Journal of Robotics, Networking and Artificial Life
Vol. 11(3); March (2026), pp. 226-235
ISSN (Online): 2352-6386; ISSN (Print): 2405-9021
https://alife-robotics.org/jrnal.html

Research Article

An Underwater Imaging System for Fish Recognition and Behavior Estimation to Optimize Sustainable Aquaculture Feeding

Raji Alahmad[1], Dominic Solpico[1], Shoun Masuda[1], Mohammad Albaroudi[1], Abdullah Alraee[1], Takahito Ishizuka[1], Kenta Naramura[1] Zhangchi Dong[1], Zongru Li[1], Hussam Alraie[2], Yuya Nishida[1], Kazuo Ishii[1]
[1]*Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan*
[2]*Middle East Technical University Northern Cyprus Campus, Güzelyurt, 99750, Mersin 10, Türkiye*
Email: raji@brain.kyutech.ac.jp, db.solpico@gmail.com, masuda.shoun528@mail.kyutech.jp, albaroudi.mohammad344@mail.kyutech.jp, alraee.abdullah-abdul648@mail.kyutech.jp, ishizuka.takahito370@mail.kyutech.jp, naramura.kenta804@mail.kyutech.jp, dong.zhangchi671@mail.kyutech.jp, li.zongru477@mail.kyutech.jp, hussam@metu.edu.tr, y-nishida@brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp

ARTICLE INFO

Article History

Received 25 November 2024
Accepted 20 February 2026

Keywords

YOLOv8
Sustainable aquaculture
Fish Detection
Underwater imaging

ABSTRACT

Aquaculture continues to expand as a response to rising seafood demand, but feeding remains a critical challenge due to its high costs and environmental impact. This study introduces an underwater imaging system that integrates video enhancement, YOLOv8-based fish detection, and velocity estimation to provide a data-driven solution for optimizing feeding strategies. Unlike conventional farmer intuition, the proposed approach offers objective monitoring of fish behavior under real aquaculture conditions. The system enhances underwater video quality by correcting color distortion, reducing noise, and sharpening contours, which improved fish detection accuracy from 69.3% to 73.2%. YOLOv8 achieved an overall detection accuracy of 85%, while velocity tracking successfully distinguished between normal and hunger-driven behaviors. These results confirm that fish velocity is a reliable indicator of feeding demand. By linking motion dynamics with feeding decisions, the system can reduce feed waste, lower costs, and improve fish health while minimizing environmental impacts. This work demonstrates the potential of integrating artificial intelligence and imaging technologies to establish standardized, sustainable, and more profitable aquaculture feeding practices. Future studies will focus on larger datasets, adaptive enhancement techniques, and real-time feeding control.

1. Introduction

With the increasing demand for seafood and declining marine fishery stocks, aquaculture has continued to expand and become a major contributor to seafood production worldwide. In 2022, the global seafood production was 214.92 million tons. Aquaculture contributed 59.1% (126.94 million tons), and capture fisheries accounted for 40.9% (87.99 million tons) [1]. This transition indicates a global reliance on aquaculture as a sustainable source of protein for human consumption, as natural fishery resources have reached or exceeded their biological limits. Therefore, the continuous rise of aquaculture represents both an opportunity and a challenge for the food production sector. However, the industry's expansion also raised sustainability concerns, including environmental degradation, biodiversity threats, and disease management, most of which stem from poor management practices. If not properly addressed, these issues could get out of control and cause harm not only to the aquatic environment but also to aquaculture production [2]. Environmental degradation includes eutrophication caused by uneaten feed, sediment built up under cages, and the spread of pollutants into surrounding waters. Biodiversity threats arise when escaped farmed fish compete with wild populations, leading to genetic mixing or displacement of native species. Furthermore, disease outbreaks can spread rapidly in densely stocked cages, causing mass mortality and major financial losses. These challenges highlight the urgent need for more efficient and sustainable aquaculture practices that balance production and environmental stewardship. Feeding optimization is one of the major challenges faced by the aquaculture industry. In many fish farming operations, feed accounts for a significant, often the largest, portion of total production expenses [3].

To promote optimal fish growth, farmers typically determine both the quantity and timing of feed distribution by observing the feeding behavior of the fish. However, these decisions are largely based on personal judgment and experience, leading to considerable variation between experienced and inexperienced farmers. The lower quality of fish managed by novice farmers often stems from inefficient feeding practices. Insufficient feeding can slow fish growth, while excessive feeding leaves uneaten feed that decomposes in the water, causing pollution and negatively impacting fish health and overall quality [4]. In addition, they incur more feeding expenses without substantially increasing the weight and quality of the fish. In the end, the farmers would make less profit or even incur losses from their operations. Feed optimization is not only an economic issue but also an environmental one. Feed represents up to 60–70% of production costs in intensive aquaculture systems, which means that even slight improvements in efficiency can lead to significant increases in profitability [5]. Simultaneously, poor feeding practices are among the largest contributors to water pollution in aquaculture farms, creating a direct link between farmer decision-making and environmental sustainability. Therefore, the ability to accurately measure feeding behavior is critical in modern aquaculture management.

