A DETECTRON2-BASED FRAMEWORK FOR REAL-TIME FISH DETECTION

Biju Angalees
Research Scholar, Department of Electronics and Computer Science, Rashtrasant Tukdoji Maharaj Nagpur University, St. Francis De Sales College, Nagpur, India
bijuangalees@gmail.com

Anupama D Sakhare
Associate Professor, Department of Electronics and Computer Science, Rashtrant Tukdoji Maharaj Nagpur University Nagpur, India
adsakhare@gmail.com

Abstract
This study presents a robust fish detection and classification approach using the Detectron2 framework, specifically implementing a Faster R-CNN architecture, on the NCFM (Nature Conservancy Fisheries Monitoring) dataset. The NCFM dataset, which contains diverse and challenging underwater fish images, was utilised to train and evaluate the Model's performance. Detectron2, a state-of-the-art object detection library, was employed due to its high accuracy and robustness in handling complex object instances, which is critical for this task. The Model was fine-tuned to detect and classify fish species in complex underwater environments, addressing occlusion, varying lighting conditions, and background clutter. Experimental results demonstrate that the proposed method achieves an F1 score of 86%, indicating a strong balance between precision and recall. This performance highlights the effectiveness of the Detectron2 framework in handling the intricacies of underwater fish detection and classification. The findings of this research contribute to the advancement of automated marine life monitoring systems, offering potential applications in fisheries management, biodiversity conservation, and ecological research. Further improvements could focus on optimising the Model for rare species detection and enhancing robustness to environmental variability.
Keywords: *Fish Detection, Deep Learning, Marine Fish Monitoring, Detectron2*

Introduction

The monitoring and classifying of fish species in marine environments play a critical role in fisheries management, biodiversity conservation, and ecological research [1]. Traditional fish identification and counting methods often rely on manual observation, which is time-consuming, labour-intensive, and prone to human error. With the rapid advancements in computer vision and deep learning, automated systems have emerged as a promising solution to address these challenges [2]. Object detection models, particularly those based on convolutional neural networks (CNNs), have successfully detected and classified objects in complex environments, including underwater scenarios [3]. Among the various object detection frameworks, two-stage detectors have demonstrated exceptional accuracy, particularly in complex scenarios with high object occlusion[4]. The Detectron2 framework, developed by Facebook AI Research (FAIR), offers a flexible and powerful implementation of state-of-the-art models, including Faster R-CNN [5]. Faster R-CNN's two-stage approach, which first involves a Region Proposal Network (RPN) to identify potential object locations, makes it particularly well-suited for the challenges of underwater image analysis. This method is highly effective when fish are small, occluded, or visually similar to the background. The Nature Conservancy Fisheries Monitoring (NCFM) dataset[6], which comprises a diverse collection of underwater fish images, serves as an excellent benchmark for evaluating the performance of such models in real-world scenarios. This dataset presents unique challenges, including occlusions, varying lighting conditions, and complex backgrounds, which accurately represent the difficulties encountered in marine environments.

This study utilises the Detectron2 framework to implement a Faster R-CNN model for detecting and classifying fish species in the NCFM dataset[7]. We aim to develop a robust and efficient system that accurately identifies fish species in challenging underwater conditions. We evaluate the Model's performance using the F1 score, a metric that balances precision and recall, and achieve a score of 86%. This result demonstrates the potential of modern two-stage detectors, accessible through Detectron2, in automating fish detection and classification tasks, contributing to the development of advanced marine monitoring systems. The findings of this research have significant implications for fisheries management, biodiversity conservation, and the broader field of marine ecology.

