International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Underwater Marine Life Study Using Yolo V8

Ms. Neha Singh[1], Palak Saini[2], Deepali Yadav[3]

[1]Professor IGDTUW
[2,3]Student IGDTUW

Abstract:
Underwater target detection, object classification, and image segmentation play an important role in ocean exploration and marine life studies for which the improvement of relevant technology is of much practical significance. Although existing target detection, image classification, and segmentation algorithms have achieved excellent performance on land they often fail to achieve satisfactory outcomes of detection and classification when in the underwater environment. In this paper, one of the most advanced target detection algorithms, YOLO v8 (You Only Look Once), was first applied in the underwater environment before being improved by combining it with some methods characteristic of the underwater environment. The state-of-the-art backbone and neck architectures and TC-YOLO/SAM were treated as the basic backbone network of YOLO v8, which makes the network suitable for underwater images .

Keywords: YOLO v8, TC-YOLO, SAM algorithm, state-of-the-art backbone, and neck architectures, Underwater target detection, object classification, image segmentation

1. Introduction

Oceans are 3/4 th part of our earth and have a large amount of unexplored life species, reservoirs, minerals, oil, and more. Underwater marine life encompasses a rich diversity of species, many of which remain undocumented due to the challenges of deep sea exploration. when we go deeper into oceans, this is very dangerous to human lives. There are many cases happening in the water. [4] The application of machine learning can revolutionize the study of these species by analyzing vast amounts of data collected from underwater sensors, autonomous under-water vehicles (AUVs), and remotely operated vehicles (ROVs). This paper outlines the potential of ML in advancing our understanding of marine ecosystems.[16][17] YOLO v8 appears to be a combination of design elements and techniques from recent object detection and instance segmentation models such as YOLO X, YOLO v6, YOLO v7, and PPYOLOE.[3][6][9] These modifications are aimed at enhancing the model's performance, accuracy, and deployment friendliness in various application scenarios. It's important to note that the effectiveness of these improvements would depend on the specific use case and data set.

This is used for various purposes such as monitoring endangered species, illegal fishing detection, and ecosystem preservation.[2][4] Very less discoveries are going on in the underwater areas. ML-based underwater life classification can help to discover and classify previously unknown underwater species and ML allows for exploration of remote and inaccessible areas.

2. Proposed model

The following three methods were proposed to improve the performance of a modified YOLO v5 network for underwater object detection:

1. TC-YOLO is integrated with YOLO V8 to improve the quality of the network to get improved images with better resolutions and enhanced picture quality.
2. SAM is used as an integration algorithm to provide better real-time segmentation of images.

2.1. Data-set

We used a test and train data-set approach to pre-train our model. We used a custom data set to pre-train the model. The images in the data set are from different sources from the internet and contain different environments and conditions under the water. The data set has 4 categories i.e.; fishes, plants, rocks, and seabed. There are 500 images in the data set and the data set is divided into 346 images of training and 154 images of testing data sets.

2.2. YOLO-V8

YOLO v8 is the latest iteration in the YOLO series of real-time object detectors, offering cutting-edge performance in terms of accuracy and speed. It includes five variants i.e.; YOLO v8n, YOLO v8s, YOLO v8m, YOLO v8l, and YOLO v8x whose sizes and other parameters increase gradually as they increase.[7][11][13] An image is first processed by the backbone for feature extraction, followed by the neck for feature fusion, and finally is outputted as the head for the prediction of objects. This working model is integrated with other models salient features to create a better function-able, fast, accurate, and fuzzy image detection, segmentation, and classification.

2.3. TC-YOLO

The modified object detection network is called TC-YOLO. It includes one Transformer module in the backbone and three CA modules in the neck. Transformer and CA modules were combined with a cross-stage partial (CSP) structure to make the attention blocks CSP-TR and CSP-CA[1][6]. Similar to most detection algorithms, the detection head was placed after the neck. The network generated provides maps at three scales as the input to the detection head. The feature map at each scale corresponded to three anchors, so in total, nine anchors were obtained by clustering the data set. The placements of these attention blocks and the overall structure of TC-YOLO are shown in Figure 1.

