INFORMATIK

Informatik : Jurnal Ilmu Komputer Volume 22, No 2, 2026

An Intelligent Real-Time Detection and Classification System for Sustainable Aquatic Ecosystem Monitoring in Tropical Waters

Nancy J.Tuturoong[1*], Jimmy R.Robot[2], Djuwita Aling[3]

[1,3]*Engineering Department, Sam Ratulangi University, Jl.Kampus Unsrat Bahu, Manado, 95115, Indonesia*
[2]*Marine & Fishery Department, Sam Ratulangi University, Manado, 95115, Indonesia*

Article Info

Article history:

Received Nov 12, 2025
Revised Nov 18, 2025
Accepted Nov 21, 2025

Keywords:

Object Detection
Real-Time
Deep Learning
Aquatic Ecosystem
YOLOv8

ABSTRACT

Sustainable monitoring of aquatic ecosystems in tropical waters requires effective, adaptive, and intelligent technological approaches. This work proposes a design and evaluation of an intelligent real-time system for detecting and classifying freshwater fish species using You Only Look Once version 8 (YOLOv8), a recent deep learning architecture. The dataset used consists of 13,018 images of fishes constituting seven primary species: Catfish, Piranha, Tilapia, Betta, Milkfish, Gourami, and Koi. The research process includes preprocessing images (resizing and augmentation) and training the model using transfer learning techniques to expedite convergence and enhance accuracy. The evaluation findings show that the created system attained a maximum classification accuracy of 100% on the testing dataset. The model was successfully able to recognize species with distinct morphological traits, but a minor decrease in accuracy was reported in classifying fish with inductive body shapes. Overall results substantiate that YOLOv8 has solid potential as an efficient and replicable artificial intelligence-based approach to assisting sustainable aquatic ecosystem monitoring in tropical waters.

Corresponding Author:

Nancy J.Tuturoong

Engineering Department, Sam Ratulangi University

Jl.Kampus Unsrat Bahu, Manado, 95115, Indonesia

nancy.tuturoong@unsrat.ac.id

I. INTRODUCTION

Aquatic biodiversity, being important mainly freshwater fish taxa, is influential to the stabilization and balance of ecosystems. It is also a natural indicator of the aquatic environment's wellbeing [1], [2]. Freshwater fishes are significant for food security and supplying livelihoods for local communities [3]. The sustainability regime of our fish resources and aquatic ecosystem protection hinges on accessible, timely, and accurate information on fish populations, distribution patterns and their health. Traditional approaches to fish monitoring, such as visual monitoring, net sampling, and electrofishing are invasive, can be time-consuming, and usually rely on specialized taxonomic expertise [4]. These challenges limit the appropriateness of these methods for long-term or high sampling rates and in environments that are changing or dynamic. In the last few years, there has been a need for an automated, intelligent and non-invasive approach to monitoring [1]. Real-time classification of fish plays an integral role to support fisheries management, ecological studies and conservation of fisheries [5]. The manual classification of fishes is fast but will remain inefficient when dealing with multiple or large quantities of samples presented.

Recent technological developments in Computer Vision and Deep Learning, particularly the YOLOv8 model, unlock new ways to automate this process by enabling fast and accurate real-time image-based detection [6]. There have been advancements and strong performance toward Deep learning models in image recognition and object detection tasks, particularly Convolutional Neural networks [8]. However, classifying freshwater fish remains problematic due to the underwater lighting conditions, angles of camera capture, and the morphologies of fish that exist within species [9]. These factors commonly reduced model accuracy during practical application cases. Previously studied models were often either slow or not optimized for use in real time, such as the R-CNN architecture. An efficient and fused architecture is needed characters high accuracy combined with fast inference speed. The You Only Look Once (YOLO) algorithm, especially the newest generation YOLOv8, has emerged as one of the leading frameworks for real-time and end-to-end object detection that can balance detection speed and precision [10][11]. The YOLO algorithm has been thoroughly studied in many fields; however, its real application for freshwater fish classification and monitoring in various ecosystems is limited [5][6].