Implementing digital transformation (DX) in aquaculture enables farmers to more precisely assess fish feeding behavior and optimize their feeding strategies. This approach can enhance fish growth performance and potentially boost profit margins. The widespread availability of digital cameras has made computer vision (CV) a commonly adopted technology in the aquaculture industry [6]. Furthermore, advancements in machine learning and artificial intelligence have facilitated the intelligent detection and analysis of fish behaviors [7]. These tools allow for the automation of tasks that were traditionally dependent on human judgment, such as the detection of feeding intensity, fish movement, and distribution within cages [8]. Artificial intelligence can also process large amounts of visual data more quickly and consistently than humans, offering opportunities to standardize feeding practices across farms, regardless of the farmers' experience level. Nevertheless, most existing studies have concentrated on recognizing feeding behaviors in controlled settings with stable lighting and clear water conditions. In contrast, the fluctuating conditions present in marine cages make it challenging to accurately assess fish feeding activity. Factors such as poor underwater visibility and the presence of non-fish objects further complicate the effectiveness of vision-based systems in detecting fish behavior. For instance, marine cages are often exposed to variable currents, plankton blooms, and suspended particles that reduce their visibility. In addition, floating debris and reflections from the water surface can generate noise in the video data. Fish may display unpredictable schooling behavior, making it difficult to distinguish between natural swimming and active feeding. These challenges necessitate the development of robust imaging systems that can operate effectively in real-world environments rather than only in laboratory settings.

This paper presents an underwater imaging system for recognizing fish-feeding behavior in marine fish cage environments. Here, we present our approach to enhancing videos of fish activity captured underwater and tracking their movement. This paper then discusses the relationship between the estimated fish motion and the farmer's feeding system to determine the optimal amount of feed to be supplied.

2. Methodology

Measuring the fish velocity is the key to understanding the fish behavior for the effective management and optimization of fishery operations. Measuring fish velocity also provides valuable insights into feeding strategies, welfare monitoring, and sustainable resource allocation in aquaculture. The biggest challenges in underwater video analysis are poor visibility, different lighting conditions, noise, and color distortion caused by wavelength-selective water absorption. These challenges introduce visual artifacts, reduce detection accuracy, and make it difficult to maintain consistent fish tracking over time. These challenges result in difficulties in fish detection and tracking. Without addressing these limitations, the reliability of fish recognition and motion analysis decreases significantly.

This research aimed to track the fish from videos of fish cages and estimate their velocity. It particularly focuses on cage-based aquaculture environments, where visual monitoring is essential for optimizing feeding and reducing operational costs. This is critical for maintaining sustainable and productive fisheries. The study was divided into three main objectives, as shown in Figure 1.

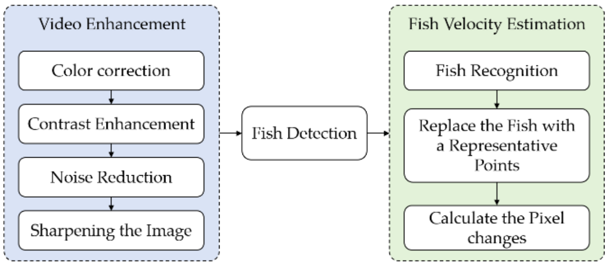

Figure 1. The flowchart of visual-based fish motion tracking

The first objective is video enhancement, which includes color correction, contrast adjustment, noise reduction, and sharpening of the frames to highlight fish contours. The second objective is fish detection, which bridges preprocessing with motion analysis by identifying the target fish within each frame. The third objective is velocity estimation, which is achieved through fish recognition, replacing detected fish with representative points, and calculating pixel displacements across frames to derive movement vectors.

2.1. *Experimental Setup*

For monitoring fish behavior in aqua cages, a dedicated observation system was installed outside the cage. This system was designed to provide continuous, noninvasive monitoring and consisted of multiple vertical cameras along with supplementary sensors. Each unit of the system was composed of underwater modules, each containing a network camera encased in watertight housing to ensure durability in marine conditions. These modules were vertically aligned at fixed intervals, creating an imaging arrangement with approximately 60% overlap for targets located 1 m from the cage. This overlap enabled the same fish to be captured simultaneously from multiple viewpoints, which is essential for accurate behavioral analysis and three-dimensional observation. The uppermost module was placed 0.5 m below the water surface to ensure that fish activity at both surface and deeper levels could be effectively recorded. Figure 2 illustrates the installation of the imaging system in the aquaculture cages.

