Article

YOLOv8-WBF: Ensemble Learning for Reliable Detection of Endangered Medaka (Oryzias)

Rahmatullah R. [1,], Armin Lawi [1,2,3], Muhammad Haerul [1], Iman Mustika Ismail [1], Irma Andriani [4], Andi Iqbal Burhanuddin [5], and Mario Köppen [6]

[1] Information Systems Study Program, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia
[2] Data Science and Artificial Intelligence Research Group, Hasanuddin University, Indonesia
[3] B.J. Habibie Institute of Technology, Parepare, Indonesia
[4] Department of Biology, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia
[5] Department of Fishery, Faculty of Fishery and Marine Sciences, Hasanuddin University, Indonesia
[6] Department of Creative Informatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan
* Correspondence : armin@unhas.ac.id)

Abstract: Reliable detection of Medaka (Oryzias) fish is essential for ecological monitoring and conservation, particularly for tracking population trends of endangered species. This study evaluates the performance of a state-of-the-art deep learning model (YOLOv8) and an ensemble approach using Weighted Box Fusion (WBF) on a manually annotated dataset of Medaka images collected from online sources. Models were trained and validated using 5-fold cross-validation, and performance was assessed using COCO metrics, including mean Average Precision (mAP), precision, recall, and bounding box regression error. The YOLOv8-WBF ensemble achieved a mAP@0.5:0.95 of 0.578, representing an 8% improvement over the best single model. It also enhanced bounding box localization and classification reliability, particularly for small and visually challenging fish instances. These accuracy gains came at the expense of computational efficiency, with inference requiring approximately five times more operations than a single YOLOv8 model. While less suited for real-time deployment, the ensemble approach offers more reliable detection for offline ecological workflows, where accuracy is prioritized over speed. By reducing missed detections of rare or occluded fish, this work contributes to more robust biodiversity monitoring and provides a baseline for developing optimized ensemble and lightweight detection models in aquatic conservation.

Keywords: Medaka (Oryzias); Deep Learning; Object Detection; YOLOv8; Weighted Box Fusion; Ensemble Learning; Ecological Monitoring; Biodiversity Conservation

Received:
Revised:
Accepted:
Published:

Citation: . Title. *Journal Not Specified* **2025**, *1*, 0. https://doi.org/

1. Introduction

Medaka fish (Oryzias) are small freshwater fish valued as ornamental species and are significant for biodiversity studies. They are species that are under danger of extinction as declared by the International Union for Conservation of Nature (IUCN), which makes protecting them an ecological priority. Identifying Medaka apart is challenging since they have subtle morphological differences and small sizes. Not only that, but the variability of aquatic environments also makes it even more challenging. Traditionally, medaka fish have been sacrificed in taxonomy studies due to their genetic value for lineage studies [1]. Their body colour serves as a social signal and reflects environmental conditions [2], making them potential indicators of ecosystem health. Research conducted using traditional methods, such as direct capture and then being put in an experimental tank to understand its anatomy and internal structure, can harm the organism and limit its application in long-term ecological and genetic studies. Reliable detection technologies can facilitate non-invasive conservation without the need for traditional conservation methods. This technique protects species and ecosystems while enabling effective ecological monitoring in conservation efforts.

Deep learning has evolved beyond digit identification to sophisticated object detection, facilitating applications in autonomous vehicles, medical diagnostics, agricultural automation, and environmental monitoring [3,4].In ecological research, object detection has been utilized for monitoring insects in agriculture [5,6] and detecting wildlife in natural habitats [7–9], where precise identification is frequently challenging due to visual similarities among species and the complexity of water environments. These issues are also present in the medaka fish species, such as Oryzias javanicus and Oryzias celebensis, which have subtle morphological distinctions that complicate the reliable identification of endemic fish.

