ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

Underwater Biological Recognition Method Based on Multi-scale Image Segmentation Technology

Yining Chen[1,*]

[1] School of Information Engineering, Minzu University of China, Beijing 100081, Beijing, China

SUMMARY: *This study investigates the applicability of a deep learning–based multi-scale image segmentation framework for underwater biological recognition in the Chinese context. The research is grounded in controlled experimental evaluations conducted using publicly available benchmark datasets of underwater organisms. The experimental results demonstrate that the proposed multi-scale segmentation approach outperforms conventional segmentation methods by more effectively capturing underwater biological targets in complex and heterogeneous environments. By leveraging multi-scale feature representation, the model exhibits a strong capability to identify organisms of varying shapes and sizes under challenging underwater conditions. Quantitative evaluation indicates that the proposed method achieves an overall F1-score of 90.7%, highlighting its effectiveness in pixel-level segmentation tasks for underwater biological recognition. The analysis further reveals that underwater recognition remains challenging due to substantial scale variability among biological targets, with certain species occupying less than 5% of the image area, while larger organisms such as fish may dominate a significant portion of the visual field. Despite these challenges, robustness assessments confirm that the multi-scale segmentation model maintains stable and high recognition performance across different scale categories. Specifically, recognition accuracy reaches 87.6% for small-scale targets and exceeds 94% for large-scale organisms. Overall, the findings demonstrate that deep learning–based multi-scale image segmentation offers significant potential for accurate identification and monitoring of underwater biological species in diverse and dynamic environments. The proposed approach provides a robust and scalable solution for underwater ecosystem monitoring, particularly in scenarios characterised by complex illumination conditions, background interference, scale diversity, and environmental variability.*

KEYWORDS: *Deep learning; multi-scale image segmentation; underwater biological recognition; underwater image analysis; pixel-level segmentation*

1 Introduction

The marine underwater system is considered an important natural resource, as it accounts not only for the environmental stability of a region but also the socioeconomic development and human well-being, which significantly depend on the marine underwater system [1]. In particular, [2] held that the marine underwater ecosystem accounts for the biodiversity conservation, maintenance of ecological balance, coastal protection and disaster mitigation,

*23302157@muc.edu.cn
https://doi.org/10.65102/is20261230

Published: 10 June, 2026

food security and livelihoods. Given the immense value of the marine underway ecosystem, [3] have pointed out that it is essential to have a solid system that could be used to identify and monitor underwater biological organisms. Such monitoring and assessments are needed not only for accessing the range of benefits associated with marine underwater ecosystems, but also fishery resource management, ecological assessment of a region and the development of an intelligent ocean observation system could not be realised in the absence of a sound, reliable biological information system [4]. A more precise view regarding the marine ecosystem could be gained from the analysis of Figure 1 that exhibits the structural model of marine underwater ecosystem functions:

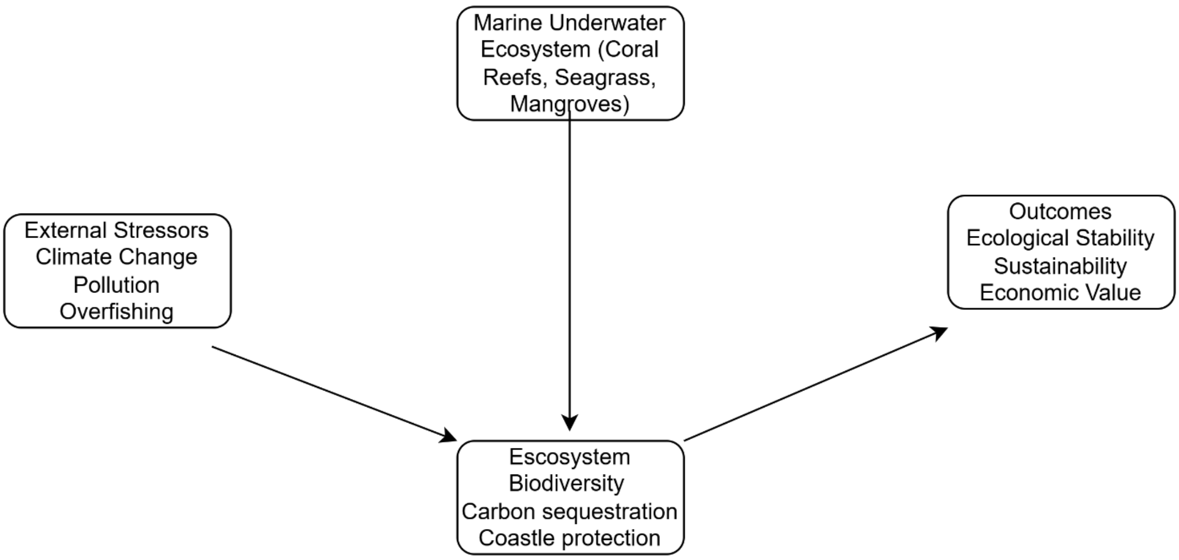

Figure 1: Structural model of Marine Underwater Ecosystem Functions

Although significant challenges in this regard have been witnessed in the past, the recent development in terms of underwater imaging equipment and the development of autonomous platforms [5]. These emerging technologies could be effectively leveraged for collecting and analysing large volumes of underwater visual data in the coastal and offshore regions [6]. However, challenges persist, as a range of technical challenges are witnessed in transforming such underwater visual data into reliable biological information [7]. To tackle such challenges, the underwater biological image recognition technology has been developed, which largely aims to analyse images. The technology thus developed has been gaining precedence largely because, unlike the traditional methods, the image-based recognition technology has been offering significant benefits in terms of scalability, non-invasiveness, and efficiency [8].

However, the images thus generated are characterised by significantly complex environments, including background disorder, lighting, colour distortions, and turbidity. The images are further complicated by the fact that the images are mostly pertaining to underwater organisms that also have significant variation in terms of size, shape, and texture [9]. Usually the organisms found could range from small planktonic species to large fishes and benthic variation. These essential characteristics in turn make the process of image recognition significantly difficult [10]. In particular, the traditional technology used for the sake of underwater biological recognition has been facing core challenges, which include degradation of underwater images, difficulties in accurate foreground extraction, and scale variations [11].

To address such challenges, the deep learning-based image segmentation technology has been developed with the aim of ensuring biological recognition, as the technology could more effectively isolate biological regions from complex backgrounds [12]. There are different deep learning-based segmentation strategies that have been developed these days; however, amongst these, multi-scale image segmentation has witnessed significant potential, as the technology could more effectively address the environmental complexity. In particular, [13] are of the view that multi-scale segmentation can more effectively detect small organisms, while at the same time the technology could target larger underwater species. Considering the significant advantages associated with the deep-learning multi-scale image segmentation technology, this research has been organised with the aim to critically analyse the deep learning–based multi-scale image segmentation method that has been adopted in China for underwater biological recognition. The research will analyse how the technology developed could be adopted to address the scale variability and environmental complexity of underwater biological images.

2 Materials and Methods

2.1 Research Design

The research design that has been adopted in this study that analyses the deep learning–based multi-scale image segmentation method that has been adopted in China for underwater biological recognition is experimental design. Experimental design has been adopted with the aim to ensure more scientific and objective analysis of the collected data and information regarding underwater biological images that are widely used in marine vision and underwater images. For the sake of this research, datasets containing underwater images that were captured in real marine environments, characterised by complex backgrounds, colour attenuation, scale variability and variable illumination. While conducting the experimental design, particular attention has been given to experimental reliability and reproducibility. For this purpose, the collected datasets have been grouped into training, validation and testing subsets. The training subsets comprised 70%, validation 15% and testing 15%.

