

Sea Turtle Recognition with Multiple Data Augmentation Methods Suitable for Marine Scenarios [†]

Yi-Chieh Hung ^{1,2}, Jhih-Ya Chan ^{1,2}, Wei-Cheng Lien ¹ , Yan-Tsung Peng ^{2,*}  and Li-Shu Chen ^{3,*}

¹ Hyson Technology Inc., New Taipei City 23553, Taiwan; redyccccc@gmail.com (Y.-C.H.); vivianvivin.vivianvivin55@gmail.com (J.-Y.C.); lien@hysontech.tw (W.-C.L.)

² Department of Computer Science, National Chengchi University, Taipei City 116011, Taiwan

³ National Museum of Marine Science and Technology, Keelung City 202010, Taiwan

* Correspondence: ytpeng@cs.nccu.edu.tw (Y.-T.P.); lschen@mail.nmmst.gov.tw (L.-S.C.)

[†] Presented at 2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBDE), Penang, Malaysia, 27–29 June 2025.

Abstract

The sea turtle is an indicator organism used in marine conservation to identify the health status of ecosystems in various marine regions. In the past, researchers had to review an 8 h underwater video every day to monitor and count sea turtle appearances. However, since sea turtles often appear for only short periods, traditional approaches of manual searching and counting require significant labor and time to ensure accurate periods of their appearance. To address this issue, we adopted the You Only Look Once (YOLO) model for object detection, utilizing real underwater videos captured from three different areas in the Taiwan Keelung City Chaojing Bay Aquatic Plants and Animals Conservation Area for training and testing. To overcome limitations, such as underwater blur, sediment interference, obstructions from other fish, and distant targets that are challenging to identify, we applied data augmentation techniques, including scaling, rotation, and depth blur, with labeled data of different fish species to improve generalization capability. The experimental results of this study showed that this method achieves a 99.4% accuracy in sea turtle detection. After 60 days of deployment across the three areas, the model reduced search time by over 99%, significantly improving efficiency and reducing workload.

Keywords: marine conservation; AI; sea turtle recognition; data augmentation

1. Introduction

Sea turtles are an essential indicator of environmental stability and biodiversity [1]. Unfortunately, many species of sea turtles are facing significant threats due to habitat degradation, pollution, climate change, and illegal poaching. Effective conservation efforts rely on accurate monitoring of sea turtle populations, migration patterns, and health status. However, traditional monitoring methods, such as manual observation and tagging, are labor-intensive, time-consuming, and can potentially disturb the animals being studied.

With the advancement of artificial intelligence technology, deep learning-based image recognition models [2] offer innovative solutions for turtle conservation. This article presents a turtle recognition model built on the You Only Look Once (YOLO) architecture [3,4] and explores various techniques to further improve the accuracy and efficiency of turtle monitoring. As one of the latest real-time object detection algorithms, YOLO delivers high computational efficiency and excellent recognition accuracy.



Academic Editors: Teen-Hang Meen,
Wei Chien and Cheng-Fu Yang

Published: 9 March 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

To improve the performance of detection models, data augmentation is widely used to address the issue of limited training data. By artificially increasing the diversity of datasets, data augmentation can improve generalization and reduce the risk of model overfitting. Common techniques include basic transformations such as flipping, rotation, cropping, and scaling, which generate variations in the training images while maintaining their original semantics [5]. More advanced methods, such as color space transformations and adversarial training, introduce novel ways to alter color properties or create adversarial examples to challenge model robustness [5]. The timing of augmentation, such as whether it is applied before or during training, also significantly improves classification outcomes, especially with techniques like image warping [6]. Furthermore, combining traditional augmentation methods with advanced techniques like random erasing or mix-up has proven to improve the accuracy and robustness of convolutional neural networks in large-scale image recognition tasks [7]. In Ref. [8], image restoration is widely used as data augmentation to enhance the accuracy of underwater object detection. Nevertheless, underwater environments pose unique challenges, such as blur or fog caused by sediments, varying lighting conditions, and target objects being too distant or too small, making achieving high accuracy in underwater detection tasks particularly challenging.

