An Approach for Fish Detection in Underwater Images

Laura A. Martinho⋆ Odalisio L. S. Neto⋆ João M. B. Calvalcanti⋆ José L. S. Pio⋆ Felipe G. Oliveira†

⋆Institute of Computing (ICOMP), Universidade Federal do Amazonas (UFAM).
Manaus, Amazonas, Brazil.
†Institute of Exact Sciences and Technology (ICET), Universidade Federal do Amazonas (UFAM).
Itacoatiara, Amazonas, Brazil.
{lam2, oldsn, john, josepio}@icomp.ufam.edu.br, felipeoliveira@ufam.edu.br

***Abstract*—Underwater images are widely used for understanding subaquatic environments. However, underwater images are severely degraded by light absorption and scattering, as it propagates in water during image acquisition causing color distortion, low contrast and noise. These problems can interfere in underwater vision tasks, such as recognition and detection. In this paper we propose an approach for fish detection in underwater environments. In order to achieve this goal, the proposed method is composed by two main steps: $i)$ Image Restoration, processing the underwater images to enhance the image quality; and $ii)$ Fish Detection, to identify the presence of fish in underwater images. Additionally, in this paper we introduce an underwater image dataset with the presence of fish. Through the experimental process using the proposed dataset, the obtained results demonstrate the precision and robustness of the proposed approach, achieving accuracy of 98.04% in the fish detection task.**

***Index Terms*—Underwater images, Fish detection, Underwater image restoration, YOLO-NAS**

I. Introduction

Exploring the underwater world has been a subject of interest in the last years, both in the academic and commercial scenarios. However, capturing clear and vibrant images in this environment has been a challenging problem. Underwater images are severely degraded due to the effects of light attenuation, scattering and color distortion, causing difficulties to the understanding of the visual information in the image [1].

In order to recognize objects in underwater images, it is important to recover lost or significantly distorted visual information at the image acquisition moment. Image enhancement is a strategy widely used to solve the aforementioned challenges, restoring details, colors and textures that are often obscured or distorted in underwater environments. In this sense, the employment of advanced algorithms and techniques may compensate the effects of light scattering and absorption. Thereby, it can result in images that reveal fine details, enhanced contrast and accurate color representation, enabling different applications in underwater domains [2].

Different applications can be found concerning the understanding and processing of underwater images. In marine biology and ecology, it assists studying and monitoring underwater ecosystems, identifying species and unraveling their behaviors. Defense and surveillance sectors employ underwater image restoration to enhance the quality of underwater images obtained from sonar systems, underwater cameras and remotely operated vehicles (ROVs), aiding in the detection of underwater mines, wreckage and potential threats. Additionally, industries such as underwater robotics, offshore engineering and underwater exploration depend on restored images for navigation, object recognition and situational awareness [3] [4] [5].

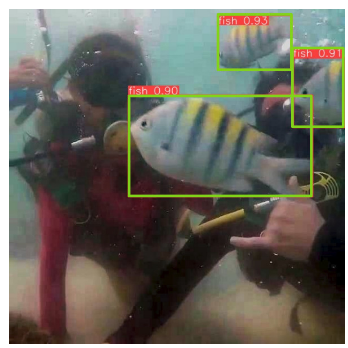


Fig. 1: Real-world underwater image in fish detection process.

Taking the applications above in account, it is possible to verify the relevance of techniques for the detection of underwater events and objects, such as fishes. Figure 2 illustrates an example of this sort of image. In this sense, this paper presents an approach for fish detection in underwater environments, even in acquisition conditions with low lighting exposure. We propose here a methodology comprising two main steps: $i)$ Image Restoration; and $ii)$ Fish Detection. For the image restoration step, it is applied a sequence of intensity transformations, in order to enhance the image quality. Meanwhile, for the fish detection it is employed the YOLO-NAS, in order to identify the presence of different types of fish in image.

This paper is organized as follows. In Section 2, the related work on underwater image processing and understanding is introduced. The proposed approach for fish detection in underwater environments is described in Section 3. The experiments and discussions are elaborated in Section 4. Finally, the paper ends with the conclusions in Section 5.