2.2 Data Sources and Data Preparation

Datasets have been generated from publicly available underwater biological image datasets. The datasets contain multiple categories of underwater organisms that go beyond fishes and contain mollusks, crustaceans, and corals. The dataset collected in the form of training, validation, and testing subsets was provided with pixel-level annotation and bounding-region labels aligned with biological targets. The process thus adopted ensured supervised training of segmentation and recognition model. As the collected images in different datasets were suffering from low contrast, colour distortion, and other distortion, to overcome these peculiarities and ensure feature learning, a standardised pre-processing pipeline has been used in all input images. Some basic pre-processing has been carried out, which helped in the standardisation of data. In this regard, first colour correction has been conducted, as a channel-wise normalisation process has been conducted. The process thus adopted helped in significantly reducing the wavelength-dependent light attenuation. All the images have been resized, and uniform resolution has been adopted for fixing the pixel value range. Additionally, pre-processing also included noise suppression, as smoothing filters were adopted to decrease background interference. Furthermore, the pre-processing also includes contrast enhancement, which has been conducted with the aim of enhancing the visibility of biological contours of images.

2.3 Data Analysis

For conducting data analysis, there are different processes adopted in this study. This includes multi-scale feature extraction that has been adopted with the aim to extract biological features of spatial resolutions; this study adopted multi-scale feature extraction. For this sake, convolutional layers having different receptive fields have been adopted that helped in extracting texture features and contextual information. In this regard, computation efficiency has been evaluated through computation of inference time and frame processing rate that helped in analysing the feasibility of real-world underwater monitoring applications. Furthermore, the study used multi-feature maps extracted by the encoder for which the skip connection and feature concentration process have been adopted in the study. Through the adoption of a fusion strategy, the spatial resolution has been preserved, while at the same time, the process helped in integrating semantic information across scales. For the sake of underwater recognising biological recognition, the segmentation strategy adopted in this study is based on segmented regions. For analysing the performance of segmentation, this study adopted intersection over union, the dice coefficient and pixel accuracy. The values for this purpose – precision, recall, F1 score and overall accuracy – have been computed for each biological category. The intersection over union has been computed using the following formula:

$$\mathrm{IoU} = \frac{|S_p \cap S_g|}{|S_p \cup S_g|}$$

In the above formula, IoU denotes intersection over union, Sp denotes the predicted segmentation region, and Sg denotes the ground-truth segmentation region. On the other hand, the Dice coefficient has been computed through the F1 score for segmentation, for which the following formula has been adopted in this study:

$$\mathrm{Dice} = \frac{2|S_p \cap S_g|}{|S_p| + |S_g|}$$

In the above formula, Sp denotes the predicted segmentation region, while Sg denotes the ground-truth segmentation region. While the pixel accuracy has been established through the following formula of pixel accuracy:

$$\mathrm{PA} = \frac{TP + TN}{TP + TN + FP + FN}$$

In the above formula, TP denotes true positive, TN denotes true negative, FP denotes false positive, and FN denotes false negative.

3 Results

3.1 Segmentation Performance

The segmentation performance results, using publicly available underwater biological image datasets has been summarised in the following Table 1:

Table 1: Segmentation Performance

Segment	IoU	Dice Coefficient	Pixel Accuracy
Fish	86.7%	92.9%	94.3%
Crustaceans	83.8%	91.1%	92.8%
Mollusks	82.5%	90.4%	92.2%
Corals	84.1%	91.3%	93.1%
Average	84.3%	91.4%	93.1%

The above Table 1 summarised the multi-scale segmentation results for different categories of species. From the analysis of the results obtained, it is pertinent to note that the multi-scale segmentation approach has been more consistent for all the biological categories. In terms of IoU, the fish category demonstrates the highest value. Furthermore, the analysis of other categories, including the Dice coefficient, also revealed that robust region-level segmentation demonstrates that noisy environments have been significantly reduced through the process of the multi-scale segmentation method. Additionally, the analysis of the entire table could also reveal that there is a small level of variance in different categories of species analysed and summarised in the above Table 1. This in turn demonstrates strong generalisation capabilities, validating the argument that multi-scale feature fusion significantly enhances small-object segmentation stability.