Collaborating with the National Museum of Marine Science and Technology, Taiwan, we obtained a dataset of sea turtle images captured in real-world environments. Marine research experts meticulously labeled the dataset to ensure accuracy. We utilized the YOLO model for turtle detection, achieving high performance, including 97% accuracy, 98% precision, and recall, as well as mean average precision at an intersection over union higher than 0.5 ($\text{mAP}@0.5 = 0.988$ and $\text{mAP}@0.5:0.95 = 0.753$). However, practical deployment in real underwater scenes reveals environmental challenges, such as low lighting conditions, blurring, and occlusion, causing some turtle clips to be incorrectly detected. In addition, according to statistics from marine conservation researchers, sea turtles appear infrequently and for only short moments, sometimes merely a few seconds within an entire month of footage. Therefore, the marine conservation team hopes to enhance detection efficiency to ensure that no time segment with sea turtle presence is missed during the identification process.

To address issues posed by underwater environments, such as blur or fog caused by sediments and lighting variations and the small target objects, we adopted four different methods: injecting noise using various blur functions, incorporating additional marker datasets of other fish, and augmenting the data by rotating and resizing objects. We also conducted a series of experiments to evaluate the impact of combining these methods on the model's performance. Experimental results show that even when the model performance is already relatively good, our method can still significantly improve the accuracy by 2.4%, recall by 1.17%, precision by 2%, $\text{mAP}@0.5$ by 0.006, and $\text{mAP}@0.5:0.95$ by 0.083.

In applications, the method adopted in this study enabled the detection of untrained 60-day data from three different sea areas and successfully found five time periods when sea turtles appeared, exactly the same as those discovered by marine researchers through manual search. Previously, researchers had to spend significant amounts of time each day manually scanning footage to identify turtle appearances. The AI-driven approach enables them to focus on the time periods detected by the model, reducing manual search time by over 99% and significantly improving observation efficiency.

2. Proposed Methods

Considering the need for the marine conservation team to expand to edge computing in the future, we chose YOLOv8 [9] as the baseline model for the sea turtle recognition task. Previous work [9,10] has pointed out that YOLOv8 can significantly improve accuracy

with fewer computing resources, making it highly suitable for deployment on edge devices or in environments with limited computational capacity. To further enhance the model's performance, we selected and adopted four performance improvement methods through observation to overcome the problem of difficult detection of objects in complex underwater environments, as shown in Figure 1. In the following sections, we will explain each method in detail.

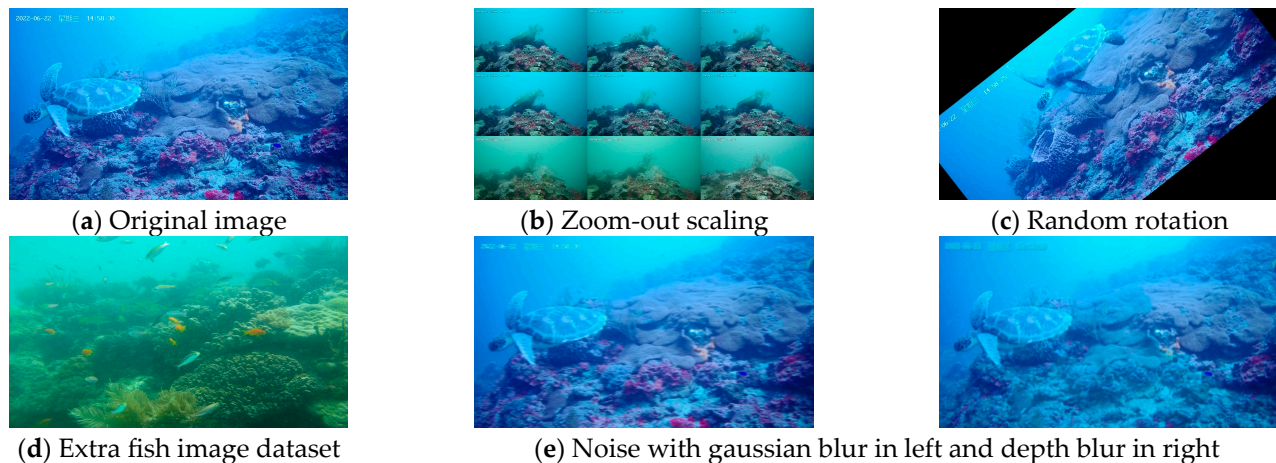


Figure 1. Comparisons of the original image with and without different performance improvement methods as data augmentation.