II. Related work

In recent years, there has been significant research on the enhancement and restoration of underwater images [6] [7]. Several work have proposed techniques focused on understanding and enhancing the visual appearance of underwater images, which have been extensively discussed in the literature.

Enhancing the quality of underwater images is a paramount task in the field of underwater imaging, especially due to the properties of the sub-aquatic environment which degrade the images. The mentioned properties involve light scattering, color distortion, poor visibility, reduced contrast and loss of details, which represent a challenge for developing new approaches and algorithms that can improve the quality of underwater images [8].

Several work tackled the problem of underwater image restoration through hybrid approaches, using fusion of different techniques. In order to reach this goal, the mentioned techniques use channels histogram adaptation, optimized contrast algorithm and histogram stretching based on the red channel [9]. The authors used different filters to reduce noise and blur in the image, removing undesired noise and enhancing the overall image perception.

Water absorbs and scatters light, leading to reduced visibility and distorted depth perception. To solve this problem, domain adaptation approaches, color compensation and depth estimation techniques were developed in [1]. The authors proposed a domain adaptation framework based on transfer learning for underwater image enhancement. They used a domain adaptation module for style transfer and a domain adaptation module for image enhancement.

In addition, with advances in deep learning and neural networks, machine learning-based approaches have achieved significant results in image restoration. Several work proposed learning methods to accomplish underwater image restoration, such as deep learning, reinforcement learning [2] and generative adversarial networks, as approached in [10]. Some work still combine contrastive learning and generative adversarial networks and released a large-scale real underwater image dataset to support both paired and unpaired training modules.

Furthermore, underwater environments are home to a diverse range of flora, fauna, and geological formations, each with distinct patterns and structures. As result, there has been a growing emphasis on underwater object detection, together the need to identify objects submerged in aquatic scenes [3].

Recognizing patterns in underwater scenarios holds significant value in ecological studies, marine biology and underwater exploration. Researchers have been actively investigating methodologies to address these challenges. For instance, a non-destructive and rapid detection method has been proposed for determining fry feeding status using a shallow underwater imaging system and a lightweight deep learning framework [4]. Additionally, a deep learning approach has been developed to enhance and restore images and videos captured in underwater scenes by incorporating an underwater scene prior [11]. Furthermore, an attention mechanism has been introduced into the feature extraction network to enhance the feature expression of fish and improve the model's robustness [12]. These advancements demonstrate the ongoing efforts to leverage deep learning techniques in underwater imagery.

In the context of underwater environments, the real-time object detection system YOLO can efficiently detect and classify objects present in images or videos. Recent research work have been proposing many methods using learning models for sub-aquatic object detection tasks. In [5] the authors worked on a deep learning-based fish detection model called YOLO-Fish that can differentiate fish from the seabed and other fish types in challenging underwater environments. Another approach explored the pruned neural network technique using MobileNetv2-YOLOv2 for underwater object detection [13].

The YOLO hability to provide fast and accurate detection results ensure its strenght in marine life exploration. Lightweight underwater object detection approach [14] that uses the MobileNetv2 architecture along with YOLOv4 algorithm achieve a balance between accuracies, improving the YOLO detection method [7]. By enhancing fish recognition in underwater images using the cumulative mean of the YOLO network, [15] conclude that their proposed method suggest that it can be useful for conservation efforts in marine ecosystems.

Several research work are performed to monitor fish in the underwater environment. An automatic sorting system [6] was developed to tackle the challenges of increasing food demand and the threat of food scarcity in the future. In [15] the YOLO-Fish method, a deep learning based fish detection model is proposed using two models, enhancing and fixing the issue of upsampling step sizes to reduce the misdetection of tiny fish. Researchers proposed in [16] a method for detection of fish with disease, using an improved YOLOv5 network for aquaculture, changing to new convolutional kernels and convolutional block attention added to the YOLOv5 algorithm.

Our main contribution in this work is the proposed deep learning-based approach for fish detection in underwater environments. In addiction, we highlight the analysis and comparison of a great amount of state-of-the-art techniques for underwater image restoration and object detection, demonstrating its robustness.