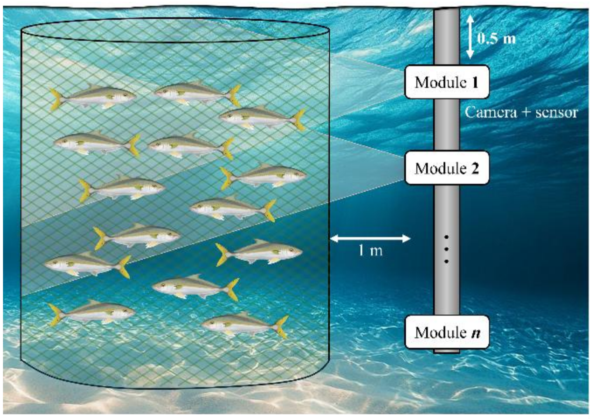

Figure 2. Installation of the imaging system in the aquaculture cages.

The collected image data formed the basis for the present study and were obtained from two different marine aquaculture sites representing realistic operational environments.

The first dataset was recorded in November 2022 from a yellowtail (*Seriola quinqueradiata*) cage operated by the Japan Fisheries Research and Education Agency in Gotō, Nagasaki, Japan. The second dataset was collected in October 2024 during a month-long field deployment on a fish-feeding vessel operated by Farm Choice Co., Ltd., in Amakusa, Kumamoto. These two datasets provide complementary perspectives on aquaculture practices under different conditions. Table 1 summarizes the experimental set-up parameters, including the camera arrangement, depth placement, and recording conditions. Together, these details form the foundation for subsequent analysis of fish behavior in aquaculture cages.

Table 1. Experimental setup parameters.

	Nagasaki	Kumamoto
Cage size (L, W, Z)	11.7 × 11.7 × 5m	12 × 12 × 13m
Fish type	Yellowtails (*Seriola quinqueradiata*)	
Fish age	6 months	1year
Number of fish	11,000	10,000
Module position	0.5 m from surface	

2.2. Video Enhancement

Enhancing underwater videos is challenging because of the unique water conditions. The wavelength-dependent absorption results in color distortion [9]. The blue color travels the longest distance in water, whereas the red color faces higher attenuation over short distances. Furthermore, image degradation is strongly affected by light attenuation and scattering [10]. Addressing these issues is essential for improving video quality and facilitating more accurate fish detection and tracking. Several steps were applied to enhance the underwater video, including color correction, contrast enhancement, noise reduction, and sharpening. The following sections describe each step of the video enhancement process.

2.2.1. Color correction:

Color correction compensates for wavelength absorption and restores the natural colors underwater [11]. The color correction process applied adjustments to the Lab color space, determined through empirical testing. The lightness channel (channel L) was saved to preserve the original brightness. A shift of 21.8 was applied to the green-red channel (channel a) to reduce green dominance and restore attenuated red components. The blue-yellow channel (b) was unchanged, ensuring that the natural distribution of blue and yellow tones remained intact. These practical adjustments enhanced the visibility of fish in underwater environments.

2.2.2. Contrast enhancement:

The histogram equalization technique was applied to improve the contrast of underwater video frames [12]. Unlike traditional histogram equalization, which applies a single global transformation to the entire image, this technique operates locally on small image regions, making it more effective in scenes with uneven illumination, such as underwater environments. The transformation function for histogram equalization in each local region is defined as Eq. (1).

$$s_k = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \quad (1)$$

Where s_k is the new pixel intensity after equalization, L denotes the total number of gray scale, $M \times N$ size of the local region, while n_j represents number of pixels with intensity j within the region.

By redistributing intensity values in this manner, the local contrast is improved while preserving the details in both dark and bright areas. This resulted in sharper fish outlines and more distinct features, facilitating both human inspection and subsequent automated image processing.

2.2.3. *Noise Reduction:*

Median filtering was used to reduce the noise in underwater video frames to preserve critical edge information [12]. This approach replaces each pixel with the mean of its neighbors according to Eq. (2).

$$\hat{f}(x,y) = median\ \{g(r,c)|(r,c)\} \in S_{xy} \tag{2}$$

Where S_{xy} denotes a neighborhood (sub-image) centered at point (x,y), and $g(r,c)$ represents the intensity values of the pixels within S_{xy}. The pixel itself at $(x,y)$is included in the computation of the median. In this work, a 3×3 neighborhood was employed. This approach is particularly effective for reducing impulsive noise, such as salt-and-pepper noise, while preserving edges that are essential for fish contour detection in underwater imagery.

2.2.4. *Image Sharpening:*

Unsharp masking was applied to sharpen video frames and highlight fish edges in underwater scenes. The method creates a blurred copy of the image, subtracts it from the original, and reintegrates the difference with a weight factor. It is expressed as:

$$g(x,y) = f(x,y) + k \cdot (f(x,y) - (f * h)(x,y)) \tag{3}$$

where $f(x,y)$ represents the original image, and $f * h$ denotes the blurred version using a Gaussian kernel h. k is the sharpening weight, and $g(x,y)$ is the output sharpened image. This technique aims to increase details visibility, which improves fish outline detection while preserving textural features important for analysis. Figure 3 shows a snapshot of the video enhancement output.