2.1. Zoom-Out Scaling Augmentation

Underwater, sea turtles often appear as small objects in the beginning frames, getting larger while coming towards the camera. To enable the model to learn the features of small-scale objects, we applied a zoom-out augmentation [11] by reducing the image size by a factor of 9. Figure 1b shows the results after using zoom-out scaling augmentation. This allows the model to increase its chances of detecting turtles even when they occupy a small portion of the frame.

2.2. Random Rotation Augmentation

From the video dataset, we observed that most sea turtles often stay in a fixed position for a period of time. This led the model to unintentionally associate specific locations with the presence of sea turtles rather than learning their intrinsic features. Such a bias reduces the model's generalization capability. To address this issue, we applied random rotation as a data augmentation technique, where images were rotated by random angles between -45° and 45° [11] as shown in Figure 1c. Adding orientation variations, this augmentation forces the model to rely on the true features of sea turtles, such as their shape, texture, and patterns, instead of memorizing relative spatial positions. This approach improves the model's robustness and ensures better generalization to new, unseen images.

2.3. Extra Image Database of Other Fish Species

To improve the model's ability to distinguish between positive and negative samples, we included another dataset of negative samples [12], a public underwater fish dataset [13], including fish images without sea turtles, as shown in Figure 1d. This enables the model to correctly distinguish between images with and without turtles while accounting for false detections caused by occlusions by other fish, thereby reducing false positive detections and improving its overall reliability.

2.4. Noise Injection

Weather and ocean currents, such as wind, waves, typhoons, and northeast monsoons, often affect underwater environments. These factors cause sediments to float in the water, resulting in blurred and visually degraded underwater images that are difficult to identify. To simulate the blurry conditions, we introduce noise into the original images, intentionally adding blur. We compared the effectiveness of two noise injection methods: (1) Gaussian Blurring and (2) Depth-Aware Blurring. We evaluated both methods to maximize detection performance under various blurry conditions, as described below.

2.4.1. Noise Injection with Gaussian Blurring

First, we introduced the common noise injection by adding Gaussian noise and Gaussian blurring to the images [5]. This augmentation technique helps the model learn sea turtle features in noisy and blurred underwater environments, improving its robustness in detecting turtles under varying visibility conditions, as shown on the left side of Figure 1e.

2.4.2. Noise Injection with Depth-Aware Blurring

Traditional Gaussian blur employs uniform blurring across the entire image, disregarding differences in depth, as distant objects appear significantly blurrier than nearby ones [14]. To simulate real-world underwater conditions more accurately, we employed a multi-scale Gaussian filtering approach [14,15], which utilizes Gaussian filters with varying kernel sizes to produce blurred images. The level of blurriness at each pixel was estimated by calculating the absolute difference between the original image and its Gaussian-blurred counterparts as follows.

$$BlurDiff(x, y) = \frac{1}{n} \sum_{i=1}^n |I(x, y) - G_{\sigma_i}(x, y)|, \quad (1)$$

Here, $G_{\sigma_i}(x, y)$ represents the Gaussian-filtered image with a standard deviation σ_i . The blurriness from all scales was averaged to obtain a blurriness map, approximating the image's depth variation. From this estimated map of blur, we applied adaptive Gaussian blurring to different image regions, so that distant objects are more blurred and closer objects retain more details, creating a more natural depth-based blur effect, as shown on the right side of Figure 1e.