The YOLO architecture is one of the most widely used deep learning-based one-stage object detectors. Among its versions, YOLOv8 has shown a particularly good balance between speed and accuracy, as demonstrated in comparative benchmarks that evaluated YOLOv8 through YOLOv11 under real-world conditions [10]. Researchers have also proposed many variants to push YOLO's performance further, especially in challenging settings. For example, CEH-YOLO adds a high-order deformable attention (HDA) module to better highlight important spatial features, an Enhanced Spatial Pyramid Pooling-Fast (ESPPF) module for improved texture and color feature extraction, and a Composite Detection module to boost detection of small or overlapping underwater objects, along with using WIoU-v3 loss to improve bounding box regression under hard conditions [11]. [12] introduces the Softplus activation function to improve training stability, an AIFI module to strengthen intra-scale feature interactions (reducing false positives and missed detections), and lightweight neck convolution modules (GSConv, VoV-GSCSP) to reduce computational overhead while maintaining accuracy. SCoralDet focuses on underwater soft coral detection, using a Multi-Path Fusion Block (MPFB) to handle varied scales and lighting colors, lightweight modules for efficiency, and an Adaptive Power Transformation label assignment strategy to better align anchors with ground truth when coral structures are complex or blurred [13]. While these studies mainly alter the internal YOLO architecture to address trade-offs between accuracy and speed, in our work, we propose the use of an ensemble method that builds on YOLOv8 without modifying its core architecture by combining the strengths of multiple detection heads or models to improve reliability under environmental variability.

Despite its strong performance, YOLO models could still make misclassifications by detecting background regions as objects or producing duplicate overlapping predictions of the same object. By default, these redundant outputs are reduced through Non-Maximum Suppression (NMS), yet this technique has well-known limitations, particularly when objects overlap or when multiple plausible predictions exist. Recent studies have come up with a new ensemble method, Weighted Boxes Fusion (WBF), that provides a more accurate and robust alternative, since it merges bounding boxes based on confidence scores and spatial alignment rather than discarding valuable predictions outright [14]. The limitations of standard non-maximum suppression (NMS) underscore the need to investigate advanced fusion strategies in object detection. Ensemble methods are particularly effective because they combine multiple models or detection heads, thereby improving uncertainty management, reducing false detections, and increasing consistency in challenging scenarios. The primary contribution of this work is the development of an ensemble approach for YOLOv8 that incorporates Weighted Box Fusion (WBF) to improve detection reliability in ecological monitoring applications, where accuracy is essential and misclassification significantly affect conservation outcomes.

In this work, we propose a novel Medaka fish dataset with images captured under diverse lighting colors to provide a realistic and challenging benchmark for ecological monitoring. Building on this resource MEDAKA-ẽL, an ensemble-based detection framework for accurate and reliable identification of Medaka fish. By training 5 different YOLOv8n models trained across a 5-fold cross-validation of the dataset, which causes the model to learn its own unique feature of the Medaka fish. At the ensemble stage, predictions from multiple YOLOv8 models trained through cross-validation are combined using Weighted Boxes Fusion (WBF). It addresses challenges such as background misclassifications and redundant overlapping bounding boxes. Unlike traditional Non-Maximum Suppression (NMS), WBF merges bounding boxes based on confidence scores and spatial alignment, preserving valuable detections and reducing false positives that achieve higher accuracy. The key contributions of this work are as follows:

- Introduced the MEDAKA-ẽL framework that uses 5 models trained across 5-fold cross-validation, and combines their outputs with Weighted Boxes Fusion (WBF) to improve detection accuracy.
- Proposed the Medaka dataset, a new collection of Medaka fish images that includes manually captured photos under different lighting colors, as well as annotated samples gathered from the internet.
- Developed an ensemble detection approach based on YOLOv8n models, where predictions are merged with WBF to keep valuable detections and reduce false positives.

2. Related Works

3. Materials and Methods

3.1. Data Gathering

Given the scarcity of publicly available Medaka fish imagery and the critical need for precise taxonomic identification in ecological monitoring, we constructed a comprehensive hybrid dataset called the *Medaka Fish Dataset*. This dataset, as shown in Figure 1, combines primary data that we collected through controlled laboratory imaging and *in situ* field observation, with curated internet-sourced images as our secondary data. This approach directly addresses the fundamental challenge specified in our introduction regarding the absence of dedicated datasets for endangered Medaka species detection while ensuring sufficient morphological diversity to support the robustness of our proposed model.

