YOLOv8-Assisted Underwater Fish Species Detection From Underwater Images

Authors: Sri Sanjana K , Samyukthaa N, vishalini P, Umesh M S, Ms Vidhyabharathi T.

Students, department of computer science with artificial intelligence and data science, KPR College Of Arts Science and Research, Coimbatore, India.

E-mail: srisanjana3112@gmail.com

vidhyabharathi.t@kprcas.ac.in

Abstract

Fish species identification plays a vital role in fisheries management, biodiversity conservation, aquaculture, and marine ecosystem monitoring. Accurate and automated detection of fish species from underwater images can significantly reduce manual effort while improving monitoring efficiency. Traditional image classification approaches often struggle with varying underwater conditions such as poor illumination, occlusion, complex backgrounds, and differences in fish orientation and size. To address these challenges, this study proposes a YOLOv8-based deep learning framework for real-time fish species detection and classification. A publicly available fish image dataset was annotated using bounding boxes corresponding to different fish species and used to train the model with GPU acceleration for improved computational efficiency. The proposed framework performs simultaneous localization and species classification, enabling accurate identification of multiple fish within a single image. Experimental evaluation was conducted using standard object detection metrics, and the proposed model achieved a mean Average Precision (mAP@0.5) of 81.91%, demonstrating strong performance in detecting and classifying fish species across diverse underwater environments. The developed system can support intelligent fisheries management, marine biodiversity assessment, ecological research, and automated monitoring applications by providing fast, reliable, and accurate fish species detection with minimal human intervention.

Keywords: Fish species detection, YOLOv8, Deep learning, underwater images, marine image classification

1.Introduction

1. Fish are an important component of aquatic ecosystems and play a significant role in marine biodiversity, fisheries, aquaculture, and underwater ecological studies. Accurate identification of fish species from underwater images is essential for monitoring aquatic environments, studying species distribution, and supporting fisheries management. However, identifying fish species manually from underwater images can be a challenging and time-consuming task, particularly when a large number of images need to be analyzed. Underwater images often contain variations in illumination, background conditions, water turbidity, fish orientation, size, and partial occlusion, which can make accurate species identification difficult.

Traditionally, fish species identification has largely depended on manual observation and expert knowledge. Although experienced researchers can identify species based on their physical characteristics, manual identification requires considerable time and may be affected by differences in human interpretation. The increasing availability of underwater imaging technologies has resulted in a large volume of visual data that requires efficient and automated analysis. Consequently, there is a need for an automated and accurate approach that can detect and identify fish species from underwater images with reduced human intervention. Recent developments in deep learning and computer vision have provided effective solutions for image-based object detection. Among these approaches, the You Only Look Once (YOLO) family of models has gained considerable attention because of its ability to perform object detection with high speed and accuracy.

In this study, YOLOv8m is employed for automated fish species detection from underwater images. The proposed approach uses a fish image dataset obtained from Kaggle containing 10 different fish species. The model is trained to detect fish and identify their corresponding species from underwater images. By utilizing the capabilities of YOLOv8m, the proposed system aims to provide an efficient approach for automatically locating and recognizing fish species in complex underwater environments. The developed model is evaluated using standard object-detection performance metrics, with the obtained mAP@0.5 of 81.91% demonstrating the effectiveness of the proposed approach for fish species detection.

The main contributions made by this study are outlined below:

1. Uses a YOLOv8m-based deep learning model as the foundation for automated fish species detection from underwater images.
2. Utilizes a Kaggle-based fish image dataset containing 10 different fish species for training and evaluation of the proposed detection model.

3. Performs simultaneous fish detection and species identification, enabling the system to locate fish in underwater images and determine their respective species.

4. Addresses challenges associated with underwater images, including variations in image quality, illumination, background, fish orientation, and visibility.

5. Evaluates the trained YOLOv8m model using object-detection performance metrics, achieving a mAP@0.5 of 81.91%.

2.Related works

Table 1.comparitive summary of existing approaches for fish species detection

Author	Method	Dataset	Accuracy / mAP	Limitations
K. M. Knausgård et al. (2020)	YOLO + CNN with Squeeze-and-Excitation	Temperate fish; Fish4Knowledge	87.74% classification accuracy	Two-stage pipeline; limited training samples and underwater variation.
R. Mandal et al. (2018)	Faster R-CNN + VGG-16	Underwater RUVs; 50 species	82.4% mAP	Two-stage region-proposal architecture is heavier and slower than one-stage YOLO.
D. Rath et al. (2018)	CNN + deep learning + image processing	Underwater fish images	96.29% accuracy	Classification-focused; does not provide direct object localization like YOLO.