3. Results

3.1. Dataset and Environment Setup

The dataset used in this study was provided by the National Museum of Marine Science and Technology, and marine conservation experts labeled images of sea turtles and various fish species. This dataset comprises images captured by underwater cameras on various dates and at different times throughout 2022. These images were collected from three distinct underwater monitoring areas in the Taiwan Keelung City Wanghaixiang Chaojing Bay Aquatic Plants and Animals Conservation Area, including the Gorgonian, sponge, and reef areas. The dataset comprises 3627 images captured by underwater camera No. 1 in the Gorgonian area, 95 by underwater camera No. 2 in the sponge area, and 34 by underwater camera No. 3 in the coral reef area. The dataset comprises a total of 3756 images, as summarized in Table 1. For the experiments, we used 70% of the labeled data for model training and the remaining 30% for testing. The experiments were conducted on a PC-based server equipped with an NVIDIA GeForce RTX 4090 graphics processing unit (GPU) (NVIDIA Corporation, Santa Clara, CA, USA) and ASPEED Graphics Family (WDDM) (ASPEED Technology Inc., Hsinchu, Taiwan) integrated graphics. GPU-accelerated computations were done using Compute Unified Device Architecture

version 12.4. The YOLOv8 model was trained using the following hyperparameters: 300 epochs, a batch size of 16, and an input image size of 640×640 . In addition, early stopping was used in the training process with a patience parameter set to 80, allowing the model to terminate if performance did not improve after 80 epochs.

Table 1. Labeled Image Dataset for Training and Testing.

Camera Index: Sea Area	Labeled Image Data		
	All	Training (70% Data)	Testing (30% Data)
Camera 1: Gorgonian Area	3627	2538	1089
Camera 2: Sponge Area	95	64	31
Camera 3: Coral Reef Area	34	22	12
Total	3756	2624	1132

3.2. Evaluation Metrics

To evaluate the performance of our method, we computed key metrics, including true positives (*TP*), false positives (*FP*), true negatives (*TN*), and false negatives (*FN*). These metrics were then used to calculate Accuracy, Precision, and Recall, which serve as indicators of performance [16]. Their calculations were conducted using the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (5)$$

Accuracy was used to measure the proportion of correct predictions among all predictions, giving the model's overall correctness. Precision indicates how reliable the positive predictions are, and recall measures the model's ability to correctly identify all actual positive cases, providing insight into its sensitivity. In addition, two metrics, mAP@0.5 and mAP@0.5:0.95, are employed. mAP@0.5 measures the average precision of the model across all classes using an IoU threshold of 0.50. This means that a predicted bounding box is considered a correct detection if it overlaps with the ground truth by at least 50%. In contrast, mAP@0.5:0.95 offers a more rigorous evaluation of the model's detection performance under varying degrees of overlap tolerance, ensuring that the evaluation reflects both loose and stricter detection criteria.

3.3. Results and Comparisons

Table 2 shows the experimental results and comparisons using various enhanced techniques. We trained YOLOv8 without applying any data augmentation techniques and evaluated its performance as a baseline. The Exp-Index 1 results indicate that the baseline model has achieved relatively high accuracy, precision, and recall without data augmentation. This is attributed to the fact that this task is relatively simple. Specifically, as target objects, sea turtles have unique visual features that make them easier to identify in underwater images. Due to the unique behavior of turtles, they tend to stay within certain areas, resulting in a relatively concentrated distribution of the objects in the training set.

These factors lead the YOLO model to a high detection performance even during the initial training process.

Table 2. Results and comparisons using different enhanced techniques for data augmentation.