3.2 Recognition Accuracy

The recognition accuracy on the basis of segmented output has been computed and summarised in the following Table 2:

Table 2: Image Recognition Accuracy

Segment	IoU	Dice Coefficient	Pixel Accuracy
Fish	93.4%	92.1%	92.7%
Crustaceans	90.7%	89.9%	90.3%
Mollusks	89.2%	88.5%	88.8%
Corals	91.6%	90.8%	91.2%
Average	91.2%	90.3%	90.7%

From the analysis of the above Table 2, it is pertinent to note that segmentation-driven recognition pipelines realised more than 90% accuracy, which in turn signifies the effectiveness of multi-scale segmentation. The greater accuracy in turn denotes a significant reduction of background interference and better discriminative feature learning resulting from the use of segmentation-based underwater imaging. More comprehensive analysis of the summarised results has been exhibited in the following Figure 2, which exhibits recognition performance metrics, including precision, recall, and F1-score for the four core categories of organisms analysed in this study.

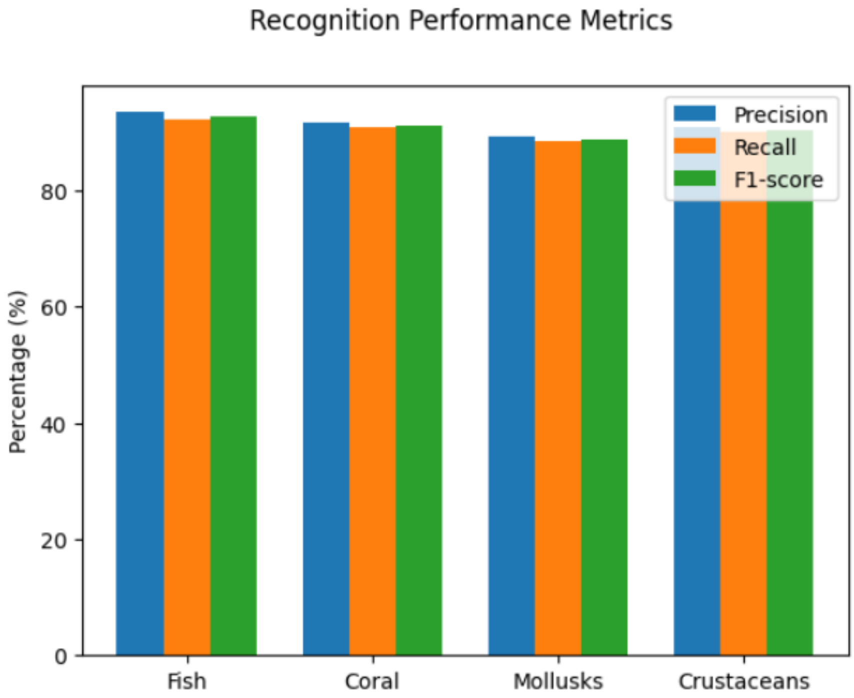

Figure 2: Recognition Performance Metrics

From the analysis of the above Figure 1, one could observe that recognition of underwater images is more consistent across different biological types. However, marginal reduction has been witnessed for small and visually ambiguous organisms. The analysis thus confirmed that through the process of multi-scale segmentation, the risks of scale-related recognition could be reduced.

3.3 Robustness of Scale Variation

The recognition accuracy of different target scales has been computed and summarised in the following Table 3:

Table 3: Robustness of Scale Variation

Target Scale	Accuracy
Small (< 5% of image area)	87.6%
Medium (5 to 20% of image area)	91.8%
Large (> 20% of the image area)	94.2%

The above Table 3 summarised recognition accuracy under different target scales. From the analysis of Table 3, it is pertinent to note that the multi-scale segmentation process for the underwater biological recognition significantly enhances small-object recognition, as 87.5% of different underwater species and organisms that could have been otherwise skipped have been recognised in the current experimental study. On the other hand, 91.8% of the medium objects and 94.2% of large underwater species and organisms could be recognised through multi-scale segmentation. From the analysis of these figures, it is thus pertinent to note that the accuracy and opportunities for recognition of biological objects underwater significantly increase when the object size increases.

