**EMITTER International Journal of Engineering Technology**
Vol. 14, No. 1, June 2026, pp. 42~59
DOI : 10.24003/emitter.v14i1.1007

# Exploring YOLO-Based Deep Learning Approaches for Fish Detection in Intelligent Aquatic Monitoring Systems

**Tresna Dewi, Riyo Irawan, Agum Try Wardhana, Muhammad Amri Yahya, Lukman Nul Hakim, and Dini Septiyani AR**

Department of Electrical Engineering
Politeknik Negeri Sriwijaya
Jalan Srijaya Negara, Palembang, 30139, Indonesia
Corresponding Author: tresna_dewi@polsri.ac.id

*Received March 2, 2026; Revised April 5, 2026; Accepted June 3, 2026*

**Abstract**

The Advancements in precision aquaculture demand robust visual monitoring systems capable of accurate, real-time fish detection in complex underwater environments characterized by turbidity, occlusion, and dynamic illumination. While YOLO (You Only Look Once) architectures have demonstrated high efficiency in object detection tasks, their comparative performance for underwater fish detection remains underexplored, particularly across recent variants such as YOLOv5, YOLOv8, and YOLOv11. This study presents a systematic evaluation of three state-of-the-art YOLO models using a curated GlowFish dataset consisting of 533 annotated images across three fluorescent species. Data were acquired under controlled but visually diverse conditions using multi-angle imaging and standardized illumination. A uniform training pipeline, consistent annotation using the COCO format, and identical hyperparameters were applied across models to ensure fair benchmarking. Key evaluation metrics include precision, recall, mAP@0.5, and mAP@0.5:0.95. Experimental results reveal that YOLOv5 achieved the highest precision (0.963) and mAP@0.5 (0.967), while YOLOv8 delivered superior recall (0.930) and more balanced detection across species classes. YOLOv11 demonstrated architectural potential but showed greater sensitivity to class imbalance and reduced confidence stability. Visual analysis and confusion matrices further confirmed model-specific trade-offs in classification reliability and localization precision. This work contributes critical empirical insights into the selection of YOLO architectures for intelligent aquaculture systems, offering practical guidance for real-time aquatic monitoring deployments. Future research will extend this framework to multi-species, multi-environment datasets, integrate spatiotemporal behavioral tracking, and investigate deployment on resource-constrained edge-AI platforms, advancing the field toward interpretable and autonomous aquatic monitoring solutions.

**Keywords**: Deep Learning for Fish Detection; Intelligent Aquaculture Monitoring; Real-Time Aquatic Surveillance; Underwater Computer Vision; YOLO Object Detection.

## 1. INTRODUCTION

Advancements in aquaculture have been critical in addressing the growing global demand for protein, yet they bring substantial challenges related to productivity, environmental sustainability, and operational efficiency. Intelligent aquatic monitoring systems have emerged as essential tools to support these objectives by enabling real-time surveillance of fish health, behavior, mortality, and population dynamics[1-5]. Traditional monitoring approaches—such as manual visual observation, sonar-based detection, and net sampling—often suffer from limitations in scalability, temporal resolution, and species-level accuracy, particularly in visually complex aquatic environments characterized by occlusion, turbidity, and dynamic lighting[6].

The integration of computer vision (CV) and deep learning (DL) has led to significant progress in the automation of fish detection and classification tasks[6-10]. In this domain, the You Only Look Once (YOLO) family of convolutional neural networks has emerged as a dominant framework for real-time object detection. Specifically, the successive developments of YOLOv5, YOLOv8, and the more recent YOLOv11 have exhibited substantial improvements in detection accuracy, multi-scale feature learning, and inference efficiency, making them particularly well-suited for deployment in aquatic monitoring systems[11-14].

YOLOv5 has demonstrated high precision and low-latency performance across various aquatic and agricultural applications. Its architectural simplicity, deployment flexibility, and robust object detection capability have been applied in tasks such as plant species identification[11], fish classification[7], and automated harvesting systems [15][16]. YOLOv8 introduces significant enhancements through its extended decoupled head architecture and attention mechanisms, which improve object localization and classification under challenging conditions such as overlapping fish schools, low contrast environments, and rapid fish movement[12][13]. Studies employing YOLOv8 have reported its effectiveness in UAV-based dead fish detection[12], foreign object recognition in aquaculture products[12], and underwater object analysis with fine-grained segmentation[14].

Although YOLOv11 has only recently emerged in the object detection landscape, preliminary evaluations suggest substantial improvements over previous iterations in terms of real-time processing speed and precision in low-visibility aquatic scenes. While academic literature on YOLOv11 remains limited due to its novelty, its anticipated adoption builds upon the empirical successes and structural refinements introduced in YOLOv8 [12][14].

The application of these YOLO models extends across various intelligent aquatic monitoring functions. Visual monitoring systems based on YOLO architectures have facilitated age prediction through otolith imaging[7], fish length estimation via mobile applications[10], behavioral response detection under variable oxygen conditions[18], and image-based health diagnostics during transport[4]. Moreover, YOLO-powered detection systems contribute to ecological studies involving fish migration[10], anomaly detection in electronic monitoring data [19], and automatic classification of aquatic species and substrates [8][9][20][21].

The role of YOLO-based vision systems becomes especially significant in sustainable aquaculture systems such as aquaponics. These integrated systems combine fish farming and hydroponic plant cultivation, requiring continuous monitoring of biomass, water quality, and nutrient cycling to maintain ecological balance [22-24]. Recent research has introduced hybrid solar-hydro energy management strategies with deep learning forecasting models—such as Random Forest and LSTM—for predictive control in aquaponic operations [25-27]. YOLO-based detection technologies support such systems by offering real-time monitoring solutions compatible with energy-constrained environments [22][26], while also contributing to system automation, data-driven feedback loops, and visual quality assurance [22][28][29].

YOLO's adaptability in aquaponic and aquaculture systems also extends to multi-species co-cultivation setups, real-time shrimp growth tracking[22], and automated disease prevention through early-stage visual indicators [24]. In addition, YOLO-based object detection has been employed in aquaculture imaging for reproductive ultrasound classification, economic forecasting of aquatic product prices, and small marine organism recognition [19].