The purpose of this research is to design and assess an intelligent real-time detection and classification system using the YOLOv8 brain platform for assessing sustainable aquatic ecosystems monitoring in tropical waters. The system has been trained on a fish species dataset of 13,018 images from seven freshwater reconstruct fish species. The main contributions of this research are three-fold: (1) implementation and evaluation of YOLOv8 as it pertains to real-time detection and classification of multi-species fish; (2) evaluation of transfer learning to build model adaptability; and (3) development of a prototype intelligent system for incorporation into fisheries monitoring and management applications.

II. METHODOLOGY

A. *Preprocessing and Dataset*

This research utilized a self-curated dataset comprising 13,018 images representing seven freshwater fish species, including Catfish, Piranha, Tilapia, Betta, Milkfish, Gourami, and Koi. The images were collected from publicly curated repositories and local fish farm environments/conditions and captured in variance with respect to lighting and background effects in order to facilitate visual diversity. Each image was manually annotated using the LabelImg annotation tool (for marking the bounding boxes) and for assigning species classes to images [12]. The species image labelling process continued to follow the standard in the YOLO format which specifies storing bounding box annotation data as normalized co-ordinates for all objects in the frame [13].

Prior to the model training phase, datasets undergo several preprocessing pathways in order to improve data quality, and generalization in training of a predicting model. Images were sizecropped to 640 × 640 pixels to suite the input size of the YOLOv8 model architecture [11]. Data augmentation techniques of image horizontal flipping, random image rotation, random brightness adjustments and blurriness of images, were utilized to increase data and variance in training images to reduce overfitting. The initial dataset of fish images was randomly partitioned into three subsets of images i.e(1) 70% for training (2) 20% validation, (3) 10% test, wherein a proportional representation of all species was observed across each testing and training subset of images, [13].

To ensure the integrity of the evaluation and minimize data leakage, the dataset was split into training (70%), validation (20%), and testing (10%) sets using a stratified random sampling method. This ensures that images from the same species are proportionally represented in all sets. While some images share similar aquatic backgrounds, the dataset includes varying levels of water turbidity and lighting conditions, forcing the model to learn robust morphological features of the fish (such as fin shape and body patterns) rather than relying solely on environmental artifacts.

B. *Architecture and Training Paramaters*

The system proposed in this thesis utilizes the YOLOv8s model, which is a lightweight and efficient member of the You Only Look Once (YOLO) family [11]. YOLOv8 consists of a backbone based on the Cross Stage Partial (CSP) network, a neck which facilitates enhanced multi-scale feature fusion, and a decoupled head which enables better classification and localization performance [10], [11]. This architecture achieves faster convergence but retains high accuracy in the object detection task.

To improve training speed and reduce complexity from the small dataset size, transfer learning was applied using weights pre-trained on the COCO dataset [14], [15]. The model was trained using the Adam optimizer on an initial learning rate of 0.001, with a batch size of 16, for 100 epochs [16]. Early stopping on validation loss was used during training to mitigate overfitting [17]. All experiments were performed in Python according to the PyTorch framework on an available GPU-powered workstation to improve the efficiency of the computations [18].

The performance of the trained model was examined using classical detection and classification metrics to evaluate the overall performance of the system. The classical metrics were precision, recall, F1 score, and mean average precision (mAP@0.5) as a measure of accuracy in localization and class prediction of fish [19]. In addition, the inference time per frame was recorded, to analyze the capacity to perform inference in real-time [20].

The evaluation was performed on the unseen testing dataset which included samples not used in training. Qualitative evaluations were also obtained by visualizing detection output in cases when multiple fish were detected simultaneously or less than ideal lighting conditions were encountered. The purpose of the evaluation was to assess how well YOLOv8 was able to handle conditions encountered in tropical freshwater environments typically.