III. Proposed Approach

Restoring and understanding visual information in underwater images are a paramount for different applications. Many applications involve underwater ecosystems, such as: archaeological sites exploration, study marine life and search and rescue in deep waters. Moreover, underwater image restoration and understanding plays a crucial role in underwater exploration, surveillance, marine biology, and industrial applications like offshore engineering and oil exploration [17].

This paper presents an approach to detect different types of fish in underwater environments. For this, the proposed approach consists of two main steps: i) Image Restoration; and ii) Fish Detection, as can be observed in Figure 2. The mentioned steps will be further detailed in the next subsections.

For the fish detection procedure, initially, the acquired underwater image (I) undergoes the Image Restoration stage, for enhancing the underwater image quality, improving the clarity, visibility, contrast, noise and color degradation. Next, in the Fish Detection stage, the restored underwater image (I^r) goes through an object detection technique, trained for the fish detection operation. As output, different types of fishes are detected in the underwater images (I^f).

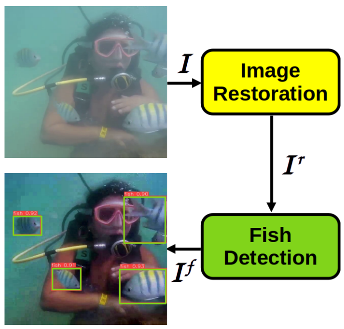


Fig. 2: Proposed methodology for the fish detection process from underwater images.

A. Image Restoration

Initially, the raw images (I) are acquired in the Red, Green and Blue (RGB) color model. Then, the degraded underwater images (raw) are restored obtaining enhanced underwater images (I^r), compensating the turbidity, blurring, and other degrading effects. For this purpose, the proposed solution is based on the fusion of intensity transformation techniques to improve the quality of underwater images. It consists of six techniques in a two-layer-based method, that takes a single image as input and the sequence of techniques is performed.

In the first layer, gamma correction and edge enhancement methods are applied, using the Unsharp Mask. In the second layer, gamma correction, edge enhancement and Contrast-limited Adaptive Histogram Equalization (CLAHE) are applied. The final processing step takes the two images, generated in the previous layers, and combine them using the linear blending as image fusion technique. Afterwards, the generated image undergoes brightness and contrast adjustment, finally achieving the final output of the proposed method.

Below we provide an overview of the mentioned techniques, while further in-depth information can be found in the reference [9].

1) Color Correction: The first technique performed is a color correction algorithm, based in [18], that applies histogram stretching in each color channel (R, G, B), represented by equation 1:

$$f(v) = (v - I^{min})(max - min)/(I^{max} - I^{min}) + min \quad (1)$$

2) Gamma Correction: The second technique is the gamma correction of underwater images. In this stage, the image pixel intensities are scaled from the range [0, 255] to [0, 1.0], as follows in equation 2:

$$I_G = I^{(1/g)} \quad (2)$$

3) Unsharp Enhancement: As the final operation in each layer, the unsharp enhancement method is applied in underwater images. For this, the Unsharp Mask filter is used to enhance edges in underwater images.

$$h(v) = I(v) - I_{smooth}(v) \quad (3)$$

where v is the gray value to be transformed, I is an underwater image and I_{smooth} corresponds to the smoothed underwater image.

4) CLAHE: In this stage, the algorithm Contrast-Limited Adaptive Histogram Equalization (CLAHE) is applied, as in equation 4:

$$I_{he} = [I^{max} - I^{min}] * P(I) + I^{max} \quad (4)$$

5) Fusion: Let I_1 and I_2 be the processed underwater images in each layer of techniques, they are combined using the linear blending, as image fusion technique, represented by the k function. Thereby, the visual features in I_1 and I_2 are fused in order to improve the quality and visibility of degraded underwater images, as seen in equation 5:

$$k(v) = (1 - \alpha)I_1(v) + \alpha I_2(v) \quad (5)$$

6) Contrast and Brightness Adjustment: The generated fused underwater image undergoes brightness and contrast adjustment, and finally yields the final output of the proposed approach. The parameters $\beta > 0$ and γ control the contrast and brightness, respectively. The equation below presents the brightness and contrast transformation:

$$I^r = \beta * I + \gamma \quad (6)$$

The generated fused underwater image undergoes brightness and contrast adjustment, and finally yields the final restored image.