Despite these advancements, current research lacks comprehensive comparative analysis of YOLOv5, YOLOv8, and YOLOv11 performance in aquatic monitoring contexts. Few studies rigorously benchmark these models under varying conditions of occlusion, turbidity, overlapping fish, and low-light environments. Furthermore, the majority of existing implementations are limited to single-environment validation, often omitting evaluation across different aquaculture modalities, such as tank-based systems, open-water farms, and closed-loop aquaponics [8][9][11-13][19].

The present study addresses this gap by conducting a thorough performance assessment of YOLOv5, YOLOv8, and YOLOv11 for fish detection in intelligent aquatic monitoring systems. The evaluation includes metrics such as mean average precision (mAP), inference time, robustness to occlusion, and adaptability to underwater environmental variability. This investigation aims to provide foundational insights for the implementation of real-time, DL-based visual monitoring frameworks in aquaculture and aquaponics, contributing to the design of intelligent, scalable, and sustainable aquatic production systems.

## 2. RELATED WORKS

The evolution of intelligent aquatic monitoring systems is driven by the convergence of traditional fish observation methods, computer vision-based detection, and deep learning-driven object recognition, with recent emphasis on YOLO-based models for real-time and scalable deployment.

Conventional fish monitoring methods rely on manual and acoustic surveys, providing essential data on fish abundance, distribution, and behavior. In-situ acoustic measurements and snorkel-based observations have long served as baseline methods for fisheries research [1], [2], [3], [5]. Dunning et al. [1] demonstrated the high variability in fish target strength in broadband measurements, while Evans et al. [2] analyzed fish avoidance of ships during acoustic surveys using quiet, uncrewed surface vessels. Similarly, Williams et al. [3] investigated fish behavioral responses to approaching underwater cameras, highlighting biases in visual surveys. Complementary approaches, such as integrated snorkel surveys with uncertainty modeling [5], provide ecological context but remain labor-intensive, invasive, and limited in temporal resolution. These challenges catalyzed the transition toward automated and non-invasive monitoring solutions.

The advent of deep learning (DL) and computer vision (CV) has transformed fish monitoring by enabling real-time detection, tracking, and behavior recognition. Early DL applications focused on species classification and event detection using multi-stage pipelines. For example, Saqib et al. [6] applied computer vision and AI for automated fishing event detection and species classification in electronic monitoring systems, while İşgüzar et al. [7] developed FishAgePredictioNet, a multi-stage framework combining segmentation, CNNs, and Gaussian process regression for age prediction from otolith images. Additionally, Lin et al. [10] introduced a machine vision approach for monitoring and quantifying fish school migration, demonstrating the potential of vision-based ecological analytics. Acharya et al. [18] extended DL to automatic anomaly detection in fisheries, whereas Shibata et al. [17] demonstrated a smartphone-based fish length estimation model using deep learning for practical field deployment. These studies collectively illustrate the shift from manual surveys to semi- and fully automated DL-driven monitoring systems.

Among various DL models, YOLO (You Only Look Once) has emerged as a dominant framework for real-time object detection in aquatic environments. YOLO-Fish demonstrated robust fish detection in realistic underwater conditions, achieving resilience to turbidity and occlusion [8]. HD-YOLO, designed for detecting Aurelia hydroids, showcased the adaptability of YOLO to marine invertebrate monitoring [9]. More recent adaptations have introduced attention mechanisms [13], microalgae segmentation in complex environments [29], and SCoralDet for soft coral detection [19]. Deep-sea applications include mobile crawler platforms leveraging YOLO for automated species classification and counting, enabling large-scale benthic monitoring [28]. In aquaculture operations, real-time dead fish detection using YOLOv8 integrated with UAVs supports autonomous farm management and mortality mitigation [12], while lightweight YOLO frameworks for recirculating aquaculture systems (RAS) enhance on-site operational safety [14]. These YOLO-based advances underscore the technology's flexibility, speed, and practical utility in underwater monitoring.

Beyond standalone detection, intelligent aquaculture and aquaponics systems integrate DL models with IoT-enabled environmental monitoring and energy optimization. Dewi et al. [16] applied fuzzy logic-based visual servoing for robotic strawberry harvesting, illustrating cross-domain CV expertise. Paul et al. [15] demonstrated YOLO's potential in multi-task agricultural detection, while Clinton et al. [11] employed YOLOv5 for plant species identification, bridging agricultural and aquatic monitoring concepts. In integrated aquaponic systems, hybrid ML/DL models have been proposed for PV output forecasting and energy management [23], [24], and smart aquaponics powered by hybrid solar-hydro energy systems with DL-based optimization has been implemented for sustainable aquaculture [24]. Research by Alarcón-Silvas et al. [21] and Suárez-Cáceres et al. [22] further highlights the importance of water quality and disease susceptibility monitoring, while Oktarina et al. [25] extended DL to agrivoltaic chili growth prediction, underscoring cross-domain AI integration. Moreover, fully automated market prediction for aquatic products using ML and DL has been demonstrated in Taiwanese wholesale systems [26], showcasing the wider economic integration of intelligent monitoring frameworks.

Despite these advances, critical gaps remain. Most YOLO-based implementations are species-specific and often validated under controlled laboratory or semi-natural conditions [10], [11], [26]. Environmental variability—including light scattering, turbidity, and occlusion—continues to challenge detection robustness, motivating synthetic data augmentation, transfer learning, and multimodal sensor fusion approaches [6], [7], [27]. Meanwhile, fully integrated, edge-deployed, and energy-aware YOLO-based pipelines for multi-species and behavior-sensitive fish monitoring are still in early development [12], [14], [23]–[25]. Addressing these gaps is essential to realize the vision of autonomous, sustainable, and ecologically intelligent aquatic monitoring systems.

## 3. ORIGINALITY

This study contributes to the advancement of YOLO-based deep learning for fish detection in aquaculture by concentrating exclusively on YOLOv5, YOLOv8, and YOLOv11, the most relevant and widely adopted architectures for real-time visual monitoring. Existing fish detection research remains fragmented, often focusing on single-species recognition or laboratory-controlled conditions that fail to capture the operational complexity of aquaculture environments. Conventional monitoring methods, whether visual or acoustic, are constrained by manual labor, fish behavioral responses, turbidity, and variable lighting, underscoring the need for robust, real-time detection models optimized for deployment in production settings.

