Enhancing Aquarium Fish Classification through YOLO: A Deep Learning Approach

2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP) | 979-8-3315-3149-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/IDAP64064.2024.10710813

1st Murat ISIK
Computer Engineering
Faculty of Engineering and Architecture
Kırşehir Ahi Evran University
Kirsehir, Turkiye
muratisik@ahievran.edu.tr

2nd Mehmet Ali YALCINKAYA
Computer Engineering
Faculty of Engineering and Architecture
Kırşehir Ahi Evran University
Kirsehir, Turkiye
mehmetyalcinkaya@ahievran.edu.tr

***Abstract*— In recent years, advancements in deep learning have revolutionized the field of image classification and object detection. This study presents a novel application of the YOLO (You Only Look Once) model for the classification of aquarium fish species. The model is designed to accurately identify and classify five distinct species: Percula Clownfish, Moorish Idol, Yellow Tang, Queen Angel Fish, and Blue Tang. Our approach leverages the YOLO v8 architecture due to its balance between speed and accuracy, making it suitable for real-time applications. The model was trained on a comprehensive dataset of annotated images collected manually from the internet, capturing various poses, lighting conditions, and background variations to enhance robustness. Experimental results demonstrate that the YOLO v8 model achieves high precision and recall rates, with an overall accuracy exceeding 95.78%. The model's performance was evaluated using accuracy, precision, recall and F1-Score metrics. This research showcases the potential of advanced neural networks in automated aquatic biodiversity monitoring and contributes to the growing body of work in marine biology and conservation technology.**

***Keywords*— computer vision, YOLO, artificial intelligence technologies**

I. INTRODUCTION

The accurate classification of fish species in aquariums is a critical task for marine biologists, aquarium enthusiasts, and conservationists. Traditional methods of species identification typically rely on manual observation and expert knowledge, which can be time-consuming and susceptible to errors, particularly in environments with a diverse array of species and complex visual backgrounds. To address these challenges, recent advancements in deep learning and computer vision have created new opportunities for automated species classification. Among the most promising techniques in this domain is the YOLO (You Only Look Once) model.

The YOLO model is well-known for its balance between speed and accuracy, making it an ideal candidate for real-time object detection and classification tasks. In contrast to traditional convolutional neural networks, which require multiple passes over an image, YOLO performs detection in a single forward pass, significantly reducing computation time while maintaining high accuracy [1]. This efficiency is particularly beneficial in applications involving video streams and real-time monitoring, including aquarium fish classification.

Underwater environments exhibit significant variability in terms of lighting, background complexity, and the presence of multiple species, all of which can impact the performance of image classification models. Recent studies have highlighted YOLO's effectiveness in various aquatic scenarios, demonstrating its robustness in handling diverse and complex backgrounds [2, 3]. In this study, a YOLO model was developed and trained to classify five species of aquarium fish: Percula Clownfish, Moorish Idol, Yellow Tang, Queen Angel Fish, and Blue Tang. The training dataset was sourced from the internet and manually labeled using MakeSenseAI, an annotation tool that facilitates efficient and accurate labeling of image datasets. The dataset included images of these species captured in various poses and lighting conditions.

The implications of this research extend well beyond aquarium management. Automated fish classification systems have the potential to play a significant role in biodiversity monitoring, marine biology research, and conservation efforts. By offering a reliable and efficient tool for species identification, the model developed in this study can aid researchers in tracking fish populations, studying behavioral patterns, and implementing effective conservation strategies.

II. RELATED WORKS

Ma et al. (2023) review the application of deep learning-based object detection techniques in fish aquaculture. The study focuses on the advantages of using deep learning models, such as YOLO, for monitoring fish behavior, counting, and measuring body length in aquaculture settings. The review highlights the challenges posed by complex environments like varying illumination and water turbidity, as well as fish characteristics such as scale variation and occlusion. The authors conclude that deep learning models provide significant improvements in accuracy and efficiency over traditional methods, making them invaluable for the sustainable management of aquaculture systems [4].

Hamzaoui et al. (2023) present an improved deep learning model for underwater species recognition in aquaculture. The study aims to enhance the detection and classification accuracy of underwater species by refining traditional deep learning approaches. Using a custom dataset, the authors demonstrate that their model significantly outperforms existing techniques in terms of precision, recall, and F1-score, particularly under challenging underwater conditions. This research underscores the potential of advanced neural networks in enhancing automated monitoring and management practices in aquaculture [2].

Lathifah et al. (2020) explore the use of the YOLO model for fast and accurate fish classification from underwater video footage. The study emphasizes the importance of real-time detection capabilities in scenarios where rapid identification is crucial. By applying YOLO, the authors achieve high classification accuracy with reduced computational complexity, making the model suitable for deployment in real-world aquatic environments. The findings suggest that YOLO is effective for applications requiring quick and reliable fish classification in dynamic underwater settings [3].