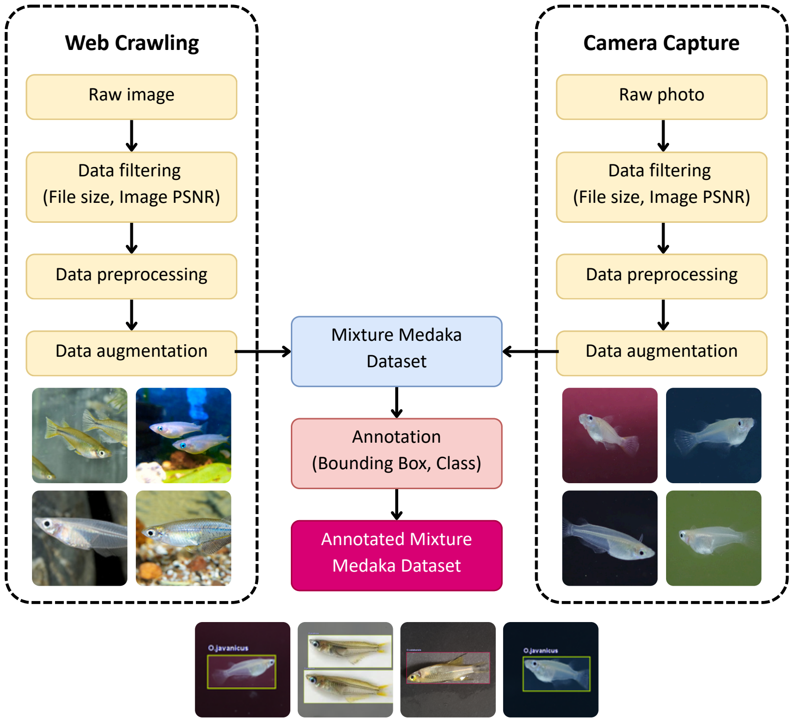

Figure 1. Representative samples from the Medaka dataset showing diversity in species, environmental conditions, and imaging scenarios. (**a**) *O. celebensis* specimens in various naturalistic settings. (**b**) *O. javanicus* specimens demonstrating morphological variation and environmental diversity.

Primary data acquisition combined (i) controlled laboratory imaging at the Genetics Laboratory, Faculty of Mathematics and Natural Sciences, Hasanuddin University, Indonesia and (ii) *in situ* observation in the Tanjung coastal freshwater-estuarine transition zone in Makassar, Indonesia. This dual setting was intentionally designed to balance morphological clarity with real-world environmental variability. In the laboratory setting, *O. celebensis* and *O. javanicus* were photographed using a Canon EOS M50 camera (24.1 MP, 6000×4000 px CMOS sensor) in modular glass aquaria with four different background colors. The camera operates at an aperture of *f/2.5* under diffuse LED lighting. This is done to help minimize light reflection and motion blur while maintaining natural colors. We also used four background colors (red, black, blue, and green) to (1) enhance contrast across pigmentation conditions, (2) avoid overfitting on single color-light pairs, and (3) approximate the natural substrate diversity.

In addition to controlled laboratory imaging, we also added *in situ* observation images to increase the variety of our data. These field images contributed heterogeneous backgrounds, such as irregular substrates, floating particles, and fluctuating lighting. However, images affected by severe turbidity, glaring light reflections, occlusions, or unclear body contours were discarded. Nearly identical temporal sequences, excessive motion blur, and noticeable chromatic aberrations were also excluded. Only minor normalization was applied, consisting of orientation adjustments and removal of unusable frames. No denoising or color correction was performed in order to preserve the genuine visual variation.

In parallel, secondary images were curated from publicly available online sources. To ensure reliability, only specimens that were clear and taxonomically identifiable were retained. Also, images with artificial backgrounds, compression artifacts, or ambiguous species characteristics were excluded to maintain the integrity of our dataset.