A. A. Muksit et al. (2022)	YOLO-Fish-1 / YOLO-Fish-2 (improved YOLOv3)	DeepFish + OzFish	76.56% / 75.70% AP	Tiny fish, complex backgrounds and varying underwater conditions remain challenging.
S. A. Santoso et al. (2024)	Faster R-CNN, SSD, MobileNet, YOLOv5 comparison	Coral fish / Chaetodontidae	92.21% mAP (SSD); 87.18% (MobileNet)	Focused on coral fish; computational speed and generalization vary by model and hardware.
P. Sarkar et al. (2023)	YOLOv3 / YOLOv4	Roboflow fish detection; 26 classes	50% mAP	Unbalanced data, unclear underwater images, occlusion and multiple fish reduce performance.
C. S. et al. (2025)	YOLOv8-TF (Transformer-enhanced YOLOv8)	SEAMPAD21; 121 evaluated species	87.9% mAP@0.5	Class imbalance, low resolution and occlusion remain difficult.
AquaYOLO (2025)	Advanced YOLO-based detector	DePondFi; DeepFish; OzFish	90.9% mAP@0.5 on DePondFi	Designed mainly for aquaculture pond monitoring; performance can vary across domains.

Based on the comparative analysis presented in Table 1, various deep learning techniques have been developed for fish detection and species recognition in underwater images. Earlier approaches such as YOLO-based and CNN-based models demonstrated promising performance but were affected by challenges including complex underwater environments, limited training data, and difficulty in recognizing fish species accurately. The YOLOv7 and YOLOv8 models achieved higher detection performance, while improved YOLOv8 architectures further enhanced feature extraction and detection accuracy. However, some advanced approaches incorporate additional modules such as transformers and attention mechanisms, which can increase computational complexity and resource requirements. These limitations highlight the need for an efficient model that can perform accurate fish species detection while maintaining reasonable computational requirements. To address these challenges, the proposed system employs YOLOv8m for detecting and classifying 10 fish species from underwater images. The proposed model achieves 81.91% mAP@0.5, demonstrating its effectiveness for underwater fish species detection.

3,.Proposed work

The proposed methodology for the fish species detection system is illustrated in Figure 1. The system begins with the collection of underwater fish images from a publicly available Kaggle dataset. The collected images are prepared and validated to ensure compatibility with the YOLOv8m framework. The images are then annotated with bounding boxes and corresponding class labels for the respective fish species. The annotated dataset is subsequently used to train the YOLOv8m model for simultaneous fish localization and species classification.

The trained model identifies the location of fish in underwater images and predicts their corresponding species. Finally, the performance of the proposed system is evaluated using standard object detection metrics, including Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95

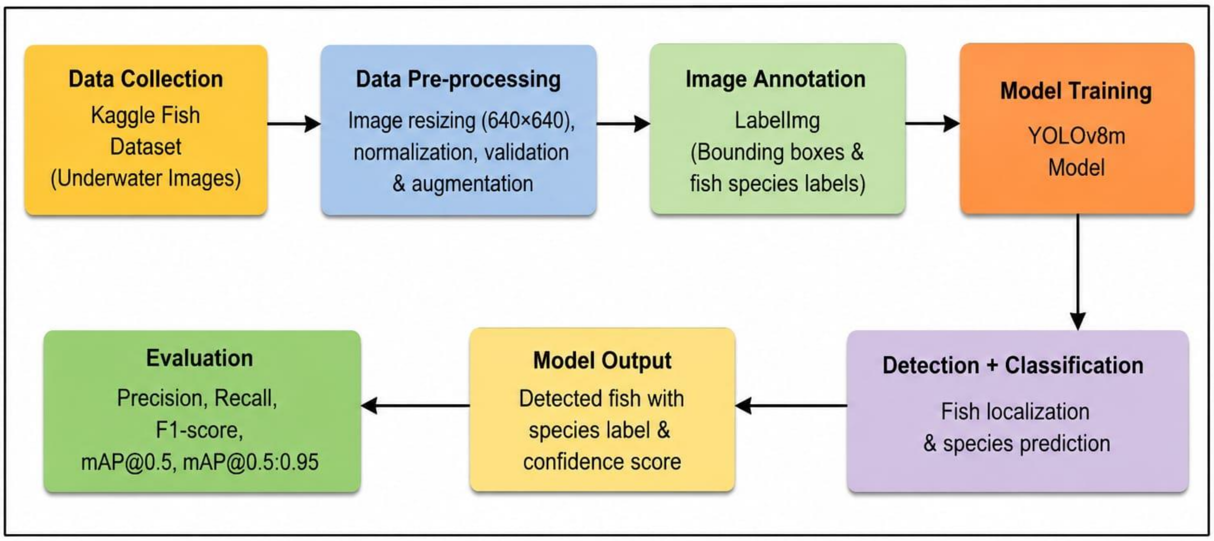


Figure 1. Workflow of the proposed fish species detection system

3.1 Data Collection

For this study, an underwater fish image dataset was obtained from Kaggle for developing the proposed fish species detection system. The dataset contains images belonging to different fish species. The evaluated classes include Black Sea Sprat, Gilt-head Bream, House Mackerel, Red Mullet, Red Sea Bream, Sea Bass, Shrimp, Striped Red Mullet, and Trout.