Figure 3. Image enhancement visualization.

2.3. ***Fish Detection***

Fish detection plays a critical role in ecological monitoring, aquaculture management, and behavioral analysis of aquatic species. Accurate detection enables researchers to study fish movement, feeding patterns, and population dynamics, which are essential for both conservation and commercial purposes. However, underwater environments present numerous challenges such as low visibility, varying light conditions, water turbidity, and background noise, which make traditional detection methods less effective. In recent years, deep learning–based object detection algorithms have shown significant promise in addressing these challenges. Among them, the You Only Look Once (YOLO) family of algorithms has become one of the most widely adopted approaches due to its balance of speed and accuracy [13-14], YOLO and its variants have been successfully applied in diverse applications, including face recognition, small object detection, and underwater object detection [15]. In this research, YOLOv8 was employed to detect the fish in video data. The process involved decomposing videos into individual frames, applying the YOLOv8 detection model to identify fish in each frame, and then reconstructing the processed frames into a video annotated with detection results. This approach provides a robust and efficient pipeline for fish detection in complex aquatic environments.

Two experiments were designed to evaluate the application of YOLOv8 for fish detection using different dataset sizes and training configurations. In the first experiment, a dataset of 150 images was used, divided into 106 images for training, 30 for validation, and 14 for testing. The model was trained for 150 epoch with a batch size of 16. All images were resized to 640×640 pixels to ensure consistency, and video data were processed at 15 frames per second (fps). This setup provided a baseline for assessing the feasibility of the YOLOv8 model under data-limited conditions.

The second experiment expanded the dataset to 600 images, with 500 allocated for training, 50 for validation, and 50 for testing. The same batch size, image resolution, and frame rate were used to maintain comparability with the first experiment. By introducing a larger dataset, the experiment allowed the model to be trained on more diverse fish appearances and environmental conditions, offering a broader foundation for learning.

To evaluate the performance of the trained models, several standard metrics were employed. The accuracy, precision, recall, and the mean Average Precision (mAP) were used as the primary detection metric, computed as the mean of the Average Precision (AP) values across all classes and Intersection over Union (IoU) thresholds.

2.4. ***Fish Velocity Estimation***

A variety of techniques, such as optical flow analysis, Kalman filtering, and the DeepSort framework, have previously been applied by researchers to estimate the fish behavior in the aqua cages [16-17]. Nevertheless, in challenging underwater environments, where fluctuations in illumination significantly affects pixel intensity, conventional tracking approaches become less effective because of the inherent difficulties in consistently extracting reliable object features.

Although YOLOv8 provides built-in multi-object tracking functionality (e.g., ByteTrack-based association), it was not adopted in this study because our objective was stable and computationally efficient velocity estimation under challenging underwater conditions rather than full multi-object trajectory management. In marine cages, frequent occlusion, rapid schooling behavior, illumination changes, and intermittent detections can cause ID switching and instability in integrated tracking frameworks.

To address these limitations, this study proposed a novel framework designed to simplify the process of motion tracking and velocity estimation.

Tracking fish movement across successive image frames generally requires following the complete fish body by identifying its features and measuring pixel displacements from one frame to the next. However, variations in lighting conditions and postures, further complicate the direct application of these methods. Consequently, substituting the entire fish with representative key points is identified as an essential step to improve robustness and computational efficiency. The representative point of each detected fish was defined as the geometric center of the YOLOv8 bounding box. After detecting a fish in each frame, YOLOv8 outputs a rectangular bounding box characterized by its top-left coordinate (x_{min}, y_{min}), The representative feature point (x_c, y_c) was computed as:

$$x_c = x_{min} + \frac{w}{2}, \quad y_c = y_{min} + \frac{h}{2} \tag{4}$$

This center point was used as the tracking reference for velocity estimation. The bounding box center was selected as the feature point because it remains stable despite fish posture changes and is more reliable than anatomical landmarks such as the head or tail, which are difficult to detect consistently in underwater conditions. It is also less sensitive to partial occlusion in dense schooling environments, as the bounding box is typically still detected even if part of the fish body is hidden.

This approach not only reduces computational complexity but also minimizes distortions that may arise from overlapping fish, occlusion, or rapid changes in orientation. By focusing on representative points rather than the entire fish body, the tracking algorithm becomes more robust and adaptable to real-world conditions. Moreover, the use of representative points ensured consistency in the velocity estimation even when the fish were only partially visible or blurred owing to sudden acceleration. Compared to conventional entire-body tracking, which often fails in low-contrast or noisy underwater scenes, this method enhances the accuracy of velocity calculation.

We implemented a simplified feature-point–based approach in which the center of each detected bounding box was used as the representative point and matched to the nearest point in the subsequent frame within a predefined distance threshold. If the displacement was below this threshold, the same ID was maintained. Velocity was calculated as:

$$v = \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{\Delta t} \tag{5}$$

Where Δt=1/15 s. This method reduces computational complexity, avoids long-term ID dependency, and provides stable short-term displacement measurements sufficient for reliable feeding behavior analysis.