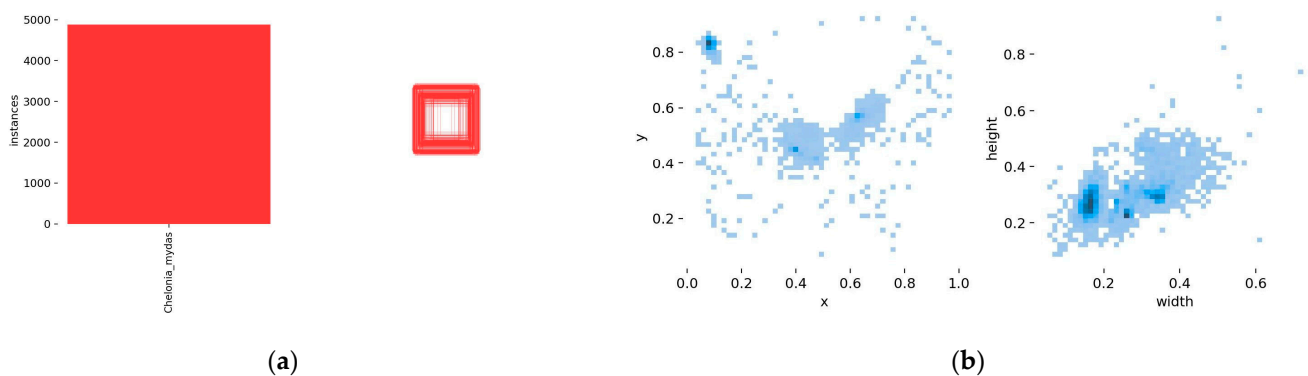
Exp-Index	Baseline Model (YOLOv8)	Applied Augmentation Method					Evaluation Metrics				
		Noise Injection		Extra Fish Dataset	Zoom Out	Rotate	Accuracy (%)	Precision (%)	Recall (%)	mAP @0.5	mAP @0.5:0.95
		Gaussian Blur	Depth Blur								
1	O						97%	98%	98%	0.988	0.753
2	O	O					97%	98.5% (+0.5%)	97.06% (−0.94%)	0.994 (+0.006)	0.826 (+0.073)
3	O		O				98.66% (+1.66%)	99.50% (+1.5%)	99% (+1%)	0.995 (+0.007)	0.837 (+0.084)
4	O		O	O	O	O	99.08% (+2.08%)	99.71% (+1.71%)	99.14% (+1.14%)	0.994 (+0.006)	0.791 (+0.038)
5	O		O		O	O	99.02% (+2.02%)	99.72% (+1.72%)	99.25% (+1.25%)	0.994 (+0.006)	0.788 (+0.035)
6	O		O	O		O	99.36% (+2.36%)	100% (+2%)	99.15% (+1.15%)	0.994 (+0.006)	0.834 (+0.081)
7	O		O	O			99.4% (+2.4%)	100% (+2%)	99.17% (+1.17%)	0.994 (+0.006)	0.836 (+0.083)
8	O			O			98.75% (+1.75%)	99.41% (+1.41%)	98.55% (+0.55%)	0.995 (+0.007)	0.785 (+0.032)
9	O	O		O	O	O	99% (+2%)	99.9% (+1.9%)	98.85% (+0.85%)	0.994 (+0.006)	0.798 (+0.045)
10	O	O			O	O	98.03% (+1.03%)	99.05% (+1.05%)	98.78% (+0.78%)	0.995 (+0.007)	0.797 (+0.044)
11	O	O		O		O	99.25% (+2.25%)	99.21% (+1.21%)	99% (+1%)	0.994 (+0.006)	0.84 (+0.087)
12	O	O		O			99.25% (+2.25%)	99.85% (+1.85%)	99.14% (+1.14%)	0.994 (+0.006)	0.832 (+0.079)

Despite the baseline model's strong performance, statistics from marine conservation researchers have highlighted that the periods when sea turtles appear are relatively rare, as shown in Table 3. Therefore, the marine conservation research team aims to improve its performance further to ensure that all periods during which sea turtles appear are identified.

To improve the model performance, we conducted a comprehensive analysis of the dataset characteristics and inspected the dataset distribution, as shown in Figure 2. Figure 2a illustrates the instance count and bounding box size, while Figure 2b shows the distribution of the position (x,y) and size (height, width) of the detected objects, i.e., sea turtles. The analysis reveals that most objects are relatively small, concentrated in a width range of 0 to 0.4 and a height range of 0.2 to 0.4. This distribution contributes to the lower mAP@0.5:0.95 metric observed in the baseline model results, as shown in Exp-Index 1 of Table 2. In addition, we analyzed and checked the images that the baseline model fails to recognize and identified four challenges: (1) blur or fogging caused by underwater sediments, (2) obstruction caused by other fish, (3) lighting variations that affect visibility, or (4) objects too distant or too small. Therefore, addressing these challenges is essential to further improve turtle recognition performance, especially for objects with small bounding boxes or blurred appearances.