For the test set, 900 images were evaluated, with 100 images corresponding to each of the nine classes. The images contain variations in fish appearance, orientation, size, background, and underwater conditions. This diversity provides suitable samples for training and evaluating the fish detection model.

Table 2.total number of images per class

Class ID	FISH SPECIES	TEST IMAGES
0	Black sea sprat	100
1	Gilt head bream	100
2	House mackrel	100
3	Red mullet	100
4	Red sea bream	100

5	Sea bass	100
6	shrimp	100
7	Striped red mullet	100
8	trout	100

3.2 Data Pre-processing

Before training and evaluation, the dataset was prepared according to the directory structure and input requirements of the YOLOv8 framework. The images and their corresponding label files were organized into the appropriate dataset directories for training and validation.

During the validation process, the YOLOv8 framework automatically checked the input images for validity. As observed during validation, some images were identified as corrupted or unreadable and were therefore ignored by the framework. This ensured that invalid images did not affect the model evaluation.

The validation process was performed on 115 images containing 293 object instances. The model evaluated the annotated fish images using their bounding boxes and class labels. The resulting evaluation metrics included Precision (P), Recall (R), mAP@0.5, and mAP@0.5:0.95 for the different fish classes.

3.3 Image Annotation

The fish images were annotated by assigning the appropriate class labels and bounding boxes to the fish instances present in each image. Each bounding box represents the location of a fish, while the associated class label identifies its corresponding species.

The annotations were prepared in a format compatible with the YOLOv8 object detection framework. Accurate annotation enables the model to learn both the spatial location and visual characteristics of the different fish species. The annotated images and labels were subsequently used for model training and validation.

3.4 Model Training

The annotated fish dataset was used to train the proposed YOLOv8m model. YOLOv8m is a single-stage object detection model that performs object localization and classification simultaneously.

During training, the model learns the visual features of the different fish classes from the annotated underwater images. The learned features enable the model to predict the bounding box, class label, and confidence score of fish detected in previously unseen images.

3.5 Model Evaluation

The trained YOLOv8m model was evaluated using the validation dataset. The performance was measured using Precision, Recall, mAP@0.5, and mAP@0.5:0.95.

From the validation results, the model achieved an overall Precision of 0.812, Recall of 0.701, mAP@0.5 of 0.819, and mAP@0.5:0.95 of 0.595. Thus, the proposed YOLOv8m model achieved an mAP@0.5 of 81.9%, demonstrating its ability to detect and classify fish species from underwater images.

4. Results and Discussion

4.1 System Configuration

The proposed fish species detection system was implemented using Google Colab as the training platform. The fish image dataset was obtained from Kaggle, and the model was developed using Python 3 with GPU acceleration. An NVIDIA T4 GPU was used to accelerate the training process. The proposed system employs the YOLOv8m object detection architecture for detecting and classifying fish species from underwater images.

The model was trained for 120 epochs with a batch size of 8. The trained YOLOv8m model was subsequently evaluated using the validation dataset. The performance of the model was analyzed using Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95.

parameter	configuration
dataset source	kaggle
Training platform	Google colab
Programming language	Python 3
GPU	NVIDIA T4
Detection model	YOLOv8m
Training epochs	120
batch size	8

Table 3. System Configuration

4.2 Performance Analysis

The performance of the proposed YOLOv8m-based fish species detection system was evaluated using the validation dataset. The overall validation results obtained from the model were 0.812 Precision, 0.701 Recall, 0.819 mAP@0.5, and 0.595 mAP@0.5:0.95. Therefore, the proposed model achieved an mAP@0.5 of 81.9%, demonstrating its capability to localize and classify fish objects in underwater images.

metric	value
precision	0.812
recall	0.701
mAP@0.5	0.819
MAP@0.5:0.95	0.595

Table 4. Overall Performance Metrics

The class-wise validation results show variations in detection performance among the different fish classes. The model obtained its highest mAP@0.5 of 0.934 for fish_9, followed by 0.894 for fish_8 and 0.830 for fish_10. The lowest mAP@0.5 was observed for fish_2 with 0.643. These differences indicate that certain fish classes are easier for the model to distinguish, while visually similar classes remain more challenging.

class	precision	recall	mAP@0.5	mAP@0.5:0:95
Fish_0	0.796	0.625	0.783	0.541
Fish_2	0.768	0.500	0.643	0.453
Fish_5	0.817	0.738	0.828	0.562
Fish_8	0.730	0.905	0.894	0.697
Fish_9	0.820	0.913	0.934	0.758
Fish_10	0.940	0.525	0.830	0.557

Table 5. Class-wise Performance Metrics

4.2.1 Precision-Confidence Curve

The Precision-Confidence curve illustrates the relationship between the confidence threshold and the precision of the proposed YOLOv8m model. As the confidence threshold increases, the precision generally improves. At higher confidence levels, the model achieves precision values close to 1.0, indicating that predictions retained at high confidence are more likely to be correct. This demonstrates that increasing the confidence threshold can effectively reduce uncertain detections and improve prediction reliability.