Velocity estimation plays a critical role in behavior modeling. As changes in velocity are directly linked to feeding activity, stress responses, and social interactions, they provide a simplified yet highly informative input for aquaculture management systems. By incorporating the statistical averages of these motion patterns, farmers and researchers can better evaluate the collective behavior of fish schools, ultimately linking motion dynamics to feeding optimization and welfare monitoring.

3. Results and discussion

This section presents the fish detection results and the velocity estimation results obtained from the proposed system. Two experiments were conducted using different dataset sizes to evaluate performance under varying conditions.

With the enhanced video, the YOLOv8 was trained using two different datasets to evaluate its robustness. Experiment 1 used 150 images (70% for training, 30% for validation, and 10% for testing), while Experiment 2 expanded the dataset to 600 images (83.3% for training, 8.3% for validation, and 8.3% for testing). The trained models were then applied to separate images to evaluate performance both before and after enhancement.

3.1. *First Experiment:*

In the first experiment, the trained model was applied to two test sets: 50 images before enhancement and the same 50 images after enhancement. Figure 4 shows the training and validation loss curves with precision, recall, and mAP metrics over training epochs.

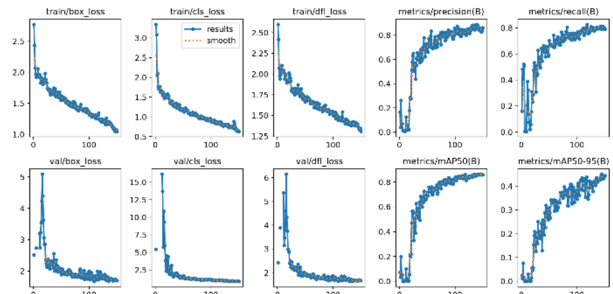

Figure 4. The training and validation loss curves with precision, recall, and mAP metrics of the first experiment.

Table 2 illustrates comparison of the precision, the recall, the mean average precision (mAP-50), and the system accuracy between the videos before and after enhancement.

Table 2. YOLOv8 results comparison between the videos before and after enhancement.

	Before enhancement	After enhancement
Precision	0.79	0.86
Recall	0.84	0.82
mAP-50 best	0.87	0.86
Accuracy	69.3%	73.2%

The table shows that while there was a slight decrease in the mAP after enhancement, the test accuracy increased from 69.4% to 73.3%. Precision also improved from 0.797 to 0.866, indicating that the enhanced video helped the model reduce false positives. Recall slightly decreased, suggesting a trade-off between precision and recall. These results highlight that image enhancement improves the reliability of detections but requires fine-tuning to maintain recall.

Further analysis of the 14-image testing set confirmed this trend. Out of 104 fish present, the model detected 77 fish (74%) before enhancement and 80 fish (76.9%) after enhancement, showing a small but consistent improvement in detection performance. Table 3 summarizes the comparison results of the deep analysis of the testing dataset before and after the enhancement.

Table 3. Comparison results of the deep analysis of the testing dataset before and after enhancement.

parameter	value
Test set	14 images
Same detection rate	3 images
Detect more fish before enhancement	5 images
Detected more fish after enhancement	6 images
Number of fish in the test set	104
Detected fish before enhancement, (detection rate)	77, (74%)
Detected fish after enhancement, (detection rate)	80, (76.9%)

The image enhancement gave a small improvement in detection, raising the rate from 74% to 76.9%. Some images improved, some got worse, but overall, a few more fish were detected after enhancement.

3.2. *Second Experiment:*

In the second experiment, the larger dataset produced more stable training and validation curves, as shown in Figure 5.

Unlike the first experiment, where fluctuations were visible due to limited training data, the second experiment shows smoother convergence. The training losses (box, classification, and distribution focal loss) decreased more steadily, while validation losses stabilized without extreme spikes.

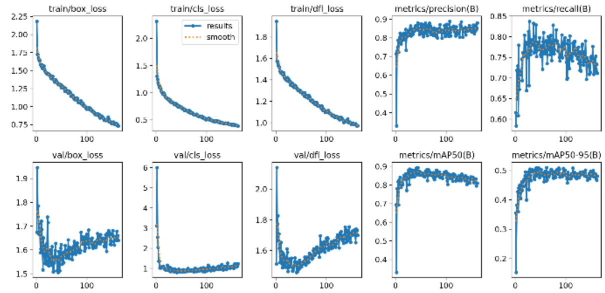

Figure 5. The training and validation loss curves with precision, recall, and mAP metrics of the second experiment.

Table 4 illustrates a comparison of the precision, the recall, the mean average precision, and the system accuracy between the videos before and after enhancement for the second experiment.

Table 4. YOLOv8 results comparison between the videos before and after enhancement in the second experiment.

