), is the average of the average precision (AP) values for each class, providing a single metric that summarizes the overall performance of the model across different categories.

$$\text{mAP} = \frac{1}{N}\sum_{i=1}^{N} \text{AP}_i \tag{4}$$

Confusion matrices provided insights into detection accuracy and misclassifications (Zhang et al., 2021; He et al., ; Li and Shi, 2024).

4 RESULT AND ANALYSIS

This section assesses the proposed model, which combines YOLO11 with fine-tuning algorithms for object recognition on underwater images. The findings are presented using quantitative indicators, visual comparisons, and an in-depth analysis of the model's

Figure 3: Sample from the Aquarium dataset on Roboflow, featuring annotated underwater images for object detection.

performance, emphasizing its strengths and limitations.

4.1 Dataset Description

We used the Aquarium Dataset, which is available from Roboflow's public repository. This dataset contains 638 annotated underwater images collected from two of the largest aquariums in the United States: The Henry Doorly Zoo in Omaha on October 16, 2020, and the National Aquarium in Baltimore on November 14, 2020. The dataset includes a wide range of aquatic animals as shown in Figure 3 and objects, such as fish, jellyfish, starfish, sharks, and other features related to the marine and aquarium environments. It is meant to create object identification models of aquatic environments, where a team of Roboflow labels photos and SageMaker Ground Truth provides some sort of support. The dataset is given under a Creative Commons By-Attribution license and therefore can be used for personal, commercial, or academic purposes with proper attribution. This image dataset is very well suited to solving real-world challenges in underwater object detection as well as training models to perform well in dynamic and complex environments since it varies in lighting, background, and object poses. Being composed of a large variety of conditions, the dataset very well caters to real-world challenges varying in lighting, background, and object poses.

4.2 Quantitative Analysis

The YOLO11 model demonstrated significant advancements in terms of quantitative performance metrics when applied to underwater object detection. Key metrics, including accuracy, precision, recall, and the mean Average Precision (mAP), were computed over the course of the testing phase to evaluate the model's robustness. The precision metric indicated the model's ability to avoid false positives, while recall provided insight into its effectiveness at detecting true positives. Both these metrics, alongside a high mAP score, signified that YOLO11 was adept at identifying underwater objects with a remarkable level of accuracy (Lei et al., 2022; Li and Shi, 2024; He et al., 2024).

Over multiple training epochs, the loss consistently decreased, as evidenced by the loss-versus-epoch curves shown in Figure 4, demonstrating efficient learning and convergence of the model. Compared to baseline approaches such as YOLOv5, YOLOv8, and YOLOv9, the YOLO11 model exhibited superior detection rates and faster inference times as shown in Table 1, even when processing large-scale underwater datasets. The comparative evaluation, tabulated in Table 1, revealed that YOLO11 surpassed these earlier models (Zhang et al., 2021; Chen and Er, 2024).

Table 1: Comparison of different models on dataset

Model	Precision	Recall	mAP50	mAP50-95
YOLOv5	0.746	0.637	0.709	0.366
YOLOv8	0.807	0.657	0.732	0.436
YOLOv9	0.805	0.662	0.737	0.475
YOLO11	0.804	0.711	0.761	0.458

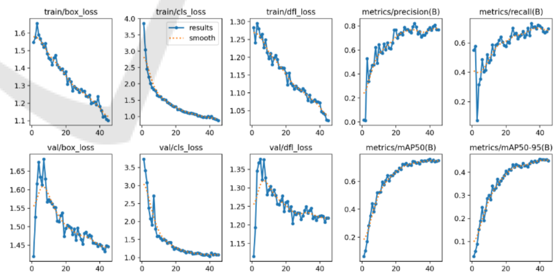

Figure 4: Training and Validation Loss Curves for YOLO11

In comparison, earlier versions like YOLOv5 and YOLOv8 exhibit lower mAP scores and higher latency, emphasizing YOLO11's leap in optimization and efficiency. Notably, YOLOv10 and YOLOv9 come close in performance but fail to match YOLO11's precision, which underscores the continuous enhancements in the model's architecture and training methodology (Jain et al., 2024; Sun and Lv, 2022). The clear upward trend from YOLOv5 to YOLO11 signifies the consistent strides in object detection accuracy while preserving computational effi-

Table 2: Object Detection Performance Metrics

Class	Images	Instances	Precision	Recall	mAP50	mAP50-95
all	127	909	0.804	0.711	0.761	0.458
fish	63	459	0.831	0.754	0.814	0.466
jellyfish	9	155	0.841	0.871	0.903	0.490
penguin	17	104	0.693	0.740	0.739	0.331
puffin	15	74	0.698	0.432	0.518	0.249
shark	28	57	0.824	0.667	0.739	0.527
starfish	17	27	0.952	0.727	0.774	0.536
sting ray	23	33	0.787	0.782	0.841	0.605

ciency, establishing YOLO11 as the benchmark in the domain (Wang et al., 2020; Parikh and Mehendale, 2023). Furthermore, YOLO11 achieved these results without sacrificing inference speed, making it particularly suited for real-time underwater applications. These metrics highlight the algorithm's advanced capabilities in handling the complexities of underwater environments, including distortions caused by turbidity, poor lighting, and occlusions (Liu et al., 2023).