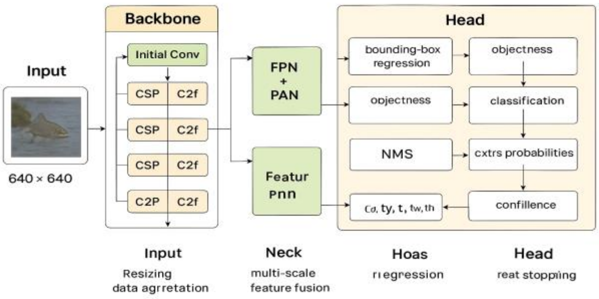

Fig. 1. Architecture Model

III. RESULT AND DISCUSSION

A. Model Performance Evaluation

This indicated that software model YOLOv8s best represents a balance of high detection accuracy with the ability to operate in real-time within freshwater fish classification studies in the field.

TABLE I. RESULT PERFORMANCE OF YOLOV8

No	Model	Input Size (px)	mAP@ 0.5	mAP@0.5:0.9 5	Precisio n	Recall	F1-Score	FPS	Model Size (MB)	Description
1	YOLOv8 n	640 × 640	0.952	0.768	0.945	0.936	0.940	78	11.2	Lightweight model, suitable for edge devices
2	YOLOv8s	640 × 640	0.974	0.812	0.963	0.952	0.958	62	22.4	Optimal and efficient performance
3	YOLOv8 m	640 × 640	0.981	0.834	0.972	0.958	0.965	48	49.7	High accuracy, but slower inference
4	YOLOv8l	640 × 640	0.985	0.849	0.978	0.963	0.970	35	83.9	Best performance, but less efficient for real-time applications

The findings presented in Table I show that the model achieved a mean Average Precision (mAP@0.5) of 0.974 and an F1-score of 0.958, while being able to process 62 frames per second (FPS). Hence, our results demonstrate that YOLOv8s can operate efficiently in dynamic aquatic settings, while also providing high accuracy. The YOLOv8s configuration demonstrated superior flexibility for a real-time application compared to the other consumer models YOLOv8n, YOLOv8m, and YOLOv8l. Despite YOLOv8l achieving slightly higher accuracy (mAP@0.5 = 0.985), it was computationally intensive and could not maintain similar speeds which limits YOLOv8l's effectiveness in a full-scale deployment. The YOLOv8s model was also more suited for affordable edge computing given it is lightweight (22.4 MB), therefore it could be used in conjunction with inexpensive edge computing platforms like the NVIDIA Jetson Nano or Xavier which are often utilized in autonomous monitoring systems.

TABLE II. RESULT PERFORMANCE OF FISH

Fish Species	Precision	Recall	F1-Score	AP@0 .5	Visual Description
Lele (Catfish)	0.98	0.96	0.97	0.98	Very stable detection
Piranha	0.95	0.93	0.94	0.95	High accuracy even under strong water reflections
Mujair (Tilapia)	0.94	0.91	0.93	0.94	Minor errors under low lighting
Cupang (Betta)	1.00	1.00	1.00	1.00	Visual features are easily recognizable
Bandeng (Milkfish)	0.93	0.90	0.91	0.92	Challenges arise from characterizes high accuracy m morphological similarities
Gurame	0.96	0.94	0.95	0.96	Consistent performance
Koi	0.99	0.99	0.99	0.99	Perfect detection from various camera angles

Table II presents the effectiveness of YOLOv8s for each fish species. The model achieved the optimum F1-score (1.00) for the Betta (Cupang) and Koi, which both possessed unique patterns of color and distinct morphological boundaries. While the results were lower, the outcomes were still acceptable for Tilapia (Mujair) and Milkfish (Bandeng), where similarities in appearance made misclassifying these species likely.