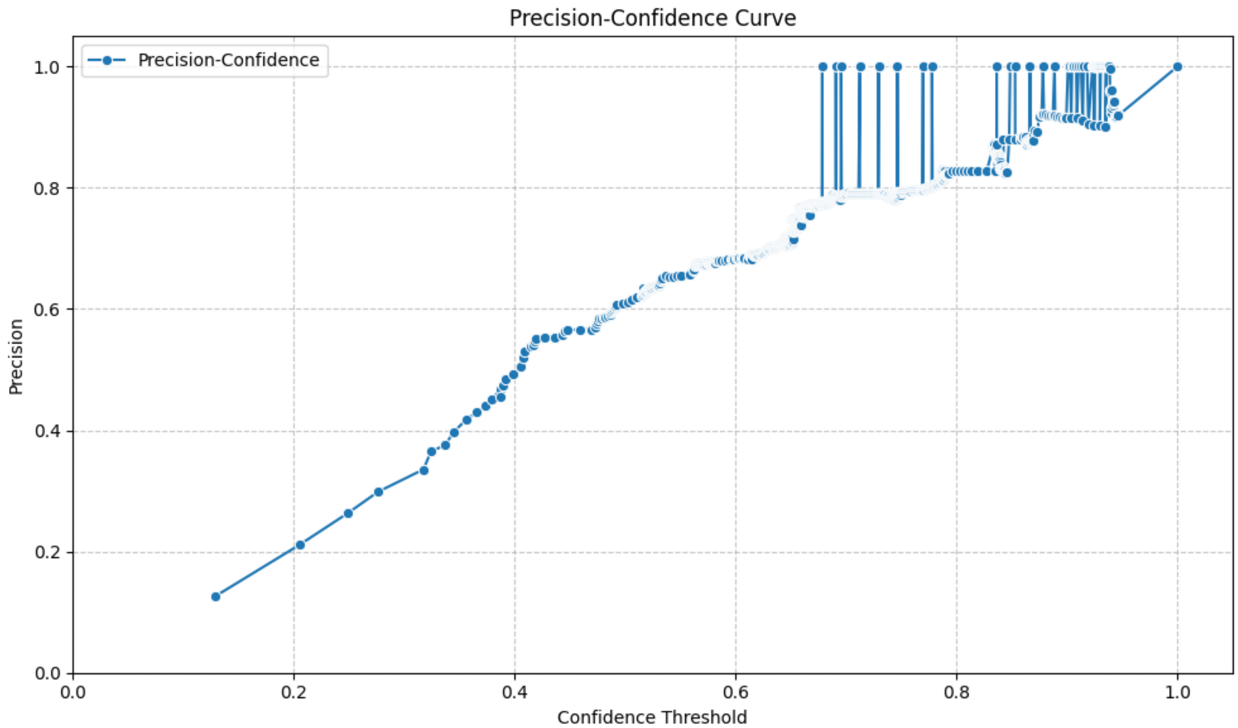


Figure 2. Precision-Confidence curve

4.2.2 Recall-Confidence Curve

The Recall-Confidence curve demonstrates the variation in recall with respect to the confidence threshold. The graph shows that recall increases progressively as the confidence threshold increases, reaching approximately 0.93 at a confidence level of around 0.82. The observed variation indicates that the selection of the confidence threshold has a significant effect on the number of correctly detected fish instances. The model therefore provides a useful balance between detection confidence and the ability to identify fish objects.