B. Fish Detection

From the restored images (I^r) it is possible to detect fish present in underwater images, identifying the correct position where the fish is, resulting in images with a bounding box for every detected fish (I^f).

For the fish detection we used the You Only Look Once (YOLO) algorithm, in version NAS, called YOLO-NAS. The YOLO algorithm is designed for efficient object detection by accurately predicting the presence and precise location of objects within an image. It approaches the detection stage as a supervised regression learning process, resulting in fast and reliable performance. The YOLO model is built upon a CNN network, enabling it to predict multiple bounding boxes and assign class probabilities to those boxes. This ability to generalize objects contributes to its effectiveness in object detection [3].

The YOLO algorithm divides the input image into an SxS grid, assigning the responsibility of object detection to the grid cell containing the center of the object. Each grid cell predicts B bounding boxes, where B represents the total number of boxes. These predictions include confidence scores, which indicate the model level of certainty regarding the presence of an object within the box, as well as its confidence in the accuracy of the prediction for that specific box. Additionally, the predictions encompass class probabilities (C), providing insights into the likelihood of the object belonging to different predefined classes [5].

YOLO-NAS is an advanced extension of YOLO that introduces the concept of automated neural architecture search. Thereby, allows the model to learn designing its own neural architecture instead of relying on pre-defined architectures, being able to automatically find the best network configuration for the object detection task. Different from the original architecture, the extension searches through a space of possible architectures, where different configurations are evaluated and refined through an optimization process. This search process is guided by an objective function, which takes into account performance metrics such as model accuracy and efficiency. By performing the automated search, it is able to discover optimized architectures, adapted to the specific features of the object detection task, achieving superior performance compared to conventional architectures, while still maintaining the core features of YOLO.

Our proposed YOLO-NAS model consists of backbone, neck and head stages. The backbone, usually is a convolutional neural network (CNN), which extracts important features from the image at different scales. Our backbone step is composed by seven convolutional layers. The neck refines these features, enhancing spatial and semantic information. Our neck step is composed by six convolutional layers. Lastly, the head uses these refined features to make object detection predictions. Our head step is composed by nine convolutional layers. During the training stage, the Adam optimization algorithm is employed, with a learning rate set to 5×10^{-4}. The training process runs for 300 epochs, and a batch size of 16 is used for efficient processing and learning. It is important to mention that the YOLO-NAS model was used in this work due to significant results obtained in object detection regarding underwater images [6] [7].

IV. Experiments

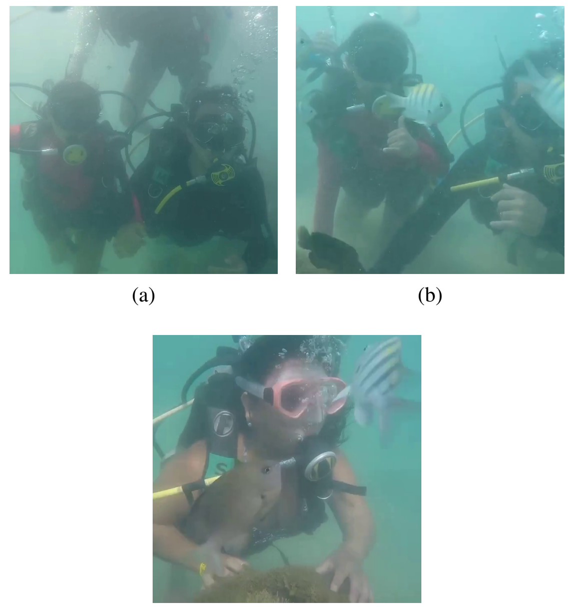
In this section the proposed approach for fish detection from images of underwater environments is evaluated demonstrating its robustness.

A. Experimental Setup

The experiments were carried out using a Lenovo laptop with an Intel® Core^{TM} i7-10750H CPU @ 2.60GHz, 16 GB DDR4-2133 main memory and NVIDIA® GeForce® RTX 3060 6 GB GDDR6. Furthermore, the OpenCV and Tensorflow frameworks were used to support the development of the proposed approach for fish detection in underwater environment. For image acquisition we used a GoPro Hero 10 Black camera, with $4k$ resolution, Exposure time of $1/120s$, Aperture value of $F2.4$ and Focal length of $2.7mm$.