The originality of this work lies in two aspects. First, it delivers a focused synthesis of YOLOv5, YOLOv8, and YOLOv11 models, analyzing their capabilities and limitations for real-time fish detection, mortality identification, and operational decision support in aquaculture. This emphasis on the current generation of YOLO architectures reflects the progression from YOLOv5's balance of speed and accuracy, to YOLOv8's enhanced feature representation and robustness, and finally to YOLOv11's emerging ability to handle complex and dynamic aquatic scenes efficiently. Second, the study identifies the key research gaps that limit the effectiveness of current YOLO-based approaches, including: (1) environmental sensitivity to turbidity, occlusion, and low-light conditions; (2) limited multi-species and behavior-aware detection, restricting comprehensive stock assessment; and (3) the need for lightweight, deployment-ready models that deliver high accuracy under real-world farm constraints.

By consolidating insights from YOLOv5, YOLOv8, and YOLOv11 and highlighting these operational challenges, this study defines a clear pathway toward scalable, accurate, and sustainable fish monitoring solutions, supporting the modernization of aquaculture with deep learning-driven visual intelligence.

## 4. SYSTEM DESIGN

This study benchmarks the performance of three advanced YOLO-based deep learning models—YOLOv5, YOLOv8, and YOLOv11—for automated fish detection in underwater environments, and the proposed method is given in Figure 1.

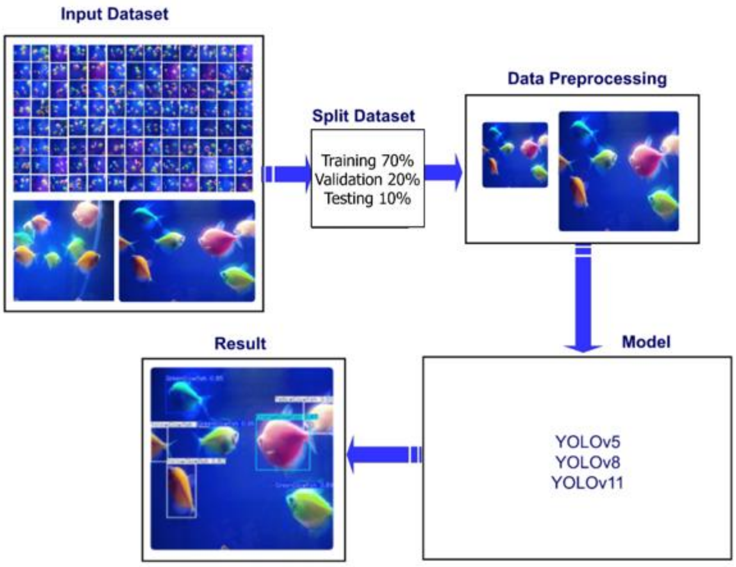



**Figure 1.** Proposed method workflow.

A standardized deep learning pipeline is adopted, encompassing data acquisition, preprocessing, model training, evaluation, and performance comparison. The aim is to assess each model's capability in terms of detection accuracy, inference speed, and robustness under varying aquatic conditions, providing insights for deploying intelligent monitoring systems in aquaculture and marine research.

### 4.1 Material

#### 1) Dataset and Experimental Setup

To evaluate the YOLO models under realistic aquatic conditions, a dataset was developed using a controlled aquarium-based environment, as illustrated in Figure 2. The experimental aquarium was equipped with adjustable LED lighting, filtration systems, and high-resolution imaging devices to simulate variable underwater conditions such as turbidity, illumination gradients, and occlusion effects.

**Figure 2.** Controlled aquarium setup for GlowFish dataset acquisition.

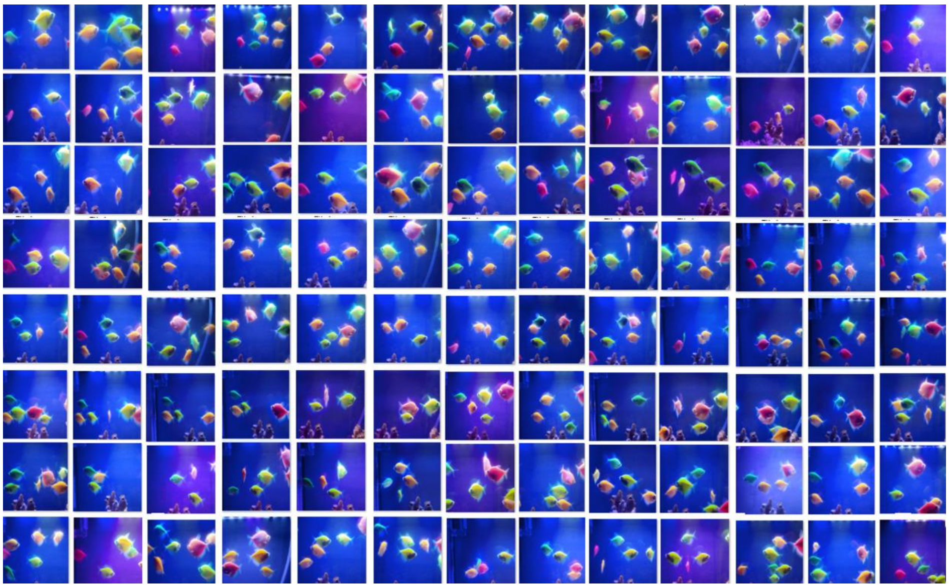


**Figure 3.** Data set captured from Figure 2.

The primary subjects are GlowFish species in Figure 2 with high visual contrast due to genetically induced fluorescence. The dataset illustrated in Figure 3 comprises 533 annotated images at a resolution of 640×640 pixels, capturing three fish categories: GreenGlowfish (1,069 instances), YellowGlowfish (1,007 instances), and OrangeGlowfish (885 instances).

To capture rich visual diversity, multi-angle imaging was applied (top-down and lateral views), ensuring varied spatial arrangements and behavioral dynamics. The dataset was split into training (70%), validation (20%), and testing (10%) subsets to maintain statistical rigor in model development.

### 2) Annotation Tools

All images were manually annotated using the COCO bounding box format via Roboflow shown in Figure 4, ensuring consistency and facilitating supervised learning. Annotated classes included the species label and precise bounding box coordinates.