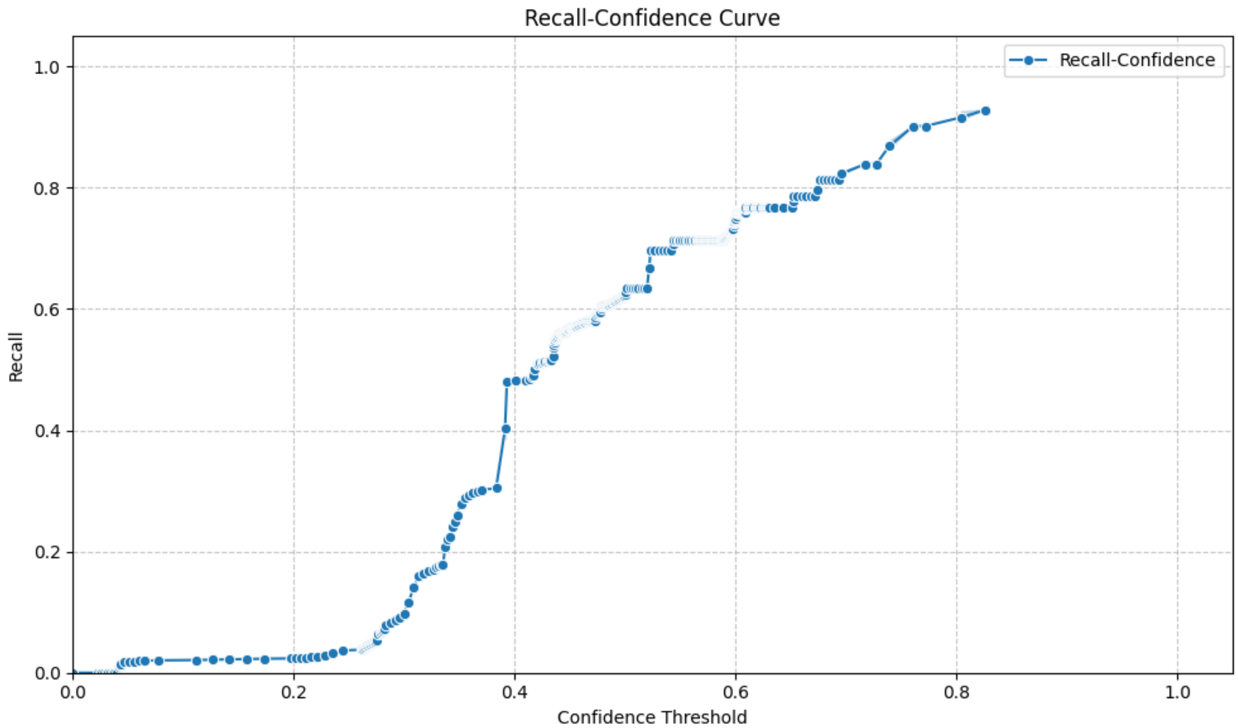


Figure 3. Recall-Confidence curve

4.2.3 F1-Confidence Curve

The F1-Confidence curve represents the relationship between the confidence threshold and the F1-score, which combines Precision and Recall. The F1-score increases considerably in the intermediate confidence range and reaches a maximum of approximately 0.75 around a confidence threshold of 0.65. Beyond this region, the curve shows some variation. This suggests that a moderate-to-high confidence threshold provides a better balance between precision and recall for the proposed fish detection system.