3.4 Computation performance

Model efficiency has been evaluated in the following Table 4, which presents the summary of the model efficiency evaluation.

Table 4: Computation performance

Metrics	Value
Average inference time	48 ms/ image
Parameters	21.6 million
FPS (GPU)	20.8

From the analysis of the above Table 4, it is pertinent to note that the inference time is < 50 ms, which supports the system, which in turn makes it a more viable imaging system for real-time application. Furthermore, the model size is acceptable for embedded marine systems. Considering all three parameters, including inference time, parameters, and FPS value, the multi-scale segmentation realised a favourable trade-off between accuracy of the biological images and computational cost. The greater efficiency highlighted in the above Table 4, in turn, demonstrates that the multi-scale segmentation is more suitable for near real-time underwater monitoring applications. In particular, the results demonstrate greater fitness of underwater monitoring that could be more suitable for autonomous underwater vehicles and smart ocean management systems.

4 Discussions

4.1 Effectiveness of Multi-Scale Image Segmentation

From the experimental study conducted, it is very clear that the deep-learning-based multi-scale image segmentation process is considered more valid and effective for capturing underwater biological recognition. The results analysed and summarised in Table 1 demonstrate that strong spatial overlap between predicted biological regions and ground-truth annotation has been noted while adopting the deep learning-based multi-scale image segmentation for underwater biological recognition. In particular, this study confirms that the multi-scale segmentation model could more effectively capture the biological targets in diverse and complex environments, having the opportunity to more effectively identify diverse shapes and sizes in complex underwater environments. The greater efficiency of the deep learning-based multi-scale image segmentation for underwater biological recognition could be attributed to fine-grained boundary information and the contextual cues that are leveraged for identification of underwater biological organisms. The findings of the current study are in line with the findings of [14] and [15], who have also found that the deep learning-based multi-scale image segmentation for underwater biological recognition is a more effective and efficient model, particularly suitable for underwater noisy background areas, as the model enjoys higher recognition performance that leads to better classification of biological species.

4.2 Impact of Multi-Scale Segmentation on Recognition Accuracy

From the analysis of the overall F1 score, it is pertinent to note that this study showed that the deep learning-based multi-scale image segmentation for underwater biological recognition realised an overall F1 score of 90.7%. In particular, for the fish and coral categories, this score has been recorded greater than 91%. These figures in turn demonstrate that the deep learning-based multi-scale image segmentation for underwater biological recognition could be adopted for more effective pixel-level segmentation. In particular, the technology could play a pivotal role in enhancing downstream recognition of tasks. It has been noted that the multi-scale segmentation process could more effectively isolate biological targets from complex underwater backgrounds. The technology thus could work in complex environments, where other technologies failed, as multi-scale segmentation decreases feature ambiguity and improves inter-class discriminations. From the analysis of the results, it is pertinent to note that the effectiveness of multi-scale segmentation significantly improves in complex morphological characteristics. This is evident from the analysis of the accurate recognition witness in the case of molluscs and crustaceans. As these organisms have visual differences and irregular boundaries, the multi-scale segmentation technology could more effectively preserve the structural details, irrespective of different resolutions. The findings of this study confirmed the findings of [16], who have also found that the deep learning-based multi-scale image segmentation for underwater biological recognition could be adopted for more effective pixel-level segmentation.