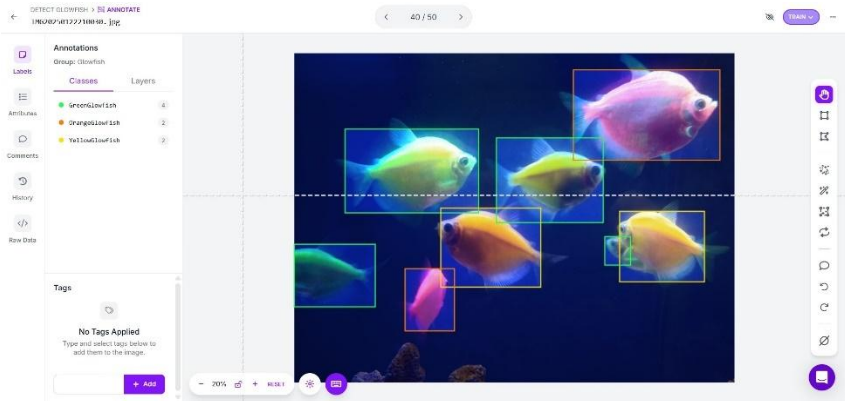



**Figure 4.** Annotation in Roboflow.

## 4.2 Methods

### 1) Preprocessing and Normalization

All images were resized to 640×640 pixels, and orientation corrections were applied. No artificial data augmentation was used, preserving the natural variability of aquatic imagery. Pixel values were normalized to standardize the input across the YOLO models.

### 2) YOLO Model Architectures

Three YOLO architectures—YOLOv5, YOLOv8, and YOLOv11—were evaluated under a consistent experimental framework to compare their object detection performance. The models represent different generations of the YOLO family and incorporate distinct architectural improvements in feature extraction, feature fusion, and detection mechanisms.

- YOLOv5 is a mature and widely adopted model that balances detection speed and accuracy, featuring PANet for feature aggregation and CSPDarknet as the backbone.
- YOLOv8 introduces an anchor-free detection head, improved decoupled heads, and an updated backbone design, enhancing object detection stability and multi-scale feature representation.
- YOLOv11, the most recent version, integrates Bidirectional Feature Pyramid Networks (BiFPN) and attention-based modules for fine-grained object discrimination in cluttered visual scenes.

Each model processes input through three main stages:

- Stem for low-level feature extraction
- Backbone for hierarchical representation
- Neck–Head module for scale-invariant detection and prediction

3) Model Training

Training was conducted using GPU acceleration, with identical hyperparameter settings across all models to ensure comparability (given in Table 1).

**Table 1.** Training Configuration and Processing Time for YOLO Models

| Model | Image Training | Batch Size | Training Time (s) | Epochs |
|---|---|---|---|---|
| YOLOv5 | 533 | 16 | 17.32 | 100 |
| YOLOv8 | 533 | 16 | 18.15 | 100 |
| YOLOv11 | 533 | 16 | 19.1 | 100 |

All models were trained using the Adam optimizer with an initial learning rate of 0.0001. Loss functions included bounding box regression loss, objectness loss, and classification loss, optimized concurrently during backpropagation.

**4) Evaluation Metrics**

The performance of each YOLO model was evaluated using the following metrics:

- **Precision (P):**

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \tag{1}$$

- **Recall (R):**

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \tag{2}$$

- **mAP:** Mean average precision at IoU threshold

$$\text{mAP} = \frac{1}{\text{C}} \sum \text{AP} \times 100\% \tag{3}$$

- **IoU (Intersection over Union):** Measures spatial overlap between predicted and ground truth bounding boxes

$$\text{IoU} = \frac{|\text{R} \cap \text{R}'|}{|\text{R} \cup \text{R}'|} \tag{4}$$

Each model's predictions were compared to ground truth annotations in the test set, with IoU ≥ 0.5 used as the threshold for a correct detection.
This methodology enables a rigorous and unbiased comparison of YOLOv5, YOLOv8, and YOLOv11 models, highlighting their relative strengths and limitations for underwater object detection. The outcomes are expected to guide the selection of suitable deep learning frameworks in real-time intelligent aquaculture monitoring and marine biodiversity assessment.

## 5. EXPERIMENT AND ANALYSIS

To investigate the feasibility and accuracy of real-time fish detection in underwater environments, three state-of-the-art object detection models—YOLOv5, YOLOv8, and YOLOv11—were comparatively evaluated. These models represent successive stages in the evolution of the YOLO (You Only Look Once) family, with each iteration introducing progressively sophisticated architectures designed to enhance detection speed, localization precision, and generalization across visual domains. The evaluation was conducted on a carefully curated multi-class GlowFish dataset, acquired under controlled aquatic conditions that simulate realistic underwater variability such as lighting gradients, occlusions, and turbidity.

The experimental analysis focuses on four critical performance metrics: Precision, Recall, Mean Average Precision at 0.5 IoU (mAP@0.5), and Mean Average Precision across IoU thresholds from 0.5 to 0.95 (mAP@0.5:0.95). These indicators collectively reflect model robustness in both object classification and spatial localization under noise-prone visual environments.

### 1) Detection Visualization and Qualitative Analysis

YOLOv5 demonstrated robust detection capabilities across all three GlowFish classes—Green, Yellow, and Orange—with confidence scores ranging between 0.60 and 0.88 (Figure 5). The model successfully localized multiple fish instances even under mild occlusion. However, detection certainty for YellowGlowfish was consistently lower, suggesting class imbalance sensitivity and potential limitations in inter-class feature differentiation. Nonetheless, bounding boxes were tightly aligned, indicative of precise spatial localization.


**Figure 5.** YOLOv5 fish detection

YOLOv8 exhibited superior detection consistency and confidence uniformity (Figure 6). Across all categories, confidence scores ranged from 0.78 to 0.90, with YellowGlowfish detection improving markedly compared to YOLOv5. The anchor-free architecture and decoupled head structure likely contributed to its heightened stability in learning complex geometric patterns in cluttered aquatic scenes. The bounding boxes were more refined and consistently centered, indicating improved feature aggregation across scales.