B. Experimental Dataset

In this paper we propose a new dataset of underwater images, acquired in an ocean environment with the presence of different types of fish. The dataset is called Fish in Underwater Images (FUI). The FUI dataset consists of 4773 underwater images with and without the presence of fish. Figure 3 presents samples of the proposed image dataset, where it is possible to verify the high level of turbidity and scattering in underwater images.



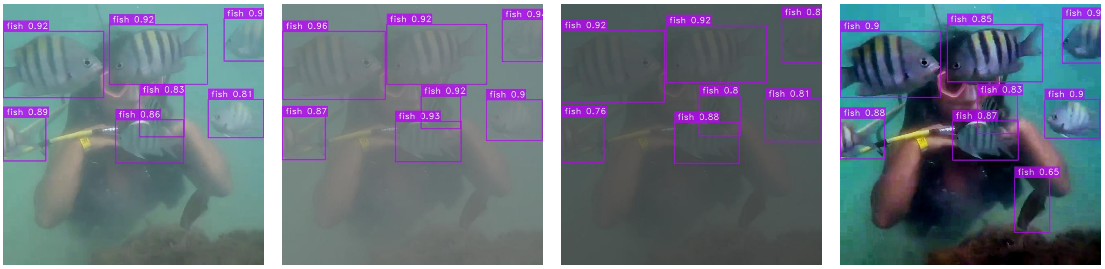
(a) (b) (c)

Fig. 3: Examples of underwater images which comprises the FUI dataset. Figures 3a, 3b and 3c represent scenes from FUI dataset.

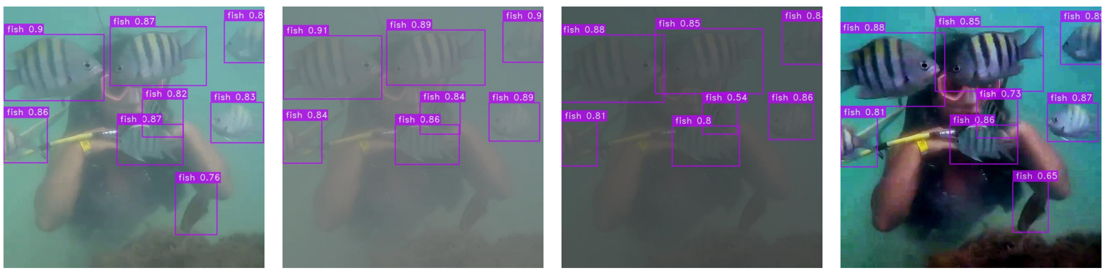
C. Fish Detection Accuracy Assessment: Qualitative Analysis

In this experiment we intend to evaluate the efficiency of the fish detection process, in a qualitative analysis. For this, we compared the proposed approach, which combines the UIET underwater image restoration technique [9] and the YOLO-NAS object detection technique, with a set of comparison approaches.

The comparison approaches are based on: i) the applying of the YOLOv8 object detection technique directly to the raw underwater image; ii) the applying of the YOLO-NAS object detection technique directly to the raw underwater image; iii) the applying of the YOLOX object detection technique directly to the raw underwater image; iv) the combining of the IBLA underwater image restoration technique [19] and the YOLOv8 object detection technique; v) the combining of the IBLA underwater image restoration technique and the YOLO-NAS object detection technique; vi) the combining



(a) Raw + YOLO-NAS (b) IBLA + YOLO-NAS (c) UDCP + YOLO-NAS (d) UIET + YOLO-NAS (Ours)