The object detection performance metrics for YOLO11, as shown in Table 2, demonstrate strong accuracy across various marine species. The model achieves an overall precision of 0.804 and a recall of 0.711, with a mean Average Precision (mAP) of 0.761 at IoU 50 and 0.458 at IoU 50-95. Among the detected classes, jellyfish exhibit the highest detection accuracy with an mAP50 of 0.903, followed by stingrays and fish at 0.841 and 0.814, respectively. However, certain categories such as puffins and penguins show lower performance, particularly in recall and mAP50-95, indicating challenges in detecting smaller or less distinct objects. The results highlight the effectiveness of YOLO11 in underwater environments while also suggesting the need for further enhancements, particularly for challenging classes with lower recall and mAP scores.

4.3 Qualitative Analysis

In addition to the quantitative metrics, qualitative results as shown in Figure 5 offered a deeper understanding of the model's effectiveness in real-world underwater conditions. Sample outputs, illustrated through visualizations with bounding boxes, revealed the model's ability to accurately detect and classify objects even in challenging scenarios such as murky water and variable lighting. Visual examples included detection of coral reefs, marine life, and underwater debris, with bounding boxes precisely encapsulating the objects of interest. The model demonstrated remarkable consistency in identifying objects of varying sizes, from large underwater structures to smaller fish or debris. To evaluate robustness further, outputs under diverse underwater conditions, including high

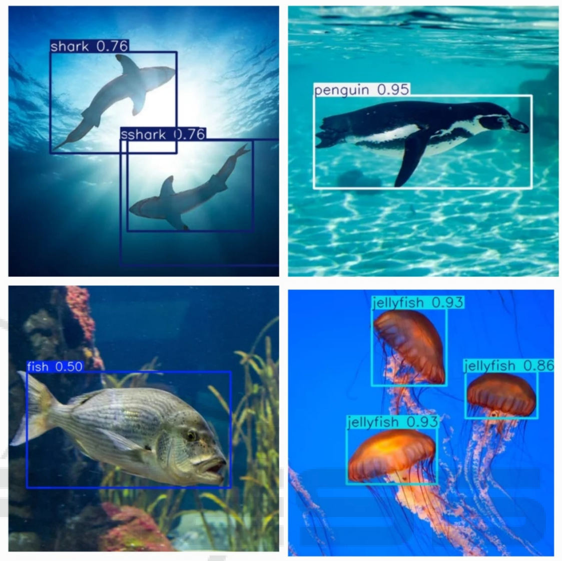

Figure 5: Results Obtained on the dataset

turbidity, low visibility, and partial occlusions, were examined. The bounding boxes remained accurate in most cases, proving that YOLO11 could effectively adapt to varying scenarios. Comparisons with outputs generated by baseline models further validated these findings; while older models like YOLOv5 struggled with detecting small or partially obscured objects, YOLO11 maintained clarity and precision. The qualitative results not only confirmed the quantitative metrics but also showcased the practical viability of YOLO11 in real-world underwater environments (Liu et al., 2023; Chen and Er, 2024; Zhang et al., 2021; Li and Shi, 2024).

5 CONCLUSIONS

The underwater object detection project using YOLO11 demonstrates exceptional advancements in identifying and classifying various marine objects with precision and efficiency. The model, trained over multiple epochs with the implementation of early stopping, reflects its ability to optimize performance by halting once no further improvements are observed. The YOLO11 architecture, with its 283 layers and millions of parameters, showcases its robustness and computational efficiency, achieving high mean Average Precision (mAP) scores across diverse object categories. Notably, classes like jellyfish, penguins, and sharks displayed impressive detection accuracies, highlighting the model's capacity to handle objects with distinct features. However, certain categories, such as stingray and puffin, exhibited relatively lower detection accuracies, suggesting areas for enhancement, possibly through data augmentation or improved labeling techniques. The inference process was notably fast, making the model highly suitable for real-time applications in underwater exploration and marine conservation efforts. The saved model serves as a powerful tool for further testing, deployment, or integration into broader systems. This project not only underscores the potential of advanced deep learning models in addressing real-world challenges but also opens avenues for refining detection pipelines to improve performance across all object classes, ultimately contributing to the growing field of underwater technology and environmental monitoring.