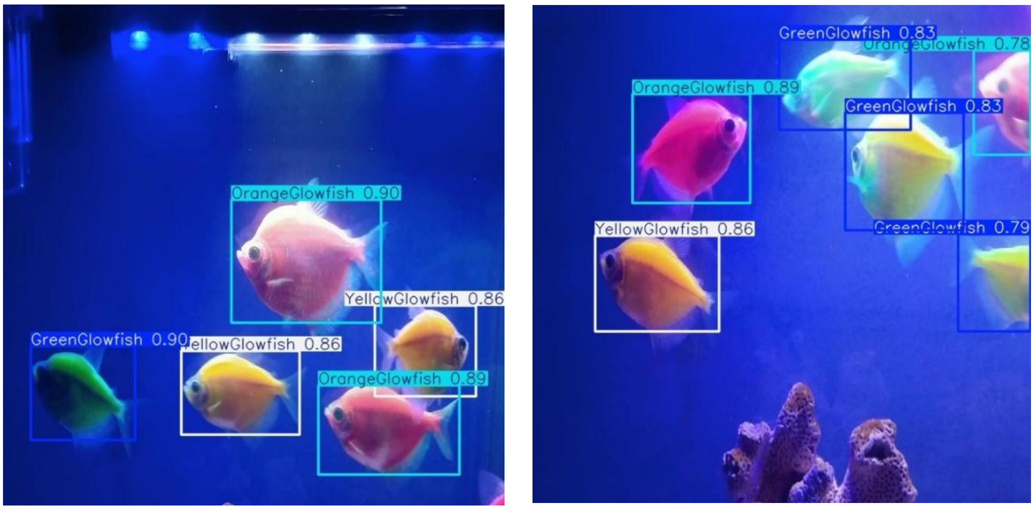



**Figure 6.** YOLOv8 fish detection

YOLOv11, though architecturally advanced, demonstrated a wider variance in confidence scores (0.63–0.82) and slightly more frequent misclassifications, particularly for YellowGlowfish (Figure 7). While it effectively localized all fish classes, its output showed reduced classification certainty and increased bounding box jitter. These patterns may be attributed to the higher architectural complexity (e.g., attention modules and BiFPN), which, when trained on relatively small datasets, can lead to overfitting or sub-optimal convergence in visually ambiguous cases.

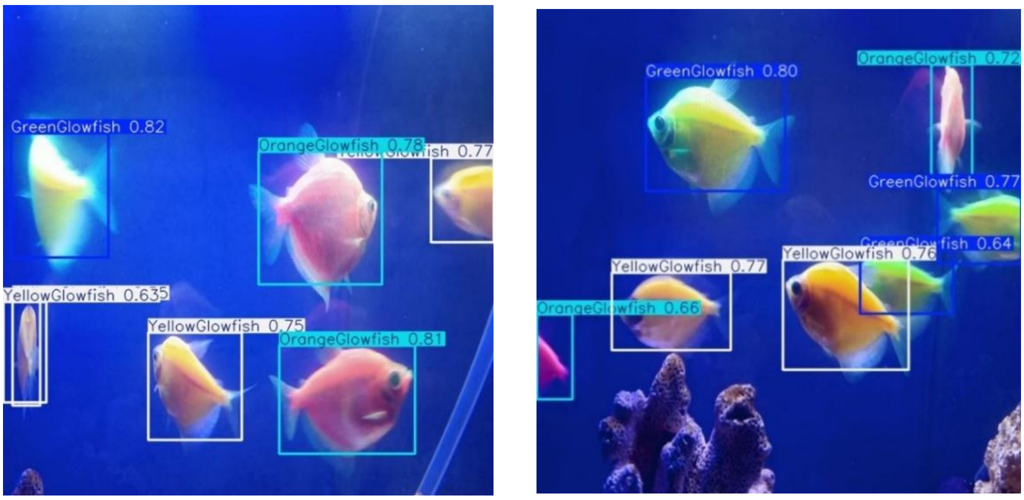



**Figure 7.** YOLOv11 fish detection

### 2) Quantitative Benchmarking and Performance Metrics

Quantitative evaluation results are presented in Table 2 and visualized in Figure 8 for YOLOv8. Each model was trained and tested under identical conditions, ensuring a fair comparison.

**Table 2.** Detection Performance of YOLOv5, YOLOv8, and YOLOv11

| Model | Precision | Recall | mAP50 | mAP50-95 |
|---|---|---|---|---|
| Yolov5 | 0.963 | 0.922 | 0.967 | 0.548 |
| Yolov8 | 0.951 | 0.93 | 0.963 | 0.546 |
| Yolov11 | 0.942 | 0.908 | 0.952 | 0.546 |

- YOLOv5 achieved the highest Precision (0.963) and mAP@0.5 (0.967), confirming its ability to reduce false positives while accurately localizing fish instances.
- YOLOv8 recorded the highest Recall (0.930), indicating a stronger ability to capture more true positives in diverse visual conditions.
- YOLOv11 showed comparatively lower scores across all metrics, though still within an acceptable range, suggesting reliable—but less optimal—performance.