(e) Raw + YOLOX (f) IBLA + YOLOX (g) UDCP + YOLOX (h) UIET + YOLOX

(i) Raw + YOLOv8 (j) IBLA + YOLOv8 (k) UDCP + YOLOv8 (l) UIET + YOLOv8

Fig. 4: Qualitative comparison on FUI dataset underwater images. Figures 4a, 4e and 4i, in the first column, were not restored (raw images) and were combined with the YOLO-NAS, YOLOX and YOLOv8, respectively. Figures 4b, 4f and 4j, in the second column, were restored using the IBLA technique, with the YOLO-NAS, YOLOX and YOLOv8, respectively. Figures 4c, 4g and 4k, in the third column, were restored using the UDCP technique, with the YOLO-NAS, YOLOX and YOLOv8, respectively. Meanwhile, Figures 4d, 4h and 4l, in the forth column, were restored using the UIET technique, with the YOLO-NAS, YOLOX and YOLOv8, respectively.

of the IBLA underwater image restoration technique and the YOLOX object detection technique; vii) the combining of the UDCP underwater image restoration technique [20] and the YOLOv8 object detection technique; $viii$) the combining of the UDCP underwater image restoration technique and the YOLO-NAS object detection technique; ix) the combining of the UDCP underwater image restoration technique and the YOLOX object detection technique; x) the combining of the UIET underwater image restoration technique and the YOLOv8 object detection technique; and xi) the combining of the UIET underwater image restoration technique and the YOLOX object detection technique.

In Figure 4 we can observe some results obtained from the proposed approach and the comparison approaches, for fish detection from underwater images. From Figure 4 it is also possible to verify the necessity of restoring the underwater images for the efficient employment of the object detection techniques. Some fish in the underwater images can not be in clear and visible condition. In this sense, the underwater image restoration techniques can contribute providing a clearer restored image, allowing accurate fish detection. Additionally, we can notice the significant results obtained by the UIET image restoration technique, as well as the significant results obtained using the YOLO-NAS and YOLOX for fish detection in underwater images.

D. Fish Detection Accuracy Assessment: Quantitative Analysis

In this experiment we intend to evaluate the accuracy of the fish detection process, in a quantitative analysis. For this, we compared the proposed approach with the same comparison techniques used in the previous experiment. For the experiments the underwater image dataset was divided into the training set of images with 80% of images, meanwhile the testing set of images is composed by 20% of images.

In Table I we can observe that the proposed approach, combining the UIET underwater image restoration technique and the YOLO-NAS object detection technique, outperform the comparison techniques for fish detection from underwater images. It is important to highlight that the proposed approach, detected fish even when the image was in degraded condition, with better accuracy in challenging image acquisition. From the results we can verify the robustness of the UIET technique, for underwater image enhancement, meanwhile the YOLO-NAS presented effective results for object detection.

TABLE I: Accuracy evaluation of the fish detection approaches in underwater environments.

Accuracy Evaluation for Fish Detection				
	RAW	**IBLA**	**UCDP**	**UIET**
YOLO-NAS	95,15%	90,73%	92,71%	98,04%
YOLOX	94,92%	91,72%	93,69%	97,17%
YOLOv8	94,85%	92,17%	94,60%	96,72%

V. Conclusion and Future Work

This paper proposes a fish detection approach regarding underwater environments, using an image enhancement pipeline to restored the underwater image quality. Finally, from the restored underwater image, the YOLO-NAS is used to detect the presence of different fish, highlighting their positions. In the experiments, state-of-the-art algorithms were used to improve the quality in the detection process.

From the experimental process we verify that the UIET image restoration step significantly improved the image quality, enabling better fish identification. The YOLO-NAS model demonstrated its accuracy in detecting different types of fish, contributing to efficient and reliable fish detection in underwater scenarios. Furthermore, taking the achieved results in account, it was possible validate the proposed methodology, proving its robustness and reliability, using the proposed FUI underwater image dataset.

Future work includes the improvement of underwater image restoration techniques, reducing the image degradation in challenging underwater environments. We intend to create two datasets of underwater images, concerning two challenging rivers, with predominantly dark and brown water colors, presenting high turbidity. Additionally, we intend to expand the investigations regarding another applications related to underwater environments.

Acknowledgments

This work was developed with support from the Motorola, through the IMPACT-Lab R&D project, in the Institute of Computing (ICOMP) of the Federal University of Amazonas (UFAM)

References

[1] Q. Jiang, Y. Zhang, F. Bao, X. Zhao, C. Zhang, and P. Liu, "Two-step domain adaptation for underwater image enhancement," *Pattern Recognition*, vol. 122, p. 108324, 2022.