Zhang et al. (2022) focus on the detection and classification of temperate fish species using a deep learning-based approach. The study integrates multiple neural network architectures, including YOLO, to address the challenges of identifying fish species with similar appearances in complex aquatic environments. The model demonstrates high accuracy and robustness, outperforming traditional methods. This research contributes to the growing body of work on applying deep learning to marine biology and conservation, particularly in temperate regions [5].

In recent years, there has been a growing body of research focusing on the application of deep learning models, particularly YOLO, for fish classification and detection in various aquatic environments. These works collectively underscore the significant advancements made in automating fish classification tasks using state-of-the-art neural networks. This study contributes to this growing field by applying the YOLOv8 model specifically to aquarium fish classification, achieving high accuracy and reliability. By focusing on a curated dataset of five distinct fish species, this research not only validates the efficacy of YOLOv8 in controlled environments but also offers insights into its potential applications in broader marine biology and conservation efforts.

III. Method

The development of the model was carried out in five distinct steps. Figure 1 provides a visual representation of these steps. Each of these steps is detailed in the sections below.

Fig. 1. The roadmap to developing the model

A. Data Collection

The dataset used for training and evaluating the model was meticulously curated from various online sources. Images representing five species of aquarium fish—Percula Clownfish, Moorish Idol, Yellow Tang, Queen Angel Fish, and Blue Tang—were selected. A total of 1,000 images per species were gathered, resulting in a comprehensive dataset comprising 5,000 images. This diverse dataset included images captured under varying lighting conditions, backgrounds, and poses, thereby ensuring the model's robustness and generalization.

Fig. 2. An image of each fish class, displayed from left to right: Percula Clownfish, Moorish Idol, Yellow Tang, Queen Angel, and Blue Tang.

B. Data Annotation

To facilitate accurate training, the collected images were manually annotated using the MakeSenseAI platform, a web-based annotation tool designed to streamline the process of labeling objects within images. Each fish in the dataset was annotated with bounding boxes to ensure precise localization and classification. This detailed annotation process is essential for the model to effectively learn the distinguishing features of each species. Figure 3 presents a screenshot from the MakeSenseAI platform during the annotation process. The annotation information was saved as text files (.txt). To train the YOLO model, each image and its corresponding annotation must adhere to a specific folder structure. Figure 4 illustrates the organization of the dataset's folder structure. The dataset is divided into two main directories: images and labels, each further subdivided into three subdirectories: train, test, and val. The train folder within the images directory contains the training images, while the corresponding train folder within the labels directory contains the annotation files. Similarly, the test folder contains the test files, and the val folder contains the validation files. Figure 5 illustrates the content of an annotation file, where the first number indicates the class (e.g., class 0), and the subsequent four numbers specify the fish's location within the image.

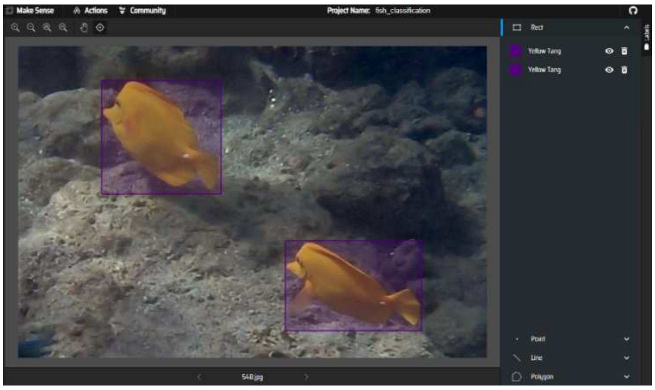

Fig. 3. A screenshot from MakeSenseAI

Fig. 4. The folder structure of the dataset

Fig. 5. The content of an annotation file

C. *Model Selection and Training*

YOLO is particularly suited for the task of classifying aquarium fish species due to its efficiency and accuracy in real-time object detection. Unlike traditional object detection methods, which involve a multi-stage process, YOLO applies a single neural network to the entire image, simultaneously predicting multiple bounding boxes and class probabilities [1, 6].

The training process involved several stages, including data preprocessing, model training, and evaluation. The dataset was divided into training, validation, and test sets, with 80% of the images allocated for training, 10% for validation, and 10% for testing. The model was trained for 100 epochs using a batch size of 32 and an input image size of 480x480 pixels. The Adam optimizer was employed with an initial learning rate of 0.001, which was progressively reduced using a cosine annealing schedule. The loss function consisted of a combination of localization loss, objectness loss, and classification loss, all optimized to enhance the model's performance across all metrics.