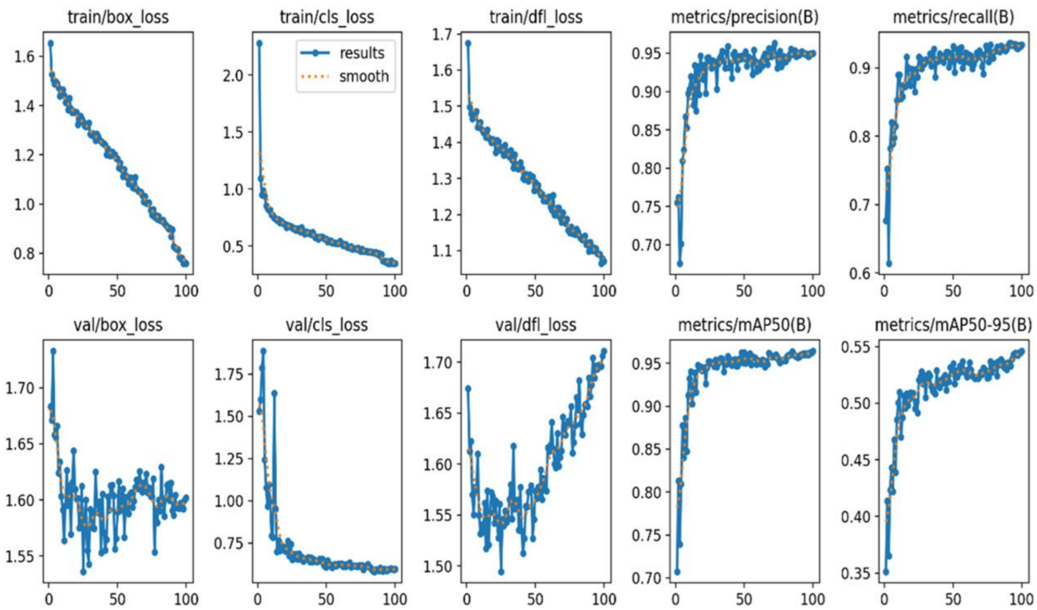



**Figure 8.** Yolov8 performance analysis.

These results suggest that while YOLOv5 remains a reliable baseline for high-confidence detection, YOLOv8 offers improved sensitivity and detection completeness, especially in visually complex scenes.

### 3) Confidence Curve and Precision–Recall Analysis

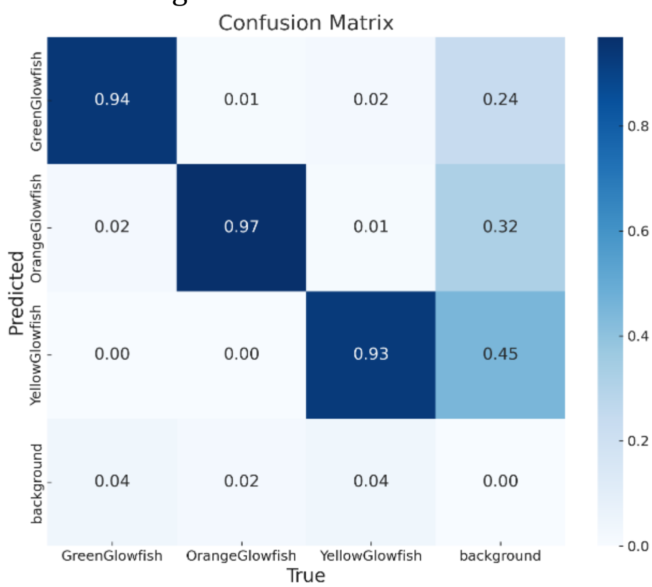
Further performance insights were extracted from a series of detection evaluation curves for YOLOv8 (the best-balanced model):

- F1-Confidence Curve peaked at 0.94 around a threshold of 0.49, indicating that optimal trade-offs between precision and recall occur at moderate confidence levels.
- Precision-Confidence Curve rose toward 1.00 at thresholds above 0.9, confirming high predictive accuracy when strict thresholds are applied.
- Recall-Confidence Curve started high (~0.97) and gradually declined with increasing confidence, illustrating that aggressive filtering may exclude true positives.
- Precision–Recall Curve displayed minimal fluctuation, with mAP@0.5 values exceeding 0.96 across all classes, confirming high classification reliability and bounding box accuracy.

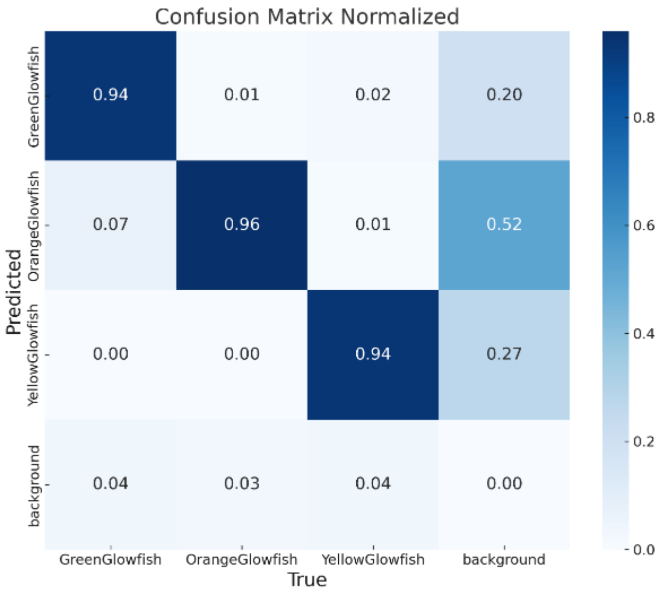
These findings confirm that a confidence threshold between 0.4 and 0.6 is optimal for real-time deployment, balancing prediction certainty with detection coverage.

### 4) Confusion Matrix Interpretation

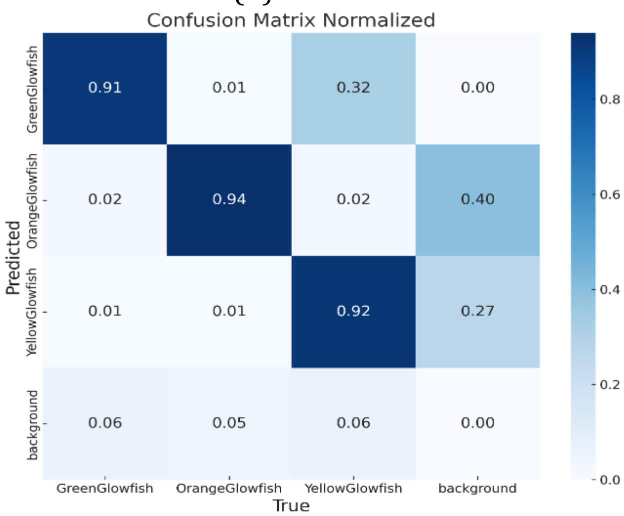
Figure 9 displays confusion matrices for YOLOv5, YOLOv8, and YOLOv11. All models performed well in classifying GreenGlowfish and OrangeGlowfish. However, YOLOv5 misclassified ~45% of YellowGlowfish as background, while YOLOv8 reduced this to ~36%, and YOLOv11 maintained ~41%. This confirms YOLOv8's enhanced capability to distinguish morphologically similar classes, likely due to its improved feature extraction and classification head design.