A total of 1,139 high-resolution images ranging from 1920x1080 to 4096x3072 pixels of Medaka fish (*Oryzias* species) were gathered from both primary and secondary sources. Then the Medaka fish in these images were annotated using Roboflow, resulting in a total of 1,280 labeled instances: 641 of *Oryzias celebensis* and 639 of *Oryzias javanicus*. Figure 2 shows the examples of annotated Medaka images from our dataset.

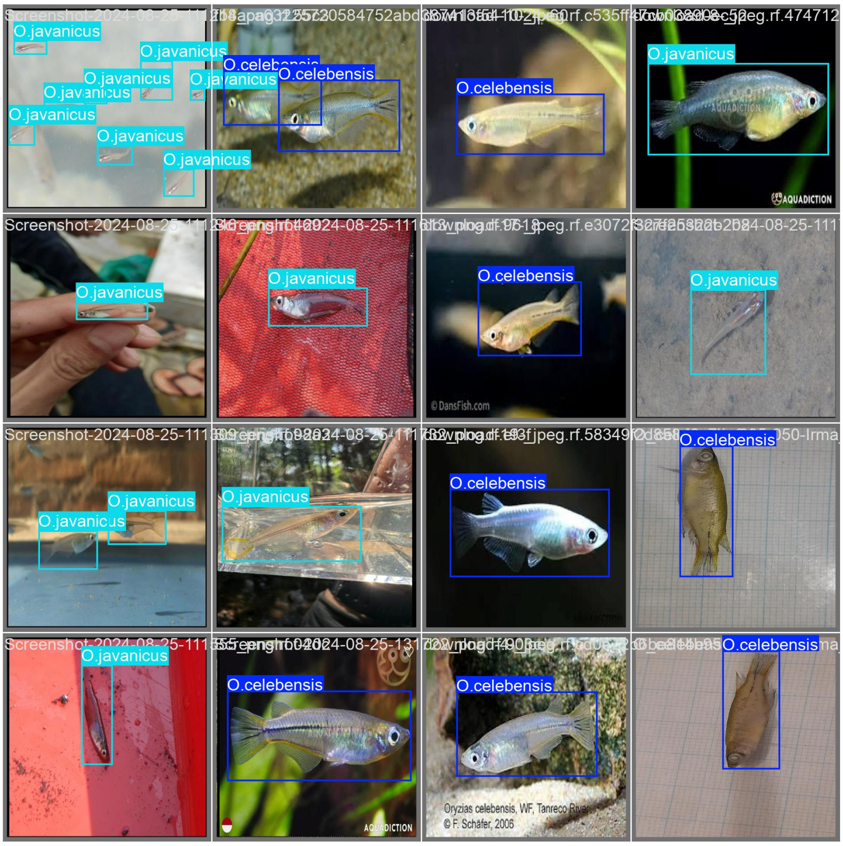

Figure 2. Examples of annotated Medaka images from the mixed dataset.

3.2. Data Filtering, Pre-processing, and Annotation

As illustrated in Figure 1, after both of raw images and raw photos were collected, we then filtered the images based on the file size and the image PSNR.

4. Results

5. Discussion

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing—original draft preparation, X.X.; writing—review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.", please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: Please add: "This research received no external funding" or "This research was funded by NAME OF FUNDER grant number XXX." and and "The APC was funded by XXX". Check carefully that the details given are accurate and use the standard spelling of funding agency names at https://search.crossref.org/funding, any errors may affect your future funding.

Institutional Review Board Statement: In this section, you should add the Institutional Review Board Statement and approval number, if relevant to your study. You might choose to exclude this statement if the study did not require ethical approval. Please note that the Editorial Office might ask you for further information. Please add "The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of NAME OF

INSTITUTE (protocol code XXX and date of approval)." for studies involving humans. OR "The animal study protocol was approved by the Institutional Review Board (or Ethics Committee) of NAME OF INSTITUTE (protocol code XXX and date of approval)." for studies involving animals. OR "Ethical review and approval were waived for this study due to REASON (please provide a detailed justification)." OR "Not applicable" for studies not involving humans or animals.

