Tools for ecosystem monitoring based on fish detection and classification using deep neural networks.

Oriol Prat[1], Pol Baños[1], Enoc Martinez[1], Joaquin del Rio[1]
1 Universitat Politècnica de Catalunya, 08800 Vilanova i la Geltrú, Spain

Abstract – This study explores the transformative impact of artificial intelligence (AI) in ecosystem monitoring, specifically object detection with YOLO (You Only Look Once), emphasising the search for optimal tools and model efficiency. The shift from manual counting to AI-based detection significantly reduces time investment. Methodologically, the YOLO model is employed, and comprehensive training strategies are outlined. The threefold data division ensures unbiased evaluation, and diverse configurations are explored for optimal model performance. Key metrics, including IoU, Precision, Recall, and mAP, along with tools like confusion matrices, contribute to a thorough understanding of the model's capabilities. Additionally, the model itself serves as a semi-automatic labelling tool.

Keywords – *Artificial intelligence, object detection, classification, YOLOv8, ecosystem monitoring*

I. INTRODUCTION

The utilisation of artificial intelligence (AI) is expanding across various knowledge domains, contributing to the acceleration of numerous work processes. One such application is in object detection and classification, particularly within the context of ecosystem monitoring. The assessment of fish populations holds significant ecological importance, as understanding their status is imperative for determining sustainable and appropriate fishing impacts while considering the resilience of their habitats. Traditionally, biologists have manually counted individual species in images to derive metrics like abundance, biomass, and biodiversity. Today, the advent of computer vision tools, such as AI-based object detection, allows for a substantial reduction in the time invested in this process. An illustrative example of the comparison between AI and human methodologies has been explored, revealing improvements in results with the application of computer vision methods [1].

In this context, it is crucial to recognize the transformative potential of AI, as it enhances the efficiency and accuracy of ecosystem monitoring. This technological advancement contributes to a more streamlined and effective approach to fish population assessment, ultimately aiding in the development of sustainable environmental management practices.

II. METHODOLOGY – KEY STEPS

There are several models that serve these methodologies through neural networks, one of the most established and stable is YOLO (You Only Look Once). YOLO divides an input image into a grid and each grid cell predicts multiple bounding boxes along with class probabilities. This grid-based approach allows YOLO to achieve impressive detection speeds. In addition, it uses a convolutional neural network architecture to extract relevant features of the input image, enabling accurate object detection [2].

In the machine learning workflow, data is typically divided into three sets: the training set (70%), used to train the model by exposing it to patterns and relationships within the data; the validation set (20%), employed for fine-tuning and adjusting hyperparameters to enhance the model's generalisation; and the test set (10%), serving as the final evaluation step to assess the model's performance on unseen data before deployment. This threefold division ensures a comprehensive and unbiased evaluation of the model's learning, tuning, and generalisation capabilities across distinct stages of development.

To optimise model performance on a custom dataset, an extensive series of training iterations was conducted, exploring diverse configurations to obtain optimal results, i.e., the most effective weight adjustments. The experimentation encompassed a range of factors, including three distinct species number groups, typically representing 30%, 60%, and 90% of individuals in each species. Additionally, five model architectures were employed; Nano, Small, Medium, Large, and Extra Large.

To further enhance model adaptability, a variety of learning rates, diverse data augmentation parameters (rotation, translation, scaling, shearing, colour jittering and other transformations), different image size and random initial seeds were systematically employed in the training process. This comprehensive approach aimed to identify the combination of characteristics that yields the best model performance for the specific dataset under consideration. *Figure 1* shows a schematic representation of steps explained above.

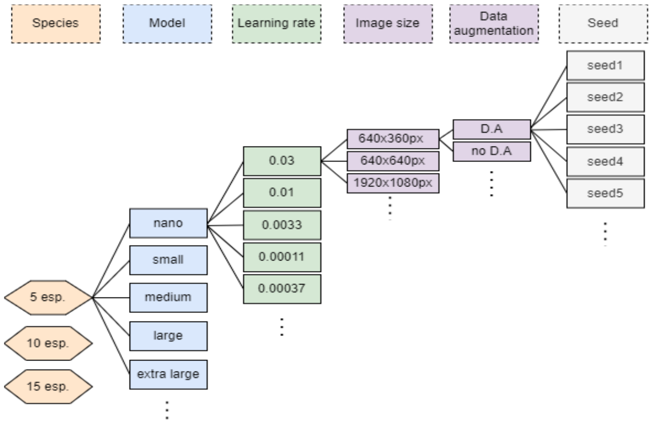

Figure 1. Schematic representation of the combinations used in each training.

Key metrics for evaluating the model include Intersection over Union (IoU), Precision, Recall, and mean Average Precision (mAP). IoU measures the accuracy of predicted bounding boxes, while Precision and Recall assess the model's ability to make accurate positive predictions comprehensively. mAP provides a holistic evaluation of precision across various IoU thresholds. Tools like confusion matrices, COCO metrics, visualisation tools, and TensorBoard contribute to a comprehensive understanding of the model's performance, along with YOLOv8's evaluation script for precision, recall, and mAP calculation.

Ultimately, when a sufficiently good model is achieved (e.g., with mAP between 0.8 and 0.9), it can be employed to automatically label new images, augmenting both the quantity and quality of the training dataset. This results in a semi-automatic labelling tool, leveraging the model's precision to efficiently expand the dataset. *Figure 2* shows the steps carried out.

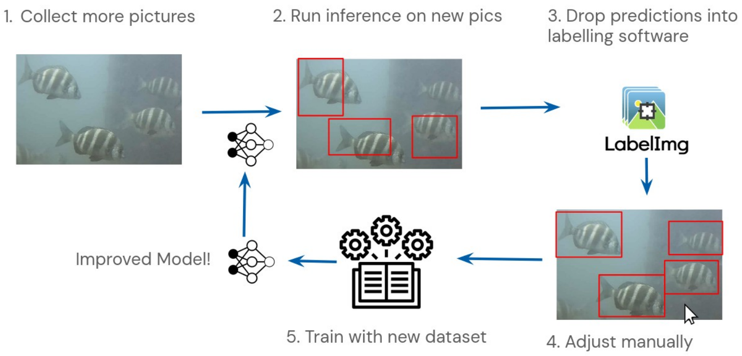

Figure 2. Following steps to label the images in a semi-automatic way.

III. CONCLUSIONS

This work culminates in the selection of best practices in the procedure employed to obtain an optimised fish detection model, after analysing and choosing hyperparameters that best fit the original dataset. The comparison of metrics and evaluation tools derived from a wide variety of training sessions enables the selection of the most optimal weights for the convolutional neural network. Additionally, the semi-automatic labelling tool allows for the improvement and optimization of the model.

REFERENCES

[1] Ditria, E. M., Lopez-Marcano, S., Sievers, M., Jinks, E. L., Brown, C. J., & Connolly, R. M. (2020). Automating the analysis of fish abundance using object detection: optimizing animal ecology with deep learning. Frontiers in Marine Science, 429.

[2] Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Computer software]. https://github.com/ultralytics/ultralytics