(a) YOLOv5




(b) YOLOv8




(c) YOLOv11

**Figure 9.**Confusion matrix comparison among YOLO models.

### 5) Model Suitability and Deployment Considerations

The analysis highlights key trade-offs:

- YOLOv5 is optimal for applications where high precision and minimal computational load are required—such as edge AI systems in aquaculture tanks or portable fish monitoring stations.
- YOLOv8 is better suited for operational scenarios that demand comprehensive coverage, higher recall, and stable generalization under visual variability—e.g., integrated aquaponics, biodiversity tracking, or mobile robotics in aquatic systems.
- YOLOv11, while architecturally more sophisticated, may benefit from larger-scale datasets or pretraining with marine-specific domains to fully leverage its attention and BiFPN structures.

### 6) Limitations and Future Work

While the proposed YOLO-based models achieved high detection accuracy in controlled aquatic environments, several limitations constrain their broader applicability. First, the dataset was limited to three fluorescent species in a static tank, lacking the environmental variability—such as turbidity, dynamic lighting, and background noise—typical of open-water or real-world aquaculture systems. Second, mild class imbalance led to reduced performance on underrepresented categories (e.g., YellowGlowfish), suggesting the need for augmentation or rebalancing strategies. Third, the models function as black-box predictors with limited interpretability, which may hinder deployment in high-stakes monitoring applications. Finally, resource efficiency and deployment feasibility were not evaluated on edge-AI hardware, where real-time performance is critical.

Future work should address these limitations by (i) expanding to multi-species, multi-environment datasets, (ii) incorporating temporal information for behavior-aware tracking, (iii) optimizing models for deployment on embedded AI accelerators, and (iv) integrating explainable AI (XAI) tools to enhance transparency and trust. These directions will support the development of scalable, interpretable, and robust fish monitoring systems for intelligent aquaculture and aquatic ecosystem management.

## 6. CONCLUSION

This study conducted a comprehensive benchmarking of three YOLO-based deep learning models—YOLOv5, YOLOv8, and YOLOv11—for real-time fish detection in intelligent aquatic monitoring systems. Using a curated GlowFish dataset collected under controlled but variable underwater conditions, we evaluated each model's performance based on precision, recall, mAP@0.5, and mAP@0.5:0.95. All models achieved high detection accuracy, with YOLOv5 attaining the highest precision (0.963) and mAP@0.5 (0.967), while YOLOv8 achieved the best recall (0.930), indicating superior coverage of true positive detections. YOLOv11 demonstrated reliable localization but exhibited increased output variance and reduced robustness to class imbalance. The proposed evaluation pipeline—including standardized preprocessing, annotation consistency, and rigorous metric analysis—proved effective for revealing the operational strengths and trade-offs among these architectures. YOLOv5 is suited for high-precision, low-resource scenarios, YOLOv8 offers optimal generalizability and balance across metrics, and YOLOv11 shows potential pending further optimization. These results provide critical empirical insights into the selection of object detection models for deployment in smart aquaculture, aquaponics, and ecological monitoring systems. The findings reinforce YOLOv8's suitability for real-time, multi-class fish detection, advancing the development of scalable and intelligent aquatic monitoring platforms

**Acknowledgments**

The authors would like to acknowledge Politeknik Negeri Sriwijaya for funding support through Grant No. 336/PL6.2.1/PG/2025, dated May 22, 2025. The authors also sincerely thank the colleagues and technical personnel in the Electrical Engineering Department for their valuable assistance and support throughout this research.

## REFERENCES

[1] J. Dunning, T. Jansen, A. J. Fenwick, and P. G. Fernandes, **A new in-situ method to estimate fish target strength reveals high variability in broadband measurements**, *Fisheries Research*, vol. 261, p. 106611, 2023. https://doi.org/10.1016/j.fishres.2023.106611

[2] T. M. Evans *et al.*, **Fish avoidance of ships during acoustic surveys tested with quiet uncrewed surface vessels**, *Fisheries Research*, vol. 267, p. 106817, 2023. https://doi.org/10.1016/j.fishres.2023.106817

[3] K. Williams, P. Goddard, R. Wilborn, D. Bryan, and C. Rooper, **Fish behavior in response to an approaching underwater camera**, *Fisheries Research*, vol. 268, p. 106823, 2023. https://doi.org/10.1016/j.fishres.2023.106823

[4] N. Shang, C. Zhang, W. Zhang, X. Zhang, and Y. Zhang, **Assessment of fish health status for waterless transportation based on image features and deep learning models**, *Aquaculture*, vol. 595, Part 2, p. 741697, 2025. https://doi.org/10.1016/j.aquaculture.2024.741697

[5] B. A. Staton, C. Justice, S. White, E. R. Sedell, L. A. Burns, and M. J. Kaylor, **Accounting for uncertainty when estimating drivers of imperfect detection: An integrated approach illustrated with snorkel surveys for riverine fishes**, *Fisheries Research*, vol. 249, p. 106209, 2022. https://doi.org/10.1016/j.fishres.2021.106209

[6] M. Saqib *et al.*, **Fishing event detection and species classification using computer vision and artificial intelligence for electronic monitoring**, *Fisheries Research*, vol. 280, p. 107141, 2024. https://doi.org/10.1016/j.fishres.2024.107141

[7] S. İşgüzar, M. Türkoğlu, T. Ateşşahin, and Ö. Dürrani, **FishAgePredictioNet: A multi-stage fish age prediction framework based on segmentation, deep convolution network, and Gaussian process regression with otolith images**, *Fisheries Research*, vol. 271, p. 106916, 2024. https://doi.org/10.1016/j.fishres.2023.106916

[8] A. A. Muksit, F. Hasan, M. F. H. Bhuiyan Emon, M. R. Haque, A. R. Anwary, and S. Shatabda, **YOLO-Fish: A robust fish detection model to detect fish in realistic underwater environments**, *Ecological Informatics*, vol. 72, p. 101847, 2022. https://doi.org/10.1016/j.ecoinf.2022.101847

[9] W. Bi, X. Sun, J. Li, and Y. Jin, **HD-YOLO: The hydroid of Aurelia detector**, *Marine Pollution Bulletin*, vol. 210, p. 117346, 2025. https://doi.org/10.1016/j.marpolbul.2024.117346