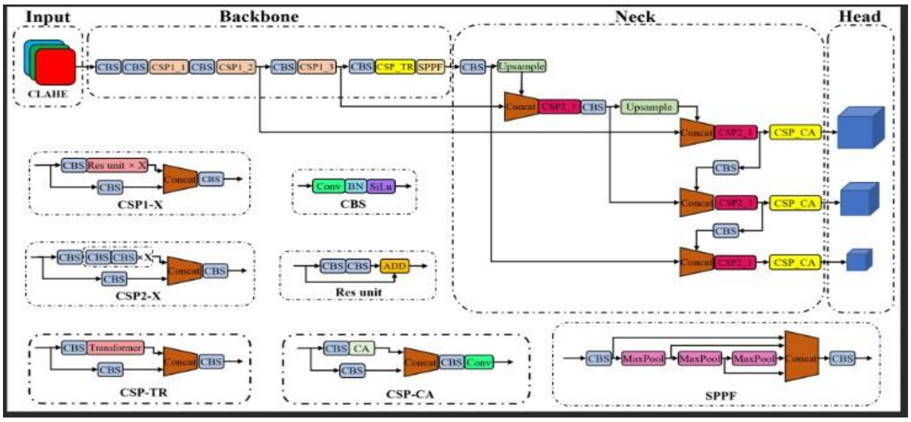


Figure 1. TC-YOLO architecture

The TC-YOLO works parallel with the YOLO-v8 architecture by becoming a part of it. The input to the section gets collected at CLAHE where each image's pixel RGBs are collected and enhanced. Then each pixel is flattened. The CSP structure, reduced the number of channels of the input feature by half without any loss of local information. The flattened feature was defined as the patch and position, encoding was carried out by passing patches through a fully connected layer. Position-encoding data and the original patch were then added as the input of the Transformer encoder. The CA modules before prediction can effectively summarize the global information for different size features after extraction and fusion. It reduces the computational cost by using the CSP-CA module with it. Then the results are up-sampled and output is provided.[7][12]

2.4. SAM model

The Aqua SAM decouples the segment image task into two sequential stages:

1. all-instance segmentation and prompt-guided selection. The first stage uses YOLO v8-seg to produce the segmentation of all instances in the image.
2. In the second stage, it outputs the region-of-interest to the corresponding to the desk screen(prompt).[15]

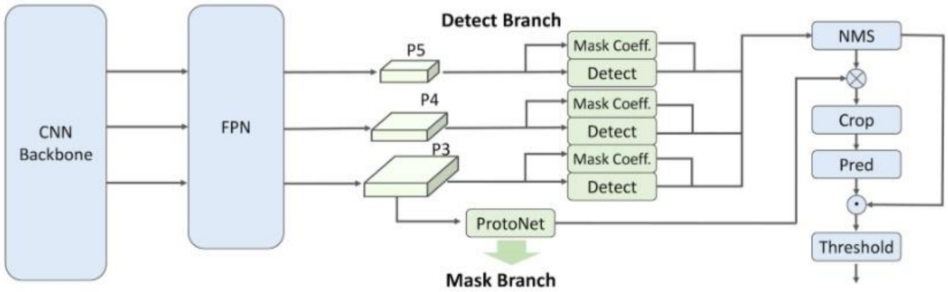
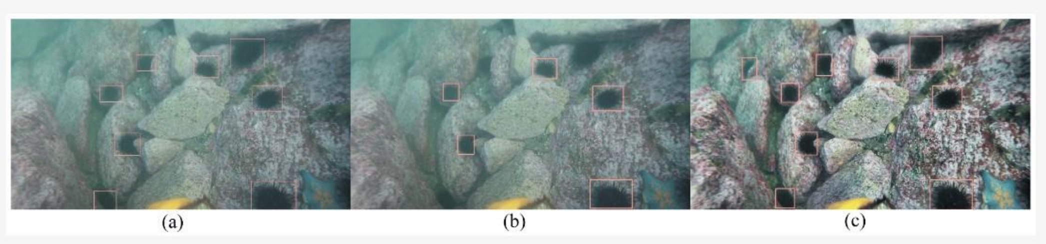