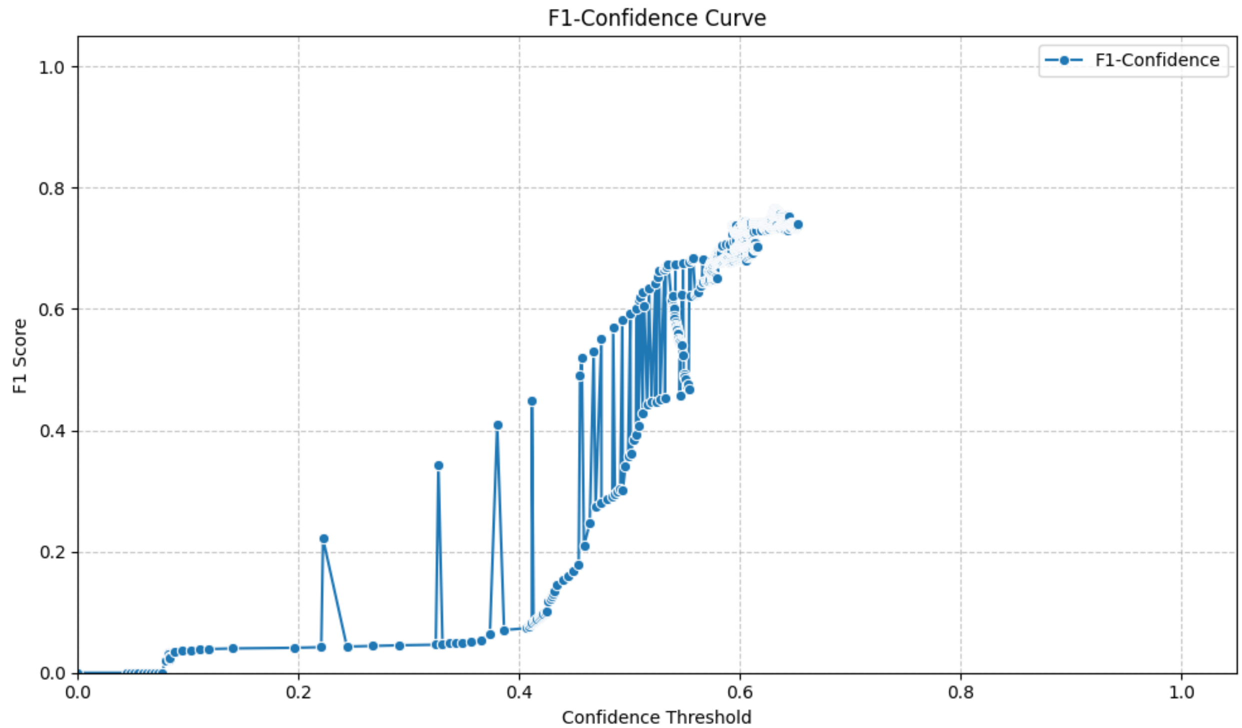


Figure 4. F1-Confidence curve

4.2.4 Normalized Confusion Matrix

The normalized confusion matrix presents the classification behavior of the YOLOv8m model across the evaluated fish classes. The diagonal values represent correctly classified instances, while the off-diagonal values indicate misclassification between classes.

The highest diagonal value observed is approximately 0.83 for fish_9, indicating strong classification performance for this class. Other relatively strong diagonal values include 0.61 for fish_5, 0.54 for fish_0, 0.53 for fish_10, and 0.50 for fish_2. Some misclassification can be observed between classes, particularly where visually similar fish species may share common characteristics in terms of shape, colour, or appearance.

For example, fish_0 shows a noticeable proportion of predictions associated with the background and other classes, while fish_5 also exhibits confusion with fish_8 and the background. These observations indicate that variations in underwater image conditions and similarities between fish appearances can affect classification performance.

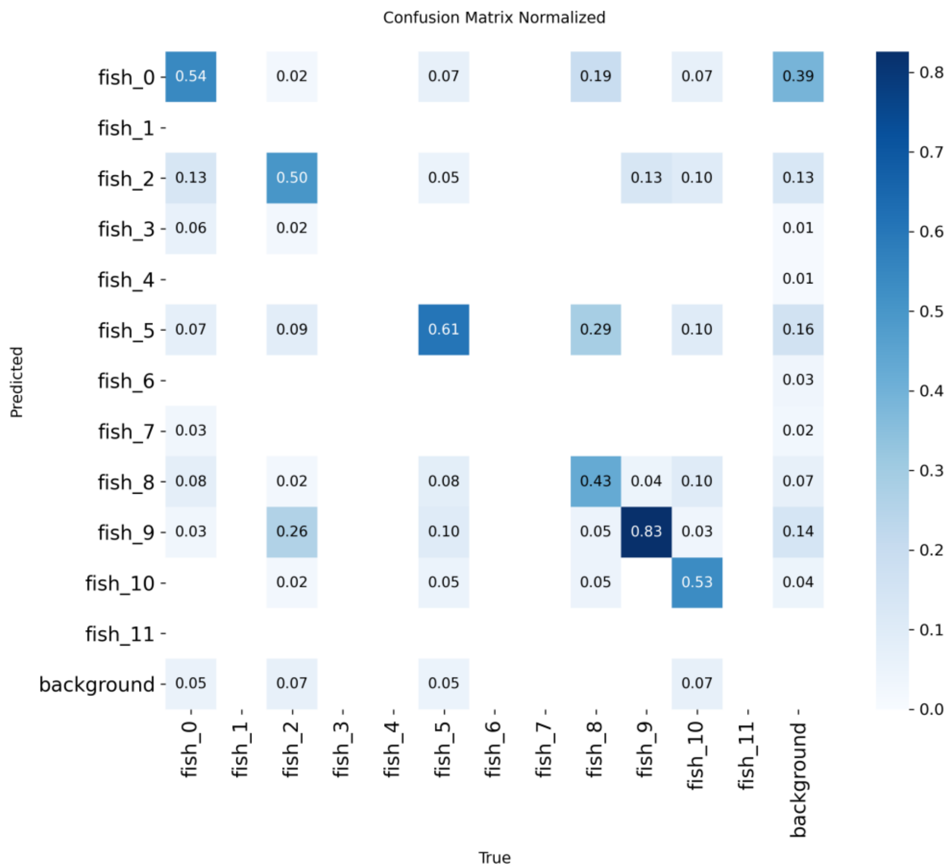


Figure 5. Normalized confusion matrix

Overall, the experimental results demonstrate that the YOLOv8m model provides effective fish species detection and classification performance, achieving an mAP@0.5 of 81.9%. The Precision-Confidence, Recall-Confidence, and F1-Confidence curves further demonstrate the effect of confidence threshold selection on model performance. The confusion matrix indicates that the model performs particularly well for some classes, while further improvement may be required for classes with greater visual similarity and lower detection performance.