D. *Implementation*

The model implementation was carried out using the Python programming language, leveraging the PyTorch, TensorFlow, and Keras libraries for neural network development. The YOLO repository from GitHub was utilized, providing a comprehensive implementation of the YOLO architecture along with medium pre-trained weights. This repository is well-documented and widely recognized within the research community [7], offering a reliable foundation for the study.

E. *Testing*

The performance of the trained model was evaluated using standard metrics for object detection, including accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly identified instances among the total instances. Precision, defined as the ratio of true positive detections to the sum of true positives and false positives, indicates the exactness of the model. Recall, the ratio of true positive detections to the sum of true positives and false negatives, reflects the model's ability to capture all relevant instances. The F1-score, calculated as the harmonic mean of precision and recall, provides a single metric that balances both precision and recall.

IV. Results

This chapter presents the results of the experiments conducted to evaluate the performance of the YOLOv8 model for aquarium fish classification using the manually collected and annotated dataset. The primary objectives were to assess the model's performance and provide a qualitative analysis of its effectiveness.

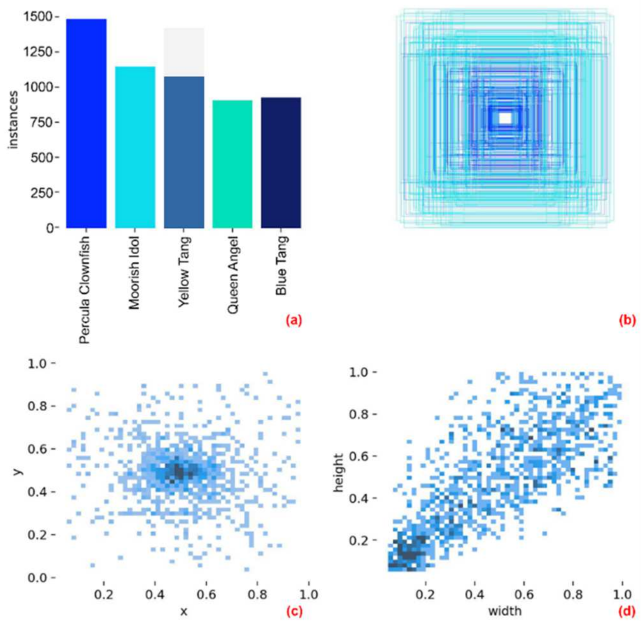

Fig. 6. The dataset characteristics

Figure 6 provides valuable insights into the distribution of the dataset and the characteristics of the bounding boxes used for the model. Figure 6a displays the number of instances for each class within the dataset. Figure 6b visualizes the distribution of bounding boxes for the fish species, where the concentration of rectangles near the center indicates that many objects tend to be located in the middle of the images. The variety of sizes and aspect ratios of the bounding boxes reflects the diversity in fish sizes and their positions within the images. Figure 6c presents a heatmap depicting the distribution of object (fish) centers based on their x and y coordinates within the images. The darker concentration in the center suggests that most objects are centered in the images, with a somewhat spread distribution, but a clear tendency for objects to be positioned closer to the center. Finally, Figure 6d shows a heatmap representing the distribution of bounding box width and height, revealing a positive correlation between the two dimensions, indicating that larger objects tend to have proportional increases in both dimensions.

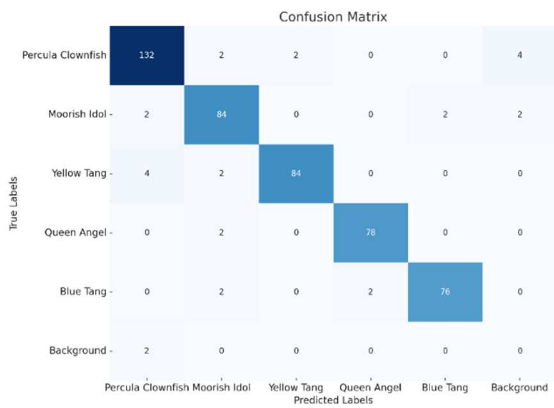

Fig. 7. Confusion Matrix

Figure 7 shows the confusion matrix generated when the test data was applied to the trained model. The "Background" class, produced by YOLO in the confusion matrix, typically represents portions of the image that do not belong to any of the defined fish classes. However, this class is not included in the calculation of metrics such as precision, recall, and F1-

score. The primary reason for its exclusion is that these metrics are designed to evaluate the model's performance in detecting and classifying the object classes of interest. Including the "Background" class could skew these metrics, leading to results that do not accurately reflect the model's ability to identify and classify the relevant objects.