Table 3. The results of Real Applications to automatically identify the sea turtle appearance segments in different sea areas.

Camera in Area	Recorded Days (8 h per day)	Total Video Time (1000 s)	Sea-Turtle Video Segment (DD-MM-YYYY, Time Period)	Sea Turtle Detection Time in Each Video Segment	
				Frame	Second (30 Frames per s)
Camera 1: Gorgonian Area	60	1728	11 August 2022 17:30–18:00	381	12.72
Camera 2: Sponge Area	60	1728	8 August 2022 16:30–17:00	418	13.72
Camera 3: Coral Reef Area	60	1728	2 August 2022 18:00–18:30	223	7.51
			12 August 2022 06:30–07:00	113	3.77
			14 August 2022 12:30–13:00	42	1.4
Total	180	5184		1177	39.12 (Save > 99% Search Time)

**Figure 2.** Dataset characteristic analysis and statistical distribution: (a) instance count and bounding box size, where the red bar represents the total count of *Chelonia mydas* instances; (b) distribution of object positions (x, y) and object sizes ($width, height$), where darker blue regions indicate a higher density of data points.

Hence, we introduced four data augmentation methods: noise injection with blur functions, incorporation of an extra labeled dataset of fish species data, image rotation, and zoom out to reduce object size. Table 2 compares different combinations of these methods, with each row representing an independent experiment marked by Exp-Index. O indicates using a specific data augmentation method in this experiment. For example, Exp-Index 2 applies the Gaussian Blur function to the baseline. The results of each set of experiments in the five-evaluation metrics are listed in the evaluation metrics columns. The numbers in parentheses represent the improvement compared to Exp-Index 1. The comparisons between different experimental combinations were observed as follows.

- Exp-Index 1: Baseline experiment using YOLOv8 for comparison with enhanced techniques.
- Exp-Index 2 to 3: These experiments compare two blurring methods: Gaussian Blur and Depth-Aware Blur (Depth Blur) to address blur or fog caused by underwater sediments and improve accuracy. The results show that the Gaussian blur fails to improve the accuracy and leads to lower recall, making it ineffective for turtle detection, as shown in Exp-Index 2. In contrast, Depth-Aware Blur improves all five metrics, increasing accuracy by 1.66%, as shown in Exp-Index 3. This improvement is attributed

to Depth Blur's effectiveness in retaining more details for closer objects and creating a more natural blur effect.

- Exp-Index 4: Building on Exp-Index 3, this experiment combines Depth-Aware Blur with three additional augmentation methods: Extra Fish Dataset, Zoom Out, and Rotate. The results show improvements in accuracy, recall, precision, and other key metrics compared to Exp-Index 3.
- Exp-Index 5: Removing the Extra Fish Dataset from Exp-Index 4 results in lower accuracy. It indicates that introducing non-turtle labeled data helps balance the dataset and enhances the model's ability to distinguish between turtles and non-turtles.
- Exp-Index 6: Removing Zoom Out from Exp-Index 4 significantly improves most of the metrics compared to Exp-Index 4 and Exp-Index 5, which shows that excessive data augmentation can introduce noise or anomalies, making it more difficult for the model to learn effective features.
- Exp-Index 7: Removing Rotate from Exp-Index 6 further improves results across all five metrics compared to Exp-Index 4 and Exp-Index 6. This suggests that although the rotated dataset increased variability in turtle orientations and bounding box positions, it may have introduced unrealistic swimming behavior, which negatively impacts model performance.
- Exp-Index 8: This demonstrates that adding only the Extra Fish Dataset without using Depth-Aware Blur or other augmentation methods can still improve model performance compared to the baseline.
- Exp-Index 9 to 12: Repeating Exp-Index 4 to 7 with Gaussian Blur instead of Depth-Aware Blur yields similar outcomes. Combining the Extra Fish Dataset, Zoom Out, and Rotate together significantly improves performance (comparing Exp-Index 9 to Exp-Index 2). Removing the Extra Fish Dataset alone reduces performance (Exp-Index 10), while omitting Zoom Out (Exp-Index 11) or both Zoom Out and Rotate (Exp-Index 12) results in better metrics compared to including them.