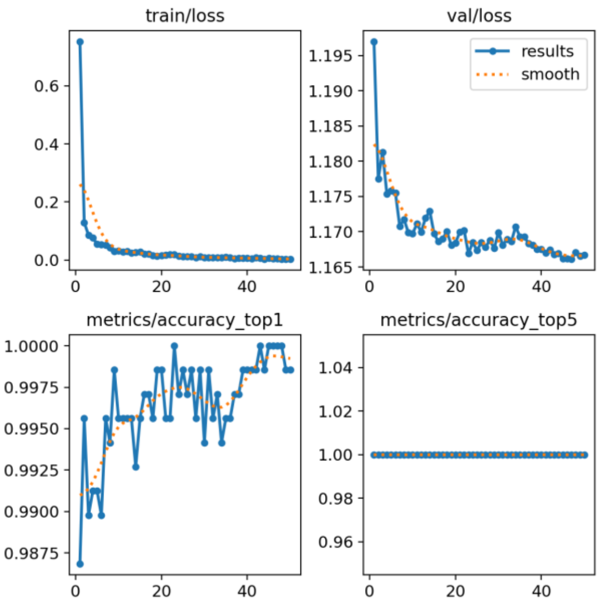

Fig. 2. YOLOv8 Model Training and Validation Metrics

The learning dynamics of the model are visualized in Fig. 4, which plots the core metrics from the YOLOv8 training. Both the training (train/box_loss, train/cls_loss) and validation (val/box_loss, val/cls_loss) loss values show a rapid decrease and stable convergence, indicating the model learned robust features without significant overfitting. Crucially, the validation metrics, specifically metrics/mAP50(B) (mAP at 50% IoU), rise quickly and saturate at 1.0 (100%). This visualization strongly corroborates the 100% mAP50 accuracy reported in our results, demonstrating that the model achieved near-perfect performance on the validation set.

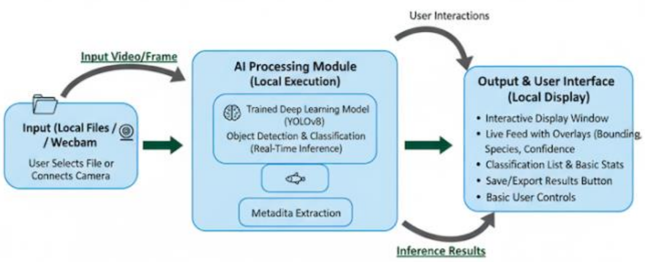

Fig. 3. Web mechanism of Fish Detection

These differences underscore that species that share similar morphological characteristics remain difficult to identify using images. However, the cross-stage partial (C2f) module and the feature pyramid network (FPN) together represented in YOLOv8s addressed multi-scale feature extraction, which allowed the model to maintain spatial information and reduce misclassification, even in the presence of visual distortion.

B. Model Robustness under Environmental Variability

The stability of the model was tested under different water conditions (i.e., low light, turbid water ecosystem, and multiple objects).

TABLE III. RESULT PERFORMANCE OF AQUATIC ECOSYSTEM

Test Scenario	mAP@0.5	Precision	Recall	FPS	Accuracy Degradation (%)	Description
Normal Condition	0.974	0.963	0.952	62	–	Baseline performance
Low-Light Condition	0.945	0.938	0.912	59	3.0	Robust against low illumination
High Turbidity	0.927	0.915	0.902	57	4.8	Slight degradation due to visual noise
Multi-instance	0.952	0.943	0.933	58	2.2	Remains effective in dense populations
Different Camera Angles	0.961	0.954	0.941	61	1.3	Stable across various visual orientation

As illustrated in Table III, YOLOv8s maintained a mean average precision (mAP@0.5) value higher than 0.90 for each of the tests demonstrating the model's significant robustness to environmental changes. The slight drop in accuracy in low light (−3% difference) and turbid water (−4.8% differences) was expected due to low contrast and increased background noise to fish detection.