[10] Y. T. Lin, J. Hu, and J. Chen, **Machine vision approach for monitoring and quantifying fish school migration**, *Ecological Indicators*, vol. 169, p. 112769, 2024. https://doi.org/10.1016/j.ecolind.2024.112769

[11] B. Clinton, A. Amperawan, and T. Dewi, **Object detection approach using YOLOv5 for plant species identification**, *Jurnal Elektronika dan Telekomunikasi (JET)*, vol. 24, no. 2, pp. 120–128, 2024. https://doi.org/10.55981/jet.643

[12] H. Zhang, Z. Tian, L. Liu, H. Liang, J. Feng, and L. Zeng, **Real-time detection of dead fish for unmanned aquaculture by YOLOv8-based UAV**, *Aquaculture*, vol. 595, Part 1, p. 741551, 2025. https://doi.org/10.1016/j.aquaculture.2024.741551

[13] Z. Zhou, Y. Hu, X. Yang, and J. Yang, **YOLO-based marine organism detection using two-terminal attention mechanism and difficult-sample resampling**, *Applied Soft Computing*, vol. 153, p. 111291, 2024. https://doi.org/10.1016/j.asoc.2024.111291

[14] C. Zhou, C. Wang, D. Sun, J. Hu, and H. Ye, **An automated lightweight approach for detecting dead fish in a recirculating aquaculture system**, *Aquaculture*, vol. 594, p. 741433, 2025. https://doi.org/10.1016/j.aquaculture.2024.741433

[15] A. Paul, R. Machavaram, Ambuj, D. Kumar, and H. Nagar, **Smart solutions for capsicum harvesting: Unleashing the power of YOLO for detection, segmentation, growth stage classification, counting, and real-time mobile identification**, *Computers and Electronics in Agriculture*, vol. 219, p. 108832, 2024. https://doi.org/10.1016/j.compag.2024.108832

[16] T. Dewi *et al.*, **Fuzzy logic-based control for robot-guided strawberry harvesting: Visual servoing and image segmentation approach**, *SINERGI*, vol. 28, no. 3, pp. 661–668, 2024. https://doi.org/10.22441/sinergi.2024.3.021

[17] Y. Shibata *et al.*, **Length estimation of fish detected as non-occluded using a smartphone application and deep learning method**, *Fisheries Research*, vol. 273, p. 106970, 2024. https://doi.org/10.1016/j.fishres.2024.106970

[18] D. Acharya *et al.*, **Towards automatic anomaly detection in fisheries using electronic monitoring and automatic identification system**, *Fisheries Research*, vol. 272, p. 106939, 2024. https://doi.org/10.1016/j.fishres.2024.106939

[19] B. C. Weinbach, R. Akerkar, M. Nilsen, and R. Arghandeh, **Hierarchical deep learning framework for automated marine vegetation and fauna analysis using ROV video data**, *Ecological Informatics*, vol. 85, p. 102966, 2025. https://doi.org/10.1016/j.ecoinf.2024.102966

[20] A. Yassir, S. J. Andaloussi, Q. Ouchetto, K. Mamza, and M. Serghini, **Acoustic fish species identification using deep learning and machine learning algorithms: A systematic review**, *Fisheries*

*Research*, vol. 266, p. 106790, 2023. https://doi.org/10.1016/j.fishres.2023.106790

[21] S. G. Alarcón-Silvas *et al.*, **Water quality, water usage, nutrient use efficiency, and growth of shrimp Litopenaeus vannamei in an integrated aquaponic system with basil Ocimum basilicum**, *Aquaculture*, vol. 543, p. 737023, 2021. https://doi.org/10.1016/j.aquaculture.2021.737023

[22] G. P. Suárez-Cáceres, L. Pérez-Urrestarazu, M. Avilés, C. Borrero, J. R. L. Eguíbar, and V. M. J. A. Fernández-Cabanás, **Susceptibility to water-borne plant diseases of hydroponic vs. aquaponic systems**, *Aquaculture*, vol. 544, p. 737093, 2021. https://doi.org/10.1016/j.aquaculture.2021.737093

[23] T. Dewi, E. N. Mardiyati, P. Risma, and Y. Oktarina, **Hybrid machine learning models for PV output prediction: Harnessing Random Forest and LSTM-RNN for sustainable energy management in aquaponic system**, *Energy Conversion and Management*, vol. 330, p. 119663, 2025. https://doi.org/10.1016/j.enconman.2024.119663

[24] T. Dewi *et al.*, **Smart integrated aquaponics system: Hybrid solar-hydro energy with deep learning forecasting for optimized energy management in aquaculture and hydroponics**, *Energy for Sustainable Development*, vol. 85, p. 101683, 2025. https://doi.org/10.1016/j.esd.2024.101683

[25] Y. Oktarina, Z. Nawawi, B. Y. Suprapto, and T. Dewi, **Towards ecological sustainability: Harvest prediction in agrivoltaic chili farming with CNN transfer learning**, *Iraqi Journal of Agricultural Sciences*, vol. 55, no. 6, pp. 1910–1926, 2024. https://doi.org/10.36103/5gdhkh84

[26] Y.-T. Lai, Y.-T. Peng, W.-C. Lien, Y.-C. Cheng, Y.-T. Lin, C.-J. Liao, and Y.-S. Chiu, **Fully automated learning and predict price of aquatic products in Taiwan wholesale markets using multiple machine learning and deep learning methods**, *Aquaculture*, vol. 586, p. 740741, 2024. https://doi.org/10.1016/j.aquaculture.2024.740741

[27] C. A. Graham *et al.*, **A practical evaluation of machine learning for classification of ultrasound images of ovarian development in channel catfish (Ictalurus punctatus)**, *Aquaculture*, vol. 552, p. 738039, 2022. https://doi.org/10.1016/j.aquaculture.2022.738039

[28] L. Ortenzi *et al.*, **Automated species classification and counting by deep-sea mobile crawler platforms using YOLO**, *Ecological Informatics*, vol. 82, p. 102788, 2024. https://doi.org/10.1016/j.ecoinf.2024.102788

[29] H. Yang, K. Lang, and X. Wang, **Identify and segment microalgae in complex backgrounds with improved YOLO**, *Algal Research*, vol. 82, p. 103651, 2024. https://doi.org/10.1016/j.algal.2024.103651