Figure 2. SAM model integrated with YOLO-V8

The forerunner approaches are based on fully convolutional networks (FCNs) that use the semi CNN-based models for feature extraction[2]. The encoded information is then given to a decoder network that learns to classify each pixel then it gradually up-samples the low-dimensional features by a series of de-convolution layers and eventually generates the pixel-wise labels. Then these label are gradually send to train and test the data set by reading the mask coefficient and detecting each pixel by its parameters. the sample pixels are also sent to a prototype network which is a kind of library used to keep the mask coefficients and other parameters. Then the detected data after getting sampled with the pre-trained data is sent to the predictor of YOLO-V8 to predict the values and respond accordingly. YOLO-V8clf then takes the mask coefficients and convolved outputs of CLAHE, to classify the objects in the image.

3. Results

In this paper, we proposed a new approach for detecting, segmenting, and classifying underwater images. Underwater object detection is difficult due to poor image quality, limited computational capacity, and underwater targets that are often small, dense, over- lapped, and obscured. YOLO v8s with the proposed TC-YOLO is only 0.8 percent larger than YOLO v8s in size and finely surpasses the state-of-the-art in underwater detection tasks. Integration of YOLO v8-seg with Aqua SAM achieves an average Dice Similarity Coefficient (DSC) of 7.13 percent. The overall precision and recall increased by 3.2 percent and 5.7 percent respectively. The proposed approach can detect dense and small targets more well: The mAPIoU=0.5:0.95 improved by 7.1 percent.



111

Figure 3. computational comparison of each stage ;(a) Actual image, (b) YOLO-V8 ,(c)YOLO-V8 with YOLO-TC

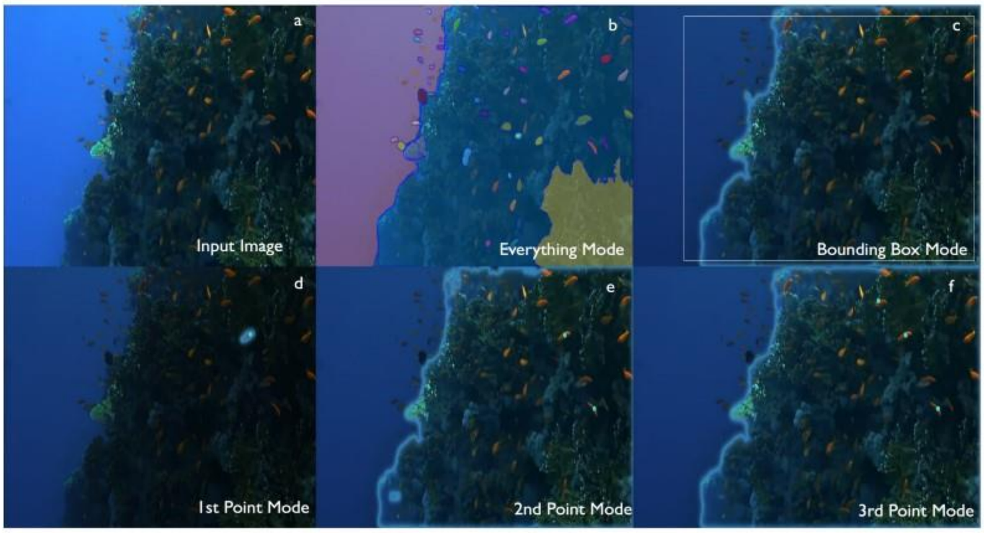


Figure 4. Integration of SAM model with YOLO-V8

The pictorial classification and breakout of each step of convolution and model work can be seen in Figures 3 and 4.

Acknowledgments:

We extend our deepest gratitude to all those who played a pivotal role in the realization of this research. Our sincere thanks go to the members of the Indira Gandhi Delhi Technical University For Women, whose insights and expertise in underwater ecology and image analysis significantly enriched this study. We are particularly indebted to Mrs. Neha Singh, whose guidance and support were invaluable throughout the research process.