The best combination, Exp-Index 7, shows that despite the strong performance of the baseline model, all metrics are further improved by using Depth-Aware Blur with adding the Extra Fish Dataset: a 2.4% accuracy, a 1.17% recall, a 2% precision, a 0.006 mAP@0.5, and a 0.083 mAP@0.5:0.95. This demonstrates that our method effectively enhances detection performance in challenging underwater conditions, such as low light and blur. For example, the enhanced model with data augmentation significantly outperforms the baseline when identifying smaller, distant turtles, as shown in Figure 3.

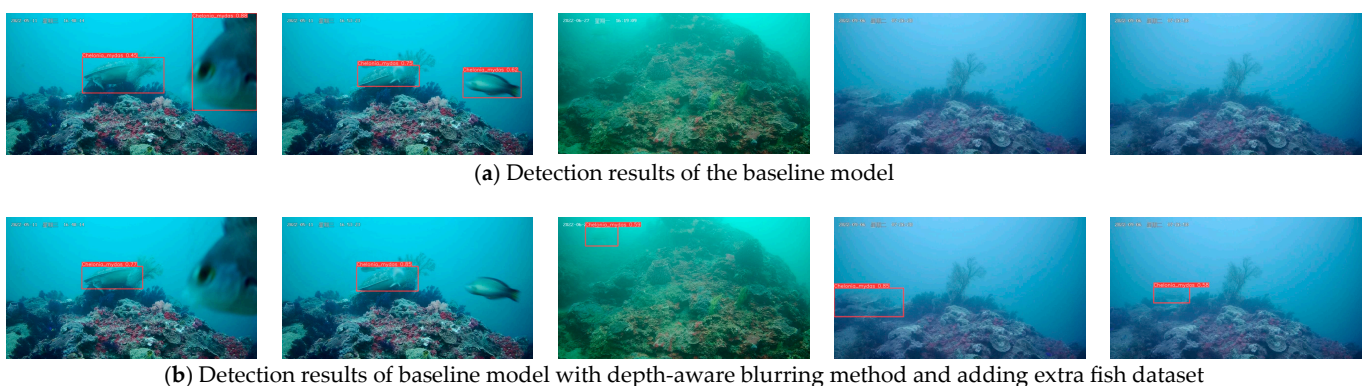


Figure 3. Comparisons between (a) baseline model and (b) baseline model with depth-aware blurring method and adding the extra fish dataset.

Using the best performer, the Exp-Index 7 method, we applied the model to detect untrained 60-day data from three sea areas, as shown in Table 3. The model successfully

identified five-time segments of sea turtle appearances, matching precisely with manual searches by marine researchers. Each sea area records video for 8 h per day, resulting in a total of 1,728,000 s of data, while the total appearance time of turtles is less than 40 s. Previously, marine researchers reviewed this vast footage to identify turtle appearances manually. With the AI model deployed, they now only need to inspect the time segments detected by the model, saving over 99% of the search and inspection time.

4. Conclusions

We developed a turtle detection method that enables marine researchers to efficiently identify time segments where turtles appear, using data augmentation techniques to enhance model accuracy and generalization ability. Despite the high effectiveness of the original model, combining depth-aware blur augmentation and incorporating the extra fish dataset further improves accuracy and stability, achieving 99.4% accuracy. In practical applications, this method saves the marine conservation team over 99% of the search time, significantly boosting efficiency. This AI-driven approach streamlines turtle monitoring, offering a sustainable solution for conserving marine species and ecosystems.