	Before enhancement	After enhancement
Precision	0.86	0.77
Recall	0.83	0.88
mAP-50 best	0.90	0.89
Accuracy	74.4%	70.3%

The results indicate that precision decreased after enhancement (0.86 → 0.77), while recall improved (0.83 to 0.88). Accuracy showed a reduction from 74.4% to 70.3%, and mAP-50 remained nearly the same (0.90 vs. 0.89). It is important to note that in the second experiment, the enhancement parameters optimized for the first experiment were directly applied. This likely explains why accuracy decreased after enhancement. In other words, enhancement parameters that improved performance under a smaller dataset did not generalize as well to the larger dataset.

Like the first experiment, a comprehensive analysis of the test set, Table 5 illustrates the comparison results of the deep analysis of the testing dataset before and after the enhancement.

Table 5. Comparison results of the deep analysis of the testing dataset before and after enhancement.

parameter	value
Test set	50
Same detection rate	16
Detect more fish before enhancement	26 images
Detected more fish after enhancement	8 images
Number of fish in the test set	589
Detected fish before enhancement, (detection rate)	576, (97.8%)
Detected fish after enhancement, (detection rate)	80, (91.3%)

This result suggests that unlike in some cases where enhancement improves visibility, in this dataset the enhancement did not provide a consistent benefit. While a few images showed better detection after enhancement, the majority either remained unchanged or worsened. The process slightly reduced the total detection rate, highlighting that enhancement techniques should be carefully adapted to the new underwater environment.

Overall performance comparison between the two experiments shows that increasing the dataset size improved the system accuracy before enhancement from 69.3% in the first experiment to 74.4% in the second experiment. This highlights the importance of larger and more diverse datasets in strengthening the robustness of the detection model. However, the decrease in accuracy after enhancement in Experiment 2 further emphasizes that enhancement techniques need to be optimized for specific conditions to fully realize their benefits.

3.3. Velocity Estimation:

To estimate the fish's motion and velocity, the trained YOLOv8 model was first used to detect the fish in each frame. After detection, every fish was represented by a single point. Each detected fish was also assigned a unique ID. If in the following frame a point was found close to the previous location compared to a specific threshold, the same ID was kept. This allowed continuous tracking of the same fish across frames. Figure 6 demonstrates velocity estimation of fish movement.

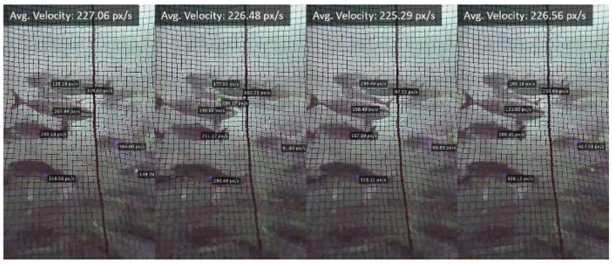

Figure 6. Estimated fish velocity.

The velocity was calculated from the distance between corresponding points in two consecutive frames, frames (i) and (i+1), This made it possible to measure the movement of each individual fish. For each frame, the system then calculated the average velocity of all detected fish present within the frame, giving an overall view of group movement.

This method was tested on two videos: the case of normal fish movement, and the case during fish feeding. Choosing two different videos serves two purposes.

1- To validate and confirm the detection results of the trained model, since the fish's random and high-velocity movements might affect the detection due to the quality reduction of the video.

Figure 7 illustrates the number of detected fish in the two described cases.

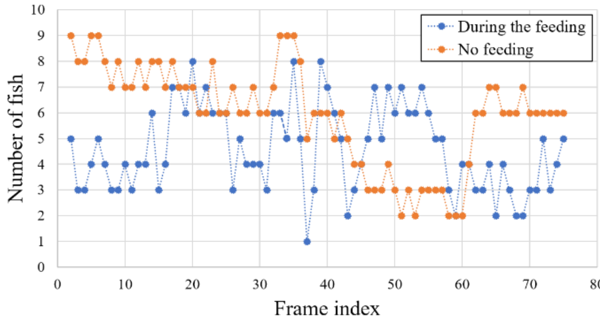

Figure 7. Number of detected fish in two cases (during the feeding presses and without feeding).

According to Figure 7, a greater number of fish were detected during the feeding period compared to when no feeding occurred. This difference is attributed to the fish movement patterns. When feeding takes place, the fish, being hungry, swim in various directions, including toward the camera's viewing axis. In such cases, the detection algorithm struggles to identify these fish accurately. Additionally, the increased swimming speed during feeding reduces video clarity, which further weakens the detection performance.

2- To assess the velocity estimation method under varying fish speed, Figure 8 presents the estimated fish velocity for the two videos scenarios: one captured during feeding and another without feeding.