Informed Consent Statement: Any research article describing a study involving humans should contain this statement. Please add "Informed consent was obtained from all subjects involved in the study." OR "Patient consent was waived due to REASON (please provide a detailed justification)." OR "Not applicable" for studies not involving humans. You might also choose to exclude this statement if the study did not involve humans.

Written informed consent for publication must be obtained from participating patients who can be identified (including by the patients themselves). Please state "Written informed consent has been obtained from the patient(s) to publish this paper" if applicable.

Data Availability Statement: We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments). Where GenAI has been used for purposes such as generating text, data, or graphics, or for study design, data collection, analysis, or interpretation of data, please add "During the preparation of this manuscript/study, the author(s) used [tool name, version information] for the purposes of [description of use]. The authors have reviewed and edited the output and take full responsibility for the content of this publication."

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflicts of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results".

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three letter acronym
LD	Linear dichroism

Appendix A

Appendix A.1

The appendix is an optional section that can contain details and data supplemental to the main text—for example, explanations of experimental details that would disrupt the flow of the main text but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data are shown in the main text can be added here if brief, or as Supplementary Data. Mathematical proofs of results not central to the paper can be added as an appendix.

Table A1. This is a table caption.

Title 1	Title 2	Title 3
Entry 1	Data	Data
Entry 2	Data	Data

Appendix B

All appendix sections must be cited in the main text. In the appendices, Figures, Tables, etc. should be labeled, starting with "A"—e.g., Figure A1, Figure A2, etc.

References

1. Mahmudi, A.; et al. Title of the study (please update with actual title). *Journal Name* **2022**, *XX*, XXX–XXX. https://doi.org/DOIifavailable.
2. Ueda, R.; Ansai, S.; Takeuchi, H. Rapid body colouration changes in Oryzias celebensis as a social signal influenced by environmental background. *Scientific Reports* **2020**. Electronic supplementary material is available online, https://doi.org/10.6084/m9.figshare.c.7358213.
3. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. https://doi.org/10.1038/nature14539.
4. Zhao, Z.; Zheng, P.; Xu, S.; Wu, X. Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems* **2019**, *30*, 3212–3232.
5. Tang, L.; et al. Deep learning-based insect detection for precision agriculture: A comprehensive review. *Computers and Electronics in Agriculture* **2023**.
6. Ciampi, S.; et al. Automatic insect detection in sticky trap images using deep learning models. *Ecological Informatics* **2023**, *73*, 101–123.
7. Roy, A.; et al. WilDect-YOLO: A one-stage detector for endangered wildlife monitoring. *Ecological Informatics* **2023**.
8. Wenhan, Z.; et al. Enhanced YOLOv5 for wildlife detection in complex environments. *Ecological Informatics* **2024**.
9. Sun, J.; Zhang, H. YOLOv7 improvements for robust wildlife monitoring. *Ecological Informatics* **2024**.
10. Sapkota, A.; Zhang, X.; Gong, Y.; Zhang, Q. Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments. *arXiv preprint arXiv:2407.12040* **2024**.
11. Chen, M.; Hou, Y.; Li, H.; Gao, J.; Liu, W.; Gong, W. CEH-YOLO: An efficient YOLO-based detection model for underwater ecological images. *Ecological Informatics* **2024**, *79*, 102513. https://doi.org/10.1016/j.ecoinf.2024.102513.
12. Yu, X.; Liu, Z.; Xu, H.; Chen, Z.; Wu, H. YOLO-SAG: An improved YOLOv8 model for underwater target detection. *Ecological Informatics* **2024**, *80*, 102570. https://doi.org/10.1016/j.ecoinf.2024.102570.
13. Wang, R.; Zhang, X.; Chen, Y.; Li, P.; Zhang, C. SCoralDet: Efficient real-time underwater soft coral detection with YOLO. *Ecological Informatics* **2024**, *80*, 102562. https://doi.org/10.1016/j.ecoinf.2024.102562.
14. Solovyev, R.; Wang, W.; Gabruseva, T. Weighted Boxes Fusion: Ensembling Boxes from Different Object Detection Models. *Image and Vision Computing* **2021**, *117*, 104–127. https://doi.org/10.1016/j.imavis.2021.104127.