Author Contributions: Y.-C.H., Y.-T.P., W.-C.L., and L.-S.C. conceived the idea. Y.-C.H., J.-Y.C., and W.-C.L. completed the conceptualization and methodology. Y.-C.H., J.-Y.C., and W.-C.L. completed the data curation and validation, and developed the software and online service systems. W.-C.L., Y.-C.H. and Y.-T.P. wrote and revised the manuscript. L.-S.C., Y.-T.P. and W.-C.L. were in charge of the projects and funding resources. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported in part by the National Museum of Marine Science and Technology, the National Science and Technology Council, Taiwan, under Grants NSTC 112-2515-S-396-003, 113-2221-E-004-001-MY3, 113-2221-E-004-006-MY2, 113-2218-E-004-001, and 113-2634-F-002-008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are not publicly available due to ownership restrictions by the National Museum of Marine Science and Technology. The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: Yi-Chieh Hung, Jhih-Ya Chan, and Wei-Cheng Lien were employed by the company Hyson Technology Inc. and declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Fuentes, M.M.; Allstadt, A.J.; Ceriani, S.A.; Godfrey, M.H.; Gredzens, C.; Helmers, D.; Ingram, D.; Pate, M.; Radeloff, V.C.; Shaver, D.J.; et al. Potential adaptability of marine turtles to climate change may be hindered by coastal development in the USA. *Reg. Environ. Change* **2020**, *20*, 104. [\[CrossRef\]](#)
2. Bhalla, S.; Kumar, A.; Kushwaha, R. Analysis of recent techniques in marine object detection: A review. *Multimedia Tools Appl.* **2025**, *84*, 20339–20437. [\[CrossRef\]](#)
3. Redmon, J. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [\[CrossRef\]](#)
4. Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A. A Comprehensive review of yolo architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716. [\[CrossRef\]](#)
5. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [\[CrossRef\]](#)
6. Wong, S.C.; Gatt, A.; Stamatescu, V.; McDonnell, M.D. Understanding data augmentation for classification: When to warp? In Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, Gold Coast, Australia, 30 November–2 December 2016.
7. Wang, J.; Perez, L. The effectiveness of data augmentation in image classification using deep learning. *arXiv* **2017**. [\[CrossRef\]](#)

8. Peng, W.-Y.; Peng, Y.-T.; Chen, C.-S. Unveiling of How Image Restoration Contributes to Underwater Object Detection. In Proceedings of the IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), Taipei, Taiwan, 15–17 September 2021; pp. 1–2.
9. Sohan, M.; Ram, T.S.; Reddy, R.; Venkata, C. *Data Intelligence and Cognitive Informatics*; Springer: Cham, Switzerland, 2024; pp. 529–545.
10. Fayaz, S.; Parah, S.A.; Qureshi, G.J. Underwater object detection: Architectures and algorithms—A comprehensive review. *Multimedia Tools Appl.* **2022**, *81*, 20871–20916. [[CrossRef](#)]
11. Alomar, K.; Aysel, H.I.; Cai, X. Data augmentation in classification and segmentation: A Survey and New Strategies. *J. Imaging* **2023**, *9*, 46. [[CrossRef](#)] [[PubMed](#)]
12. Lu, Y.; Lin, Y.; Wu, H.; Xian, X.; Shi, Y.; Lin, L. SIRST-5K: Exploring massive negatives synthesis with self-supervised learning for robust infrared small target detection. *IEEE Trans. Geosci. Remote. Sens.* **2024**, *62*, 1–11. [[CrossRef](#)]
13. Roboflow Fish Image Dataset. Available online: <https://public.roboflow.com> (accessed on 28 May 2024).
14. Peng, Y.T.; Zhao, X.; Cosman, P.C. Single underwater image enhancement using depth estimation based on blurriness. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Québec City, QC, Canada, 27–30 September 2015; pp. 4952–4956.
15. Peng, Y.-T.; Cosman, P.C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **2017**, *26*, 1579–1594. [[CrossRef](#)] [[PubMed](#)]
16. Jian, M.; Yang, N.; Tao, C.; Zhi, H.; Luo, H. Underwater object detection and datasets: A survey. *Intell. Mar. Technol. Syst.* **2024**, *2*, 1–12. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.