TABLE I. QUALITATIVE RESULTS

Class	Precision	Recall	F1-Score
Percula Clownfish	0.96	0.97	0.96
Moorish Idol	0.91	0.95	0.93
Yellow Tang	0.98	0.93	0.95
Queen Angel	0.98	0.98	0.98
Blue Tang	0.97	0.95	0.96
Accuracy overall	0.96		

The quantitative results are summarized in Table 1. The evaluation metrics—precision, recall, and F1-score—provide a comprehensive assessment of the model's ability to correctly identify and classify the fish species. Figure 8 presents examples of the proposed YOLOv8 model's predictions on the test data. This figure illustrates the model's capability to accurately detect and classify various aquarium fish species within different images.

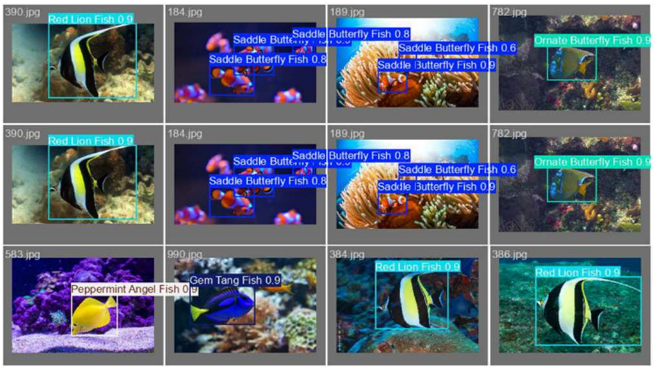

Fig. 8. The examples of the model from test data

V. DISCUSSION

The dataset contains 5,000 images, with 4,000 used for training, 500 for validation, and 500 for testing. Although Figure 6a shows more fish than the total number of images, this is because some images contain multiple fish. Additionally, Figure 6 illustrates that most fish are centered within the images, and their bounding box sizes are well-suited to the image proportions.

The results from Table 1 demonstrate that the YOLOv8 model performs exceptionally well across all fish classes, achieving an overall accuracy of 96%. The Queen Angel class achieved the highest F1-score (0.98), indicating that the model excels at detecting and classifying this species with both high precision (0.98) and recall (0.98). Similarly, the Yellow Tang and Blue Tang classes also show strong performance, with F1-scores of 0.95 and 0.96, respectively.

The Moorish Idol class, while still performing well, exhibits slightly lower precision (0.91) compared to the other classes, resulting in an F1-score of 0.93. This suggests that although the model is effective at detecting Moorish Idol, there is a slightly higher rate of false positives for this class compared to others.

Overall, the model demonstrates robust performance, with consistently high precision, recall, and F1-scores across all classes, making it a reliable tool for the accurate identification of these fish species. The slight variation in precision for the Moorish Idol class suggests a potential area for further refinement or additional data collection to improve detection accuracy for this species.

VI. CONCLUSION

In this study, we developed and evaluated a YOLOv8-based model for the classification of aquarium fish species. The results demonstrate that the model achieves high accuracy, precision, recall, and F1-scores across all classes, with an overall accuracy of 96%. The model's robust performance in real-time detection and classification tasks highlights its potential as a valuable tool for applications beyond aquarium management, including biodiversity monitoring, marine biology research, and conservation efforts. However, the slight variation in precision observed for the Moorish Idol class suggests areas for potential refinement.

VII. FUTURE WORK

Future work could focus on expanding the model to include additional species and integrating it with underwater robotic systems for in-situ monitoring, further enhancing its applicability in real-world scenarios.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779-788.

[2] M. Hamzaoui, M. Ould-Elhassen Aoueileyine, L. Romdhani, and R. Bouallegue, "An Improved Deep Learning Model for Underwater Species Recognition in Aquaculture," Fishes, vol. 8, no. 10, p. 514, 2023.

[3] H. M. Lathifah, L. Novamizanti, and S. Rizal, "Fast and accurate fish classification from underwater video using you only look once," in IOP Conference Series: Materials Science and Engineering, 2020, vol. 982, no. 1: IOP Publishing, p. 012003.

[4] H. Liu, X. Ma, Y. Yu, L. Wang, and L. Hao, "Application of deep learning-based object detection techniques in fish aquaculture: a review," Journal of Marine Science and Engineering, vol. 11, no. 4, p. 867, 2023.

[5] K. M. Knausgård et al., "Temperate fish detection and classification: a deep learning based approach," Applied Intelligence, vol. 52, no. 6, pp. 6988-7001, 2022.

[6] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," multimedia Tools and Applications, vol. 82, no. 6, pp. 9243-9275, 2023.

[7] K. Y. Chen, M. Toro-Moreno, and A. R. Subramaniam, "GitHub is an effective platform for collaborative and reproducible laboratory research," *arXiv preprint arXiv:2408.09344,* 2024.