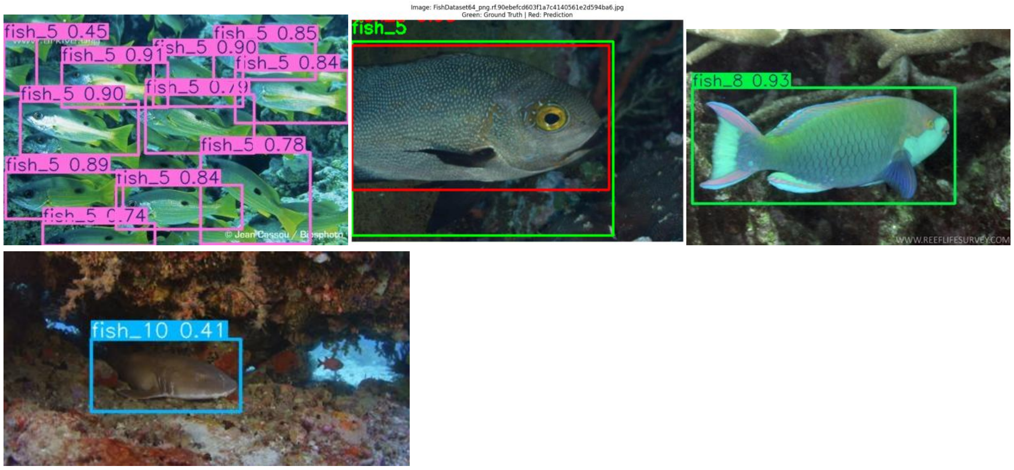


Figure 6. Sample Prediction Images

The sample prediction images demonstrate the performance of the YOLOv8m-based fish species detection model on underwater images. The model successfully identifies fish and localizes them using bounding boxes, with the predicted species label and confidence score displayed above each detection.

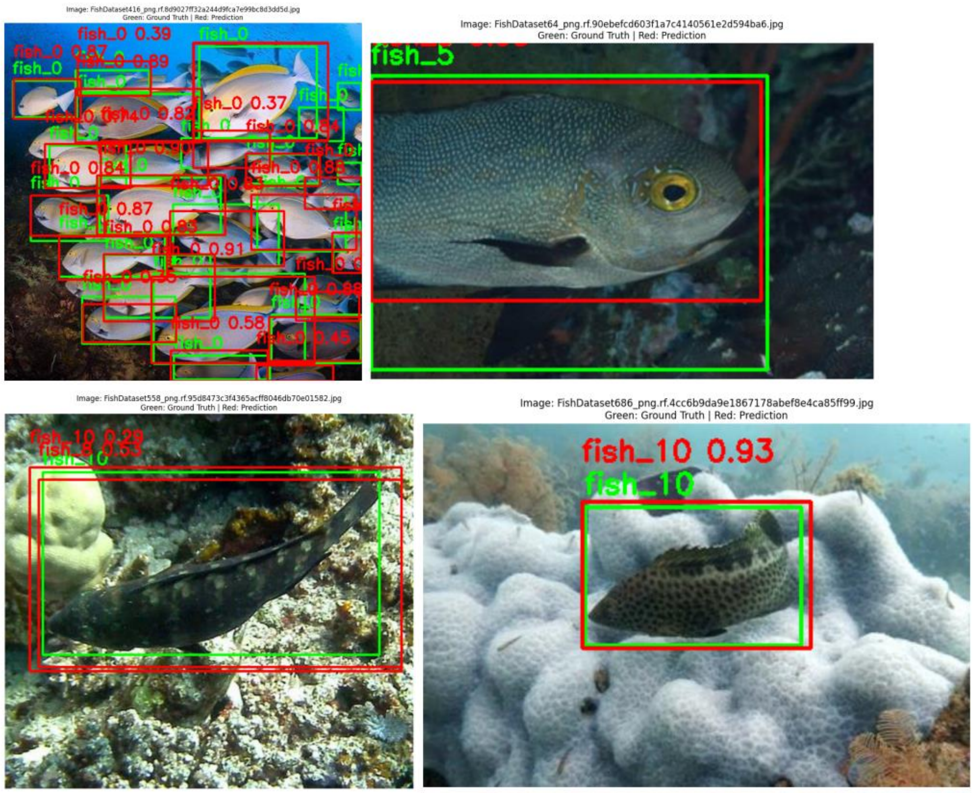
In the first image, the model detects fish_10 with a confidence score of 0.41.

In the second image, multiple instances of fish_5 are detected with confidence scores ranging approximately from 0.45 to 0.91, demonstrating the model's ability to detect multiple fish in a single image.

In the third and fifth images, fish_8 is detected with high confidence scores of 0.93 and 0.92, respectively.

The fourth image compares ground-truth and predicted bounding boxes, showing that the predicted box closely overlaps the actual fish region. This indicates effective object localization by the model.

Overall, the prediction results show that the trained YOLOv8m model can detect and localize different fish species in underwater environments, even under challenging conditions such as complex backgrounds, low visibility, and multiple fish appearing in the same image. The varying confidence scores also indicate that detection confidence depends on factors such as image quality, fish visibility, and background complexity.