The computational resources provided by our Professor enabled us to leverage the capabilities of YOLO v8 for marine life detection and recognition, and for this, we are immensely thankful.

We appreciate the funding and resources provided, without which this study would not have been possible. Our acknowledgments would be incomplete without mentioning the local marine community and diving teams whose local knowledge and expertise were instrumental in conducting our underwater surveys.

Abbreviations

The following abbreviations are used in this manuscript:

YOLOYou Only Look Once

SAMSegment Anything Model

CNNConvolutional Neural Network

CSP Cross stage partial structure
RGB Red Green Blue
CA Convolved Array stucture

References

1. Thao NgoGia, Yinghao Li, Daopeng Jin, Jinyuan Guo, Jiang Li Qirong Tang; Real-Time Sea Cucumber Detection Based on YOLOv4-Tiny and Transfer Learning Using Data Augmentation,ICSI 2021.
2. Jalal, A.; Salman, A.; Mian, A.; Shortis, M.; Shafait, F. Fish detection and species classification in underwater environments using deep learning with temporal information. Ecol. Inform. 2020,57, 101088.
3. Jäger, J.; Rodner, E.; Denzler, J.; Wolff, V.; Fricke-Neuderth, K. SeaCLEF 2016: Object Proposal Classification for Fish Detection in Underwater Videos. In Proceedings of the CLEF (Working Notes), Évora, Portugal, 5–8 September 2016;
4. B.W. McLaren et al. A small no-take marine sanctuary provides consistent protection for small-bodied by-catch species, but not for large-bodied, high-risk species J. Exp. Mar. Biol.Ecol.(2015)
5. Kang, T.; Kim, B.; Jung, Y. Deep learning platform architecture for monitoring image-based real-time construction site equipment and worker. KIBIM Mag. 2021,
6. Lei, F.; Tang, F.; Li, S. Underwater Target Detection Algorithm Based on Improved YOLOv5. J. Mar. Sci. Eng. 2022, 10, 310.
7. Lee, Y.; Kim, Y. Comparison of CNN and YOLO for Object Detection. J. Semicond. Disp. Technol. 2020,
8. Ferdous, M.; Ahsan, S.M.M. PPE detector: A YOLO-based architecture to detect personal protective equipment (PPE) for construction sites. PeerJ Comput. Sci. 2022, 8, e999.
9. Kim, K.; Kim, K.; Jeong, S. Application of YOLO v5 and v8 for Recognition of Safety Risk Factors at Construction Sites. Sustainability 2023, 15, 15179.
10. Raja, M.C. Utilizing YOLO 8x Models, Deep Learning-Based Head protector Detection for Construction Workers. Res. Sq. 2023. under review.
11. Huang, Z.J., Li, L.T., Krizek, G.C. and Sun, L.H.(2023) Research on Traffic Sign Detection Based on Improved YOLOv8. Journal of Computer and Communications, 11, 226-232.
12. Y. Yao, G. Yang, X. Sun, and L. Li, “Detecting video frame-rate up- conversion based on periodic properties of edge-intensity,” J. Inf. Secur. Appl., vol. 26, Feb. 2016.
13. Raja, M.C. Utilizing YOLO 8x Models, Deep Learning-Based Head protector Detection for Construction Workers. Res. Sq. 2023. under review.
14. Lee, Y.; Kim, Y. Comparison of CNN and YOLO for Object Detection. J. Semicond. Disp. Technol. 2020,
15. Lei, F.; Tang, F.; Li, S. Underwater Target Detection Algorithm Based on Improved YOLOv5. J. Mar. Sci. Eng. 2022.
16. Kong, S.; Fang, X.; Chen, X.; Wu, Z.; Yu, J. A real-time underwater robotic visual tracking strategy based on image restoration and kernelized correlation filters. In Proceedings of the Chinese Control and Decision Conference, Shenyang, China, 9–11 June 2018; pp. 6436–6441.
17. Jing, C.; Lin, Z.; Li, J. Detection and tracking of an underwater target using the combination of a particle filter and track-before-detect. In Proceedings of the OCEANS—Shanghai, Shanghai, China, 10–13 April 2016; pp. 1–5