[2] S. Sun, H. Wang, H. Zhang, M. Li, M. Xiang, C. Luo, and P. Ren, "Underwater image enhancement with reinforcement learning," *IEEE Journal of Oceanic Engineering*, pp. 1–13, 2022.

[3] X. Wang, X. Jiang, Z. Xia, and X. Feng, "Underwater object detection based on enhanced yolo," in *2022 Int. Conf. on Image Processing and Media Computing*, 2022, pp. 17–21.

[4] H. Yang, Y. Shi, and X. Wang, "Detection method of fry feeding status based on yolo lightweight network by shallow underwater images," *Electronics*, vol. 11, p. 3856, 2022.

[5] A. A. Muksit, F. Hasan, M. F. Hasan Bhuiyan Emon, M. R. Haque, A. R. Anwary, and S. Shatabda, "Yolo-fish: A robust fish detection model to detect fish in realistic underwater environment," *Ecological Informatics*, vol. 72, p. 101847, 2022.

[6] A. Kuswantori, T. Suesut, W. Tangsrirat, G. Schleining, and N. Nunak, "Fish detection and classification for automatic sorting system with an optimized yolo algorithm," *Applied Sciences*, vol. 13, no. 6, 2023.

[7] X. Wang, X. Jiang, Z. Xia, and X. Feng, "Underwater object detection based on enhanced yolo," in *2022 Int. Conf. on Image Processing and Media Computing*, 2022, pp. 17–21.

[8] R. Liu, X. Fan, M. Zhu, M. Hou, and Z. Luo, "Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 30, no. 12, pp. 4861–4875, 2020.

[9] L. A. Martinho, F. G. Oliveira, J. M. B. Cavalcanti, and J. L. S. Pio, "Underwater image enhancement based on fusion of intensity transformation techniques," in *Latin American Robotics Symp.*, 2022.

[10] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.

[11] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Recognition*, 2020.

[12] S. Li, B. Pan, Y. Cheng, X. Yan, C. Wang, and C. Yang, "Underwater fish object detection based on attention mechanism improved ghost-yolov5," in *2022 7th Int. Conf. on Intelligent Computing and Signal Processing (ICSP)*, 2022, pp. 599–603.

[13] A. F. Ayob, K. Khairuddin, Y. M. Mustafah, A. R. Salisa, and K. Kadir, "Analysis of pruned neural networks (mobilenetv2-yolo v2) for underwater object detection," in *11th National Technical Seminar on Unmanned System Technology 2019*. Springer Nature Singapore, 2021, pp. 87–98.

[14] W. S. Q. H. Minghua Zhang, Shubo Xu and Q. Wei., ""lightweight underwater object detection based on yolo v4 and multi-scale attentional feature fusion"," in *Remote Sensing*, vol. 13, no. 22, 2021.

[15] J.-H. Park and C. Kang, "A study on enhancement of fish recognition using cumulative mean of yolo network in underwater video images," *Journal of Marine Science and Engineering*, vol. 8, p. 952, 2020.

[16] Z. Wang, H. Liu, G. Zhang, X. Yang, L. Wen, and W. Zhao, "Diseased fish detection in the underwater environment using an improved yolov5 network for intensive aquaculture," *Fishes*, vol. 8, no. 3, 2023.

[17] S. Sun, H. Wang, H. Zhang, M. Li, M. Xiang, C. Luo, and P. Ren, "Underwater image enhancement with reinforcement learning," *IEEE Journal of Oceanic Engineering*, pp. 1–13, 2022.

[18] J.-M. M. A. B. P. Nicolas Limare, Jose-Luis Lisani and C. Sbert, "Simplest color balance," *IPOL Journal*, pp. 297–315, 2011.

[19] Y.-T. Peng and P. Cosman, "Underwater image restoration based on image blurriness and light absorption," *IEEE Transactions on Image Processing*, vol. PP, p. 1, 2017.

[20] F. M. S. B. M. C. P. Drews-Jr, E. Nascimento, "Transmission estimation in underwater single images," *IEEE ICCV - Workshop on Underwater Vision*, p. 825–830, 2013.