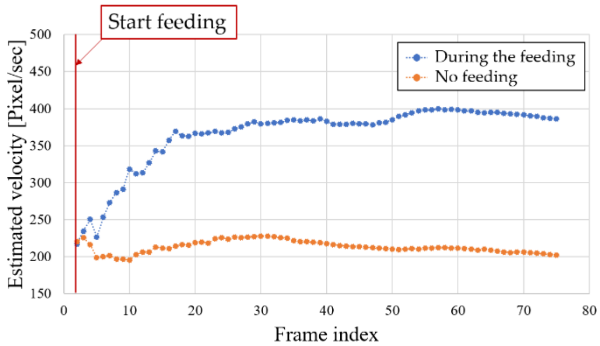

Figure 8. Estimated fish velocity.

In the figure, the orange line corresponds to the normal (non-feeding) condition, while blue line represents the feeding condition. The results show that the estimated fish velocity is higher during feeding compared to normal behavior. This velocity was measured in pixels per second by pixel per second.

4. Conclusion

This research introduced a novel underwater imaging system that combines video enhancement, deep learning–based fish detection, and velocity estimation. to improve aquaculture feeding practices. Unlike conventional farmer intuition, the proposed system offers an objective, data-driven approach for optimizing feeding operations.

Experimental results demonstrated that image enhancement improved overall detection reliability, with precision gains outweighing minor losses in recall. The YOLOv8 model, trained on real-world datasets, effectively detected fish under challenging underwater conditions, while velocity tracking confirmed clear behavioral differences between normal and feeding states. These findings validate the feasibility of using fish velocity as a practical indicator for hunger-driven activity and feed demand.

The broader significance of this work lies in its potential to reduce feed waste, mitigate water pollution, and enhance fish welfare, critical elements for the long-term sustainability of aquaculture. By shifting from subjective human judgment to AI-assisted monitoring, the system establishes a framework for standardized feeding strategies across farms, regardless of farmer experience.

Future research should focus on scaling the dataset, adapting enhancement parameters to diverse environmental conditions, and integrating real-time feedback into automated feeding systems. Ultimately, this approach moves aquaculture closer to intelligent, closed-loop management where technology directly supports both economic profitability and environmental stewardship.

Acknowledgements

This research was supported by the research program on development of innovative technology grants from the Project of the Bio-oriented Technology Research Advancement Institution (BRAIN) (No. JPJ007097). We would like to acknowledge Dr. Hiroyasu Mizoguchi, Dr. Akihiko Matsuda, and Dr. Yuichiro Fujinami of the Japan Fisheries Research and Education Agency, and Mr. Tokuo Suetsugu and Yoshinori Yatsunami of Belltechne Co., Ltd., for their collaboration in this research.

References

1. Food and Agriculture Organization of the United Nations (FAO), via World Bank, Capture and aquaculture production, Our World in Data, 2025, https://ourworldindata.org/grapher/capture-and-aquaculture-production, accessed February 20, 2026.
2. M. Martinez-Porchas and L. R. Martinez-Cordova, World aquaculture: environmental impacts and troubleshooting alternatives, The Scientific World Journal 2012, 2012.
3. W. R. Rola and M. R. Hasan, Economics of aquaculture feeding practices: a synthesis of case studies undertaken in six Asian countries, in Economics of aquaculture feeding practices in selected Asian countries, ed. M. R. Hasan, Food and Agriculture Organization, Rome, 2007, pp. 1–31.
4. M. Sun, S. G. Hassan and D. Li, Models for estimating feed intake in aquaculture: a review, Computers and Electronics in Agriculture 127, 2016, pp. 425–438.
5. S. Duan, S. Gao, X. Xu, Q. Wei, Z. Tao, G. Xu, Q. Liu, B. Wei and C. He, The impact of feeding rates on the growth, stress response, antioxidant capacity, and immune defense of koi (Cyprinus carpio var. koi), Fishes 10, 2025, 181.
6. M. Saberioon, A. Gholizadeh, P. Cisar, A. Pautsina and J. Urban, Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues, Reviews in Aquaculture 9(4), 2017, pp. 369–387.
7. C. Zhou, K. Lin, D. Xu, L. Chen, Q. Guo, C. Sun and X. Yang, Near infrared computer vision and neuro-fuzzy model-based feeding decision system for fish in aquaculture, Computers and Electronics in Agriculture 146, 2018, pp. 114–124.
8. M. Cui, X. Liu, H. Liu, J. Zhao, D. Li and W. Wang, Fish tracking, counting, and behaviour analysis in digital aquaculture: a comprehensive survey, Reviews in Aquaculture 17(1), 2025, e13001.
9. C. Li et al., An underwater image enhancement benchmark dataset and beyond, IEEE Transactions on Image Processing 29, 2020, pp. 4376–4389.
10. P. Sahu, N. Gupta and N. Sharma, A survey on underwater image enhancement techniques, International Journal of Computer Applications 87(13), 2014, pp. 19–23.
11. S. S. Sankpal and S. S. Deshpande, A review on image enhancement and color correction techniques for underwater images, Advances in Computational Sciences and Technology 9(1), 2016, pp. 11–23.
12. R. C. Gonzales and R. E. Woods, Digital image processing, 4th edn., Pearson, 2018.
13. M. Albaroudi, R. Alahmad, A. Alraee and K. Ishii, Evaluating tree branch recognition algorithm in pruning robots under augmented environmental conditions, Proceedings of the International Conference on Artificial Life and Robotics (ICAROB), 2025, pp. 356–360.
14. A. Alraee, R. Alahmad, H. Alraie, M. Albaroudi and K. Ishii, Efficient ball position estimation for tennis court robot assistants using dual-camera system, Proceedings of the International Conference on Artificial Life and Robotics (ICAROB), 2025, pp. 378–381.
15. Z. Li, H. Alraie, D. Solpico, Y. Nishida, K. Ishii, Y. Yatsunami, M. Fuchigami and T. Suetsugu, Recognition of fish in aqua cage by machine learning with image enhancement, Proceedings of the 2024 IEEE/SICE International Symposium on System Integration (SII), 2024, pp. 637–643.
16. G. Wang, A. Muhammad, C. Liu, L. Du and D. Li, Automatic recognition of fish behavior with a fusion of RGB and optical flow data based on deep learning, Animals 11, 2021, 2774.
17. S. Liu, J. Zhang, H. Zheng, C. Qian and S. Liu, An improved DeepSORT-based model for multi-target tracking of underwater fish, Journal of Marine Science and Engineering 13, 2025.