4.3 Robustness to Scale Variation

In the underwater organisms, challenges are witnessed because of the significant differences in the target scale, as some of the species occupy less than 5% of the image, as compared to large fishes that could account for a significant portion of the total image. As per the robustness analysis conducted and exhibited in Table 3, it is pertinent to note that the deep learning-based multi-scale image segmentation for underwater biological recognition maintains significantly stable recognition accuracy across different types of scale categories. The accuracy recorded for the small targets in this experimental study has been 87.6%, while for the large organism the robustness accuracy has been recorded at more than 94%. From the analysis offered in this regard, the deep learning-based multi-scale image segmentation for underwater biological recognition is a more effective image monitoring in scale-sensitive environments. The multi-scale model could more effectively detect and recognise small biological targets, which are mostly ignored in the case of the single-scale method of deep learning-based image recognition. The significantly higher accuracy of the deep learning-based multi-scale image segmentation for underwater biological recognition recorded in the current study reiterates the findings of different researchers who have also presented almost similar contentions. For example, [17] and [18] have also found that the deep learning-based multi-scale image segmentation for underwater biological recognition enjoys greater robustness and accuracy.

4.4 Implications in Terms of Marine Monitoring

From the analysis of the key results elaborated in this study, the deep learning-based multi-scale image segmentation for underwater biological recognition offered significant opportunities for identifying and monitoring underwater species and organisms in diverse underwater environments. In particular, the emerging technology could play a more effective role in the monitoring of underwater ecosystems that are more challenging due to different illumination, biological scales, background complexities, and environmental variations. The deep learning-based technology emerging in the form of multi-scale segmentation could be particularly important for monitoring of biodiversity, intelligence-based fisheries development and management, and long-term ecological management. In particular, the computational efficiency calculated for the deep learning-based multi-scale image segmentation for underwater biological recognition indicates its greater usefulness for real-time implementation in GPU-enabled underwater platforms. If the technology is effectively used, it could play a pivotal role in future marine resource management and marine conservation.

5 Conclusion

This study examined how a deep learning–based multi-scale image segmentation method could be adopted in China for underwater biological recognition. The findings of the study are based on an experimental study for which publicly available datasets regarding underwater organisms have been adopted in the study. From the experimental study conducted, it is very clear that the deep-learning-based multi-scale image segmentation process is considered more valid and effective for capturing underwater biological recognition. This study confirms that the multi-scale segmentation model could more effectively capture the biological targets in diverse and complex environments, having the opportunity to more effectively identify diverse shapes and sizes in complex underwater environments. From the analysis of the overall F1 score, it is pertinent to note that this study showed that the deep learning-based multi-scale image segmentation for underwater biological recognition realised an overall F1 score of 90.7%. The deep learning-based multi-scale image segmentation for underwater biological recognition could be adopted for more effective pixel-level segmentation. Additionally, the study found that in the underwater organisms, challenges are witnessed because of the significant differences in the target scale, as some of the species occupy less than 5% of the image area, as compared to large fishes that could account for a significant portion of the total image. From the robustness analysis, it is pertinent to note that the deep learning-based multi-scale image segmentation for underwater biological recognition maintains significantly stable recognition accuracy across different types of scale categories. The accuracy recorded for the small targets in this experimental study has been 87.6%, while for the large organism the robustness accuracy has been recorded at more than 94%. The study thus shows that the multi-scale image segmentation for underwater biological recognition offered significant opportunities for identifying and monitoring underwater species and organisms in diverse underwater environments. In particular, the emerging technology could play a more effective role in the monitoring of underwater ecosystems that are more challenging due to different illumination, biological scales, background complexities, and environmental variations.

About the Author

Yining Chen was born in Jinzhou, Liaoning, China, in 2001. He obtained a bachelor's degree from Qingdao University of Science and Technology in China. He currently studying at School of Information Engineering, Minzu University of China. His main research direction is Computer vision.

References

[1] C. Wang, W. Duan, C. Luan, J. Liang, L. Shen and H. Li, "USNet: Underwater image superpixel segmentation via multi-scale water-net," *Frontiers in Marine Science, 1411717,* 11, 2024.