However, there have been some prediction variations and errors made by the model, as illustrated in Figure X. False predictions primarily occur in cases involving complex underwater conditions, overlapping targets, or camouflage, often accompanied by lower confidence values ranging from 0.29 to 0.58.

This is largely due to the fact that visually similar benthic backgrounds, severe intra-class occlusion, and fine-grained visual differences between specific fish species are subtle and challenging for the network to distinguish. Furthermore, edge truncation during tight frame cropping leads to minor bounding box regression discrepancies despite successful feature identification.

5.conclusion

In this study, an automated fish species detection system was developed using the YOLOv8m deep learning model. The proposed system is designed to identify and classify fish species from underwater images. The dataset was obtained from Kaggle and contains images belonging to 10 different fish species. The images were prepared and organized into suitable training, validation, and testing datasets. Preprocessing and annotation procedures were performed to improve the quality of the training data. The YOLOv8m model was selected because of its strong object detection capability and efficient performance. The model was trained using Google Colab with an NVIDIA T4 GPU environment. A batch size of 8 was used during the training process. The model was trained for 120 epochs to achieve better learning and detection performance. The trained model successfully learned the visual characteristics of different fish species. It was able to detect fish objects and identify their corresponding species from underwater images. The experimental results demonstrate that the proposed approach provides reliable fish detection performance. The model achieved a mAP@0.5 of 81.91%, indicating good overall detection accuracy. The prediction results also show that the model can correctly localize fish using bounding boxes. The system performs well even when fish appear at different positions and orientations. It can also handle variations in fish appearance and underwater image conditions to a reasonable extent. The use of YOLOv8m provides a suitable balance between detection accuracy and computational efficiency. The developed system can reduce the need for manual identification of fish species. It can therefore be useful for applications such as marine biodiversity monitoring and underwater research. The approach can also support fisheries management and ecological studies. However, variations in illumination, water quality, background complexity, and overlapping fish may affect detection performance. Future work can focus on increasing the size and diversity of the dataset to improve generalization. Advanced image enhancement techniques can also be incorporated to improve detection in low-quality underwater images. Further improvements can be achieved by experimenting with optimized YOLO architectures and attention-based techniques. Overall, the proposed YOLOv8m-based system demonstrates the potential of deep learning for automatic and efficient fish species detection from underwater images.

Reference

[1] F. F. Khan, X. Li, A. J. Temple, and M. Elhoseiny, "FishNet: A Large-scale Dataset and Benchmark for Fish Recognition, Detection, and Functional Trait Prediction," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 20496–20506, 2023.

[2] N. Garcia-d'Urso, A. Galan-Cuenca, P. Pérez-Sánchez, P. Climent-Pérez, A. Fuster-Guillo, J. Azorin-Lopez, M. Saval-Calvo, J. E. Guillén-Nieto, and G. Soler-Capdepón, "The DeepFish computer vision dataset for fish instance segmentation, classification, and size estimation," Scientific Data, vol. 9, Art. no. 287, 2022.

[3] Z. Zhao, Y. Liu, X. Sun, J. Liu, X. Yang, and C. Zhou, "Composited FishNet: Fish Detection and Species Recognition From Low-Quality Underwater Videos," IEEE Transactions on Image Processing, vol. 30, pp. 4719–4734, 2021, doi: 10.1109/TIP.2021.3074738.

[4] W. Xu and S. Matzner, “Underwater Fish Detection using Deep Learning for Water Power Applications,” 2018.

[5] D. Rathi, S. Jain, and S. Indu, “Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning,” 2018.

[6] M. Y. Ouis and M. Akhloufi, “YOLO-Based Fish Detection in Underwater Environments,” Environmental Sciences Proceedings, vol. 29, no. 1, Art. no. 44, 2024, doi: 10.3390/ECRS2023-16315.

[7] Z. Zhang, Y. Qu, T. Wang, Y. Rao, D. Jiang, S. Li, and Y. Wang, “An Improved YOLOv8n Used for Fish Detection in Natural Water Environments,” 2024.

[8] J. Wang and N. Yu, “UTD-Yolov5: A Real-time Underwater Targets Detection Method based on Attention Improved YOLOv5,” 2022.

[9] S. Anwar and C. Li, “Diving Deeper into Underwater Image Enhancement: A Survey,” 2019.

[10] X. Cong, Y. Zhao, J. Gui, J. Hou, and D. Tao, “A Comprehensive Survey on Underwater Image Enhancement Based on Deep Learning,” 2024.

[11] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLOv8,” version 8.0.0, 2023.

[12] A. R. Murtadlo, E. Adriono, and Y. E. Windarto, “Real-Time Fish Detection with YOLOv8 on a Semi-Autonomous Underwater Vehicle for Aquaculture Monitoring,” 2025.