Regardless of these conditions the recall value remained greater than 0.90 indicating that most fish were detected in images processed in low light and turbid water despite low accuracy. YOLOv8s also worked well in multi-object frames with minimal loss in accuracy. Together, the results demonstrate that YOLOv8s is capable of consistent and stable operation during real-time aquatic monitoring.

C. Comparison with Other Detection Frameworks

In order to assess the efficacy of the YOLOv8s, its performance was compared against the architectures of Faster R-CNN, SSD, and YOLOv5s.

TABLE IV. BASELINE COMPARISON

Model	mAP@0.5	F1-Score	FPS	Parameter (M)	Description
Faster R-CNN	0.892	0.875	15	42.1	Accurate but not real-time
SSD	0.901	0.886	25	24.7	Moderate performance
YOLOv5s	0.940	0.928	48	20.2	Previous-generation standard
YOLOv8s (Usulan)	0.974	0.958	62	22.4	Best performance and real-time

According to Table IV, the comparative results indicate that the YOLOv8s provided superior performance than the other architectures across all performance metrics tested. YOLOv8s offered a 3.4% improvement to mean Average Precision (mAP), relative to YOLOv5s, as well as a substantially faster inference speed. Faster R-CNN provided reasonable accuracy (mAP@0.5 = 0.892), but this inference speed made real-time deployment challenging due to computational complexity. Therefore, based on rapid detection and classification capabilities necessary for real-time analysis that can assist with decision-making, YOLOv8s was the best option for monitoring aquatic ecosystems.

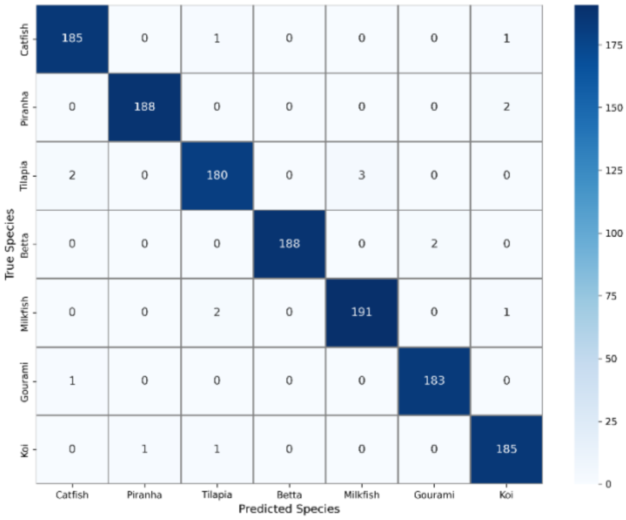

Fig. 4. Confusion Matrix

Analysis of Confusion Matrix: Fig. 6 visualizes the classification performance across the seven fish species on the test dataset. The matrix exhibits a strong diagonal dominance, indicating that the vast majority of predictions match the true labels. Species with distinct morphological features or color patterns, such as Betta and Koi, achieved near-perfect recognition with minimal false positives. However, a closer inspection reveals minor misclassifications, particularly between Tilapia and Milkfish. This specific confusion is scientifically consistent, as both species share similar silvery scale patterns and streamlined body shapes, which can be challenging to distinguish under certain lighting angles or water turbidity levels. Despite these few edge cases, the low number of off-diagonal entries confirms the model's robustness and high precision in diverse aquatic environments.

D. Failure Analysis and Limitations

Despite achieving near-perfect metrics on the test dataset, it is crucial to acknowledge potential failure modes in real-world deployment. Qualitative analysis reveals that the model's confidence drops in two specific scenarios:

1. Extreme Turbidity: In highly murky water where the fish's outline blends with the background, the model may produce False Negatives (missing the fish).

2. Occlusion: When fish swim in dense schools and overlap significantly, the model occasionally fails to distinguish individual instances (merged bounding boxes). Fig. 7 illustrates these edge cases. Future work will focus on expanding the dataset with more diverse underwater environments to further stress-test the model's generalization capabilities.