Authors Introduction

Dr. Raji Alahmad

He is currently a researcher at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his MEng degree in 2018 and his DEng in 2022 at Kyushu Institute of Technology, Japan. His research interests are AI algorithms, Field Robotics, Underwater communication, Image processing, and Neural networks.

Dr. Dominic Solpico

He obtained his M.S. in Electronics Engineering at the Ateneo de Manila University, Philippines in 2015. He obtained his Ph.D. at Kyushu Institute of Technology, Japan in 2022. He worked as a researcher at the Department of Human Intelligence Systems of the same university from 2022 to 2025. His research interests are in the fields of Intelligent Aquaculture, Wireless Sensor Networks (WSNs), and Field Robotics.

Mr. Shoun Masuda

He received his Bachelor's degree in 2024 from Hiroshima Institute of Technology in Japan. He is currently a master's student at Kyushu Institute of Technology

Mr. Mohammad Albaroudi

He received his BSc in telecommunication engineering from Al-Wataniya Private University WPU, Syria, in 2018. M.Sc. degree in Radio and Mobile Telecommunication Systems from the Higher Institute of Applied Sciences and Technology, Syria, in 2021. He is currently pursuing a Ph.D. at Kyushu Institute of Technology, Japan. His research interests are Machine learning algorithms, AI, Image processing, and Antenna design.

Mr. Abdullah Alraee

He received his B.Sc. degree in telecommunication engineering from Al-Wataniya Private University WPU, Syria, in 2018, and M.Sc. degree from Kyushu Institute of Technology, Japan, in 2024. He is pursuing a Ph.D. at Kyushu Institute of Technology, Japan. His research interests are Robotics, Artificial Intelligence, Image processing, Automation, and machine learning.

Mr. Takahito Ishizuka

He received his Bachelor's Degree in Engineering from Fukui University of Technology in 2024. He is currently a master student in Kyushu Institute of Technology, Japan

Mr. Kenta Naramura

He received his Bachelor's degree in Engineering in 2024 from National Institute of Technology, Tsuyama College in Japan. He is currently a master student in Kyushu Institute of Technology, Japan.

Mr. Zhangchi Dong

He received his bachelor's degree from Nanjing University of Aeronautics and Astronautics, China, in 2021. In 2024, he started his master's studies at Kyushu Institute of Technology, majoring in Human Intelligence Systems. His research interests include computational fluid dynamics simulation and robotics.

Mr. Zongru Li

He earned a bachelor's degree from Zhengzhou University, china in 2021 and is currently pursuing a master's degree at Kyushu Institute of Technology in Japan.

Dr. Hussam Alraie

He received the M.Sc. degree in Telecommunication Engineering from Homs University, Syria, and the Ph.D. degree from Kyushu Institute of Technology, Japan, in 2024. He is currently a postdoctoral researcher and part-time lecturer at Middle East Technical University, Northern Cyprus campus (METU-NCC), Türkiye. His research interests are Underwater communication, Image and signal processing, Robotics, Machine learning, and Artificial Intelligence.

Dr. Yuya Nishida

He is an Associate Professor at the Graduate School of Life Science and System Engineering, Kyushu Institute of Technology, Japan. His research area is in the field of robotics, its application, and data processing.

Prof. Kazuo Ishii

He is currently a Professor at the Department of Human Intelligence Systems of Kyushu Institute of Technology, Japan. He obtained his M. S. degree in 1993 and his D. Eng. degree in 1996 at The University of Tokyo. His research interests are in the fields of Underwater Robotics, Field Robotics, Neural Networks, and Intelligent Systems.