[2] X. Cheu, H. Chen, D. Zhou, J. Tian, Q. Gao, J. Zhou and S. Liu, "A multimodal optical dataset for underwater detection, segmentation, enhancement, and reconstruction," *Scientific Data, 1554,* 12, 2024.

[3] X. Wei, M. Zhang, W. Song and H. Mei, "A systematic review of deep learning-based underwater object detection," *Neurocomputing,* 527, pp. 204-232, 2023.

[4] W. Ding, L. Ma, R. Liu and X. Fan, "Semantic-aware texture-structure feature collaboration for underwater image enhancement. 2022 International Conference on Robotics and Automation (ICRA)," 2022.

[5] S. Padmapriya, A. Umamgeswari, S. Deepa and B. J. Fritha, "Underwater image denoising and detection using deep learning," *Journal of Intelligent & Fuzzy Systems,* 45(4), pp. 7129-7144, 2023.

[6] Y. Wei, Y. Wang, B. Zhu, B. Lin, D. Wu, X. Xue and R. Wang, "Underwater detection: A brief survey and a new multitask dataset," *International Journal of Network Dynamics and Intelligence, 100025,* 3(4), 2024.

[7] F. Zhang, J. Hu and Y. Sun, "Underwater fish image recognition based on knowledge graphs and semi-supervised learning feature enhancement," *Scientific Reports, 29396,* 2025.

[8] C. Xiao, Z. Zhou and Y. Hu, "A lightweight semantic segmentation model for underwater images based on DeepLabv3+," *Journal of Imaging, 162,* 11(5), 2025.

[9] M. Ding, G. Li, Y. Hu, H. Liu, Q. Hu and X. Huang, "Semi-Supervised underwater image enhancement method using multimodal features and dynamic quality repository," *Journal of Marine Science and Engineering, 1195,* 13(6), 2025.

[10] M. Zhou, L. Cai, J. Jia and Y. Gao, "Multi-scale aware turbulence network for underwater object recognition," *Frontiers in Marine Science, 1301072,* 11, 2024.

[11] X. Zuo, J. Jiang, J. Shen and W. Yang, "Improving underwater semantic segmentation with underwater image quality attention and multi-scale aggregation attention," *Computer Vision and Pattern Recognition*, 12(2), 2025.

[12] M. Contini, V. Illien, B. Barde, S. Poulain, S. Bernard, A. Joly and S. Bonhommeau, "From underwater to aerial: a novel multi-scale knowledge distillation approach for coral reef monitoring," *Computer Vision and Pattern Recognition, 2502,* 2025.

[13] H. Ge, P. Sun and Y. Lu, "A new dataset, model, and benchmark for lightweight and real-time underwater object detection (EFCWM-Mamba-YOLO)," *Neurocomputing, 130891,* 651, 2025.

[14] B. Chen, W. Zhao, Q. Zhang, M. Li, M. Qi and Y. Tang, "Semantic segmentation of underwater images based on improved SegFormer," *Frontiers in Marine Science,* 13(2), 2025.

[15] H. Zhang, C. Dai, C. Chen, Z. Zhao and M. Lin, "Underwater fish species recognition based on feature fusion visual attention graph convolutional networks," *Intelligent Marine Technology and Systems*, 31(4), 2025.

[16] B. Alawode, Y. I. Ganapathi, S. Javed, N. Werghi, M. Bennmound and A. Mahmood, "AquaticCLIP: A Vision-Language Foundation Model for Underwater Scene Analysis," *Computer Vision and Pattern Recognition, 1785,* 2025.

[17] L. S. Saoud, L. Seneviratne and I. Hussain, "MARS: Multi-Scale Adaptive Robotics Vision for Underwater Object Detection and Domain Generalization," *Robotics, 2312,* 2023.

[18] S. Lian, Z. Zhang, H. Li, W. Li, L. Tianrao, S. Kwong and R. Cong, "Diving into Underwater: Segment Anything Model Guided Underwater Salient Instance Segmentation and A Large-scale Dataset," *Computer Vision and Pattern Recognition, 2406,* 2024.