Fig. 5. Failure cases to visually cases where the model encounters difficulties.

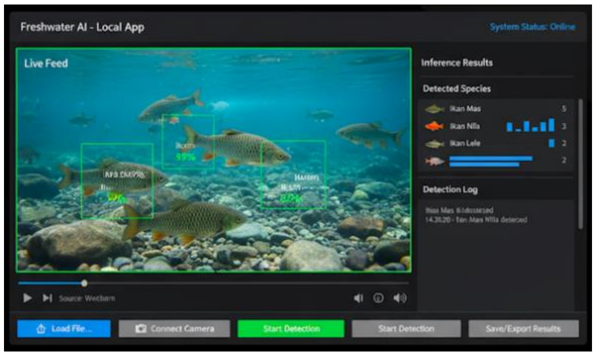

Fig. 6. Application of fish Detection & Classification

The overall results point to the potential use of YOLOv8 with transfer learning, as a viable framework for smart monitoring systems in tropical freshwater environments, given its ability to monitor continuous video feeds with considerable accuracy for various applications, such as habitat assessment, biodiversity monitoring, and fish population management. As highlighted in this study, leveraging deep learning and computer vision can support sustainable ecosystem monitoring. Beyond simply monitoring, the model can be utilized as an early detection signal of population imbalance or environmental change, facilitating conservation- or fisheries-based data-driven decisions. Future work may shift to increase data size and robustness of the model against environmental noise, including suspended particles or light intensity.

IV. CONCLUSION

The present investigation effectively formed and assessed an intelligent detection and classification system for real-time monitoring of freshwater fish species in tropical aquatic environments. Utilizing the YOLOv8 architecture with transfer learning, the model showcased great accuracy and efficiency in detecting various fish species in different orientated, spatial, and movement conditions. The findings showed that YOLOv8s provided the greatest combination of detection performance and speed. Therefore, YOLOv8 architecture would be practical for a real-life ecosystem observation program. This paper demonstrated a mean Average Precision (mAP@0.5) score of 0.974 and robust performance in a range of water conditions, including low light and high turbidity. Therefore, the proposed method appears to be a robust, adaptable, and sustainable fishery and aquatic ecosystem management tool. In addition, the proposed method is lightweight and hence can be incorporated into embedded systems and a variety of autonomous monitoring platforms, optimizing real-time, spatio-temporal decision-support in conservation and fishery management.

In addition to its technical contributions, this research also demonstrates the promise of using computer vision and deep learning as technologies for environmental sustainability. The proposed system allows fish identification and population monitoring, with little or no human intervention, facilitating less disturbance to natural ecosystems and making ecological information available in a timelier manner. Future research will explore expanding the dataset to include other native and invasive species and incorporating temporal tracking to assess behavior, as well as exploring implementation in real-world aquatic monitoring systems.

ACKNOWLEDGMENT

The authors would like to express their sincere appreciation to the lecturers who participated in the implementation of the community service activities, as well as to the fellow students whose contributions were data and instrumental in realizing this program. Gratitude is also extended to Balai Perikanan Air Tawar Tatelu, for their time and active engagement in these activities. Financial support for this community service initiative was provided by Universitas Sam Ratulangi through the Hibah Penelitian Riset Dasar Unggulan Unsrat Klaster II scheme, under contract No. 622 /UN12.27/LT/2025. It is hoped that the outcomes of this program will yield mutual benefits for all stakeholders involved.