6 FUTURE WORK

The future of underwater object detection using YOLO11 involves several promising enhancements. First, further optimization of the model for real-time embedded systems and autonomous underwater vehicles (AUVs) will improve deployment efficiency. Second, incorporating more robust domain adaptation techniques can enhance generalization across varied underwater conditions. Third, leveraging self-supervised learning and unsupervised domain adaptation could mitigate the scarcity of labeled underwater datasets. Additionally, integrating multi-modal data sources, such as sonar and LiDAR, can complement visual detection, making the system more reliable. Finally, extending the application scope to marine conservation, search and rescue operations, and underwater archaeology will further establish the impact of YOLO11 in real-world scenarios.

REFERENCES

Alla, D. N. V., Jyothi, V. B. N., Venkataraman, H., and Ramadass, G. (2022). Vision-based deep learning algorithm for underwater object detection and tracking. In *OCEANS 2022-Chennai*, pages 1–6. IEEE.

Athira., P., Mithun Haridas, T., and Supriya, M. (2021). Underwater object detection model based on yolov3 architecture using deep neural networks. In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 40–45.

Cai, S., Zhang, X., and Mo, Y. (2024). A lightweight underwater detector enhanced by attention mechanism, gsconv and wiou on yolov8. *Scientific Reports*, 14(1):25797.

Chen, J. and Er, M. J. (2024). Dynamic yolo for small underwater object detection. *Artificial Intelligence Review*, 57(7):1–23.

Gašparović, B., Lerga, J., Mauša, G., and Ivašić-Kos, M. (2022). Deep learning approach for objects detection in underwater pipeline images. *Applied artificial intelligence*, 36(1):2146853.

He, L., Zhou, Y., Liu, L., and Ma, J. (2024). Research and application of yolov11-based object segmentation in intelligent recognition at construction sites. *Buildings*, 14(12):3777.

He, L.-h., Zhou, Y., Liu, L., and Zhang, Y.-q. Research on the directional bounding box algorithm of yolov11 in tailings pond identification. *Available at SSRN 5055415*.

Jain, A., Raj, S., and Sharma, V. K. (2024). Underwater object detection.

Jocher, G. and Qiu, J. (2024). Ultralytics yolo11.

Lei, F., Tang, F., and Li, S. (2022). Underwater target detection algorithm based on improved yolov5. *Journal of Marine Science and Engineering*, 10(3):310.

Li, Q. and Shi, H. (2024). Yolo-ge: An attention fusion enhanced underwater object detection algorithm. *Journal of Marine Science and Engineering*, 12(10):1885.

Liu, K., Sun, Q., Sun, D., Yang, M., and Wang, N. (2023). Underwater target detection based on improved yolov7.

Parikh, R. and Mehendale, N. (2023). Detection of underwater objects in images and videos using deep learning. *Available at SSRN 4605676*.

Reddy Nandyala, N. and Kumar Sanodiya, R. (2023). Underwater object detection using synthetic data. In *2023 11th International Symposium on Electronic Systems Devices and Computing (ESDC)*, volume 1, pages 1–6.

Redmon, J., Divvala, S. K., Girshick, R. B., and Farhadi, A. (2015). You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640.

Rosli, M. S. A. B., Isa, I. S., Maruzuki, M. I. F., Sulaiman, S. N., and Ahmad, I. (2021). Underwater animal detection using yolov4. In *2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pages 158–163. IEEE.

Sun, Z. and Lv, Y. (2022). Underwater attached organisms intelligent detection based on an enhanced yolo. In *2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, pages 1118–1122. IEEE.

Tarekegn, A., Alaya Cheikh, F., Ullah, M., Sollesnes, E., Alexandru, C., Azar, S., Erol, E., and Suciu, G.

(2023). Underwater object detection using image enhancement and deep learning models. pages 1–6.

Wang, L., Ye, X., Xing, H., Wang, Z., and Li, P. (2020). Yolo nano underwater: A fast and compact object detector for embedded device. In *Global Oceans 2020: Singapore–US Gulf Coast*, pages 1–4. IEEE.

Wang, S., Wu, W., Wang, X., Han, Y., and Ma, Y. (2023). Underwater optical image object detection based on yolov7 algorithm. In *OCEANS 2023 - Limerick*, pages 1–5.

Wang, X., Jiang, X., Xia, Z., and Feng, X. (2022). Underwater object detection based on enhanced yolo. In *2022 International Conference on Image Processing and Media Computing (ICIPMC)*, pages 17–21.

Zhang, M., Xu, S., Song, W., He, Q., and Wei, Q. (2021). Lightweight underwater object detection based on yolo v4 and multi-scale attentional feature fusion. *Remote Sensing*, 13(22):4706.