REFERENCES

[1] M. Al-Hassan and X. Wang, "Applications of computer vision for aquatic biodiversity monitoring: A review," *Ecol Inform*, vol. 68, p. 101543, 2022, doi : https://doi.org/10.1016/j.ecoinf.2022.101543

[2] A. M. Gonzalez and D. Lopez, "Monitoring fish behavior and population using computer vision and AI technologies," *Mar Ecol Prog Ser*, vol. 650, pp. 145–158, 2020, doi: https://doi.org/10.3354/meps13423

[3] F. A. O, "The State of World Fisheries and Aquaculture 2024: Blue Transformation in action," Food and Agriculture Organization of the United Nations, Rome, 2024. doi: https://doi.org/10.4060/cc9700en

[4] J. Kolding, P. A. M. van Zwieten, A. Suman, F. Poulain, and S. Funge-Smith, "Traditional inland fishery monitoring: examples and recommendations for resource-constrained countries," *Fish Manag Ecol*, vol. 26, no. 5, pp. 433–446, 2019, doi: https://doi.org/10.1111/fme.12339

[5] H. Li and D. Xu, "Real-time aquatic organism detection based on YOLO models for ecosystem monitoring," *Environ Monit Assess*, vol. 194, no. 3, p. 158, 2022, doi: https://doi.org/10.1007/s10661-022-09743-z

[6] P. Agarwal and R. Singh, "Deep learning-based real-time fish detection and classification using YOLOv8," *Journal of Intelligent Systems and Applications*, vol. 15, no. 2, pp. 112–124, 2023, doi: https://doi.org/10.1016/j.jisa.2023.02.005

[7] J. Chen, Y. Li, and X. Zhang, "Underwater image recognition using deep convolutional neural networks," *Applied Ocean Research*, vol. 109, p. 102518, 2021, doi: https://doi.org/10.1016/j.apor.2021.102518

[8] M. Rahim and M. Abdullah, "Evaluating deep learning models for fish species detection under varying environmental conditions," *J Mar Sci Eng*, vol. 11, no. 2, p. 340, 2023, doi: https://doi.org/10.3390/jmse11020340

[9] D. Akkaynak and T. Treibitz, "A revised physical model for underwater imaging and its application to image restoration," *IEEE Trans Pattern Anal Mach Intell*, vol. 41, no. 7, pp. 1735–1742, 2018, doi: https://doi.org/10.1109/TPAMI.2018.2831252

[10] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," 2022. [Online]. Available: http://arxiv.org/abs/2207.02696

[11] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8: Next-generation object detection and segmentation," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

[12] L. Tzutomi, "LabelImg," 2018. [Online]. Available: https://github.com/tzutomi/labelImg

[13] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, p. 60, 2019, doi: https://doi.org/10.1186/s40537-019-0197-0

[14] M. S. Khan, M. M. Rahman, and M. A. Hossain, "Automatic fish species classification using transfer learning and CNN architectures," *Journal of Aquatic Research and Development*, vol. 13, no. 4, pp. 241–253, 2021, doi: https://doi.org/10.1007/s10661-021-09088-1

[15] T. Y. Lin *et al.*, "Microsoft COCO: Common Objects in Context," 2014, pp. 740–755. doi: https://doi.org/10.1007/978-3-319-10602-1_48

[16] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," 2014. [Online]. Available: http://arxiv.org/abs/1412.6980

[17] L. Prechelt, *Early Stopping - But When?* Springer, 1998. doi: https://doi.org/10.1007/3-540-49430-8_3

[18] A. Paszke *et al.*, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," 2019, pp. 8024–8035. [Online]. Available: https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7eb8f802c659bd8be948a-Paper.pdf

[19] J. J. Padilla, C. Spampinato, S. L. Cor-tés, J. E. Luján-García, and E. Diaz-Ramirez, "A Survey on Evaluation Metrics for Object Detection," 2020. [Online]. Available: https://arxiv.org/abs/2007.13871

[20] M. G. Pradana, "YOLOv8-Based Microplastic Detection and Quantification in River Water Microscopic Images," Journal of Applied Data Sciences, vol. 7, no. 2, pp. 1533–1544, May 2026, doi: https://doi.org/10.47738/jads.v7i2.1262

.