Methodology

In this research, we utilised the Detectron2 library to implement a Faster R-CNN model for fish detection and classification using the NCFM dataset, the dataset fish image distribution is shown in Figure 1, which comprises diverse underwater fish images captured under challenging conditions. The dataset was preprocessed by converting annotations to the COCO format (compatible with Detectron2), splitting them into training, validation, and test sets (70:20:10), and applying data augmentation techniques like cropping, rotation, and flipping. The Faster R-CNN model was configured with an input size of 832x832 pixels and trained for 200 epochs using an SGD (Stochastic Gradient Descent) optimiser with momentum and a learning rate of 0.001. The training involved transfer learning from a model pre-trained on ImageNet, minimising the multi-task loss (which combines RPN objectness and box regression losses with the final classification and box regression losses), and validation to prevent overfitting. The Model achieved an F1 score of 86%, demonstrating strong performance in detecting and classifying fish species despite challenges like occlusions and low-light conditions. Post processing with non-maximum suppression (NMS) further refined the results, ensuring accurate and concise predictions.

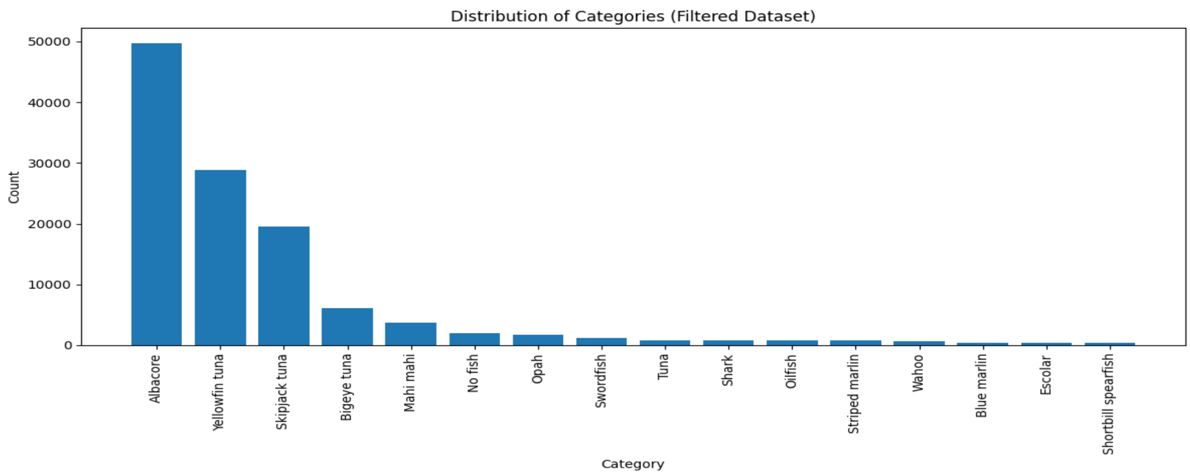

Figure 1 Distribution of Fish Images in NCFM Dataset

Results and Discussion

The results of this study demonstrate the effectiveness of the Faster R-CNN model implemented via Detectron2 in detecting and classifying fish species in the challenging NCFM dataset, achieving an F1 score of 86%. Using a higher input resolution of 832x832 pixels significantly enhanced feature extraction, enabling the Model to capture finer details and improve detection accuracy, particularly for distinguishing visually similar species and detecting small or partially occluded fish. However, this also increased computational demands, necessitating high-performance hardware. The NCFM dataset presented challenges, including mislabeled images, low-light conditions, and low-quality images, which introduced noise and reduced visibility. Despite these issues, the Model exhibited robustness, leveraging data augmentation techniques such as histogram equalisation and contrast adjustment to improve performance. The achieved F1 score reflects a strong balance between precision and recall, highlighting the Model's ability to minimise false positives and detect most fish instances, even in complex underwater environments. However, limitations such as the misclassification of visually similar species and difficulties in detecting small or occluded fish were observed. These could be addressed in future work through multi-scale training, attention mechanisms, and expanded training data. Compared to single-stage detectors, which prioritise speed, the two-stage Faster R-CNN architecture demonstrated a strong capability in handling the dataset's specific challenges. Its Region Proposal Network was particularly effective at generating high-quality proposals even for small or partially occluded fish, which are common failure points for other methods. While this approach is more computationally intensive than real-time models, the resulting accuracy gains are critical for robust scientific monitoring applications. The success of this approach has important implications for fisheries management, biodiversity conservation, and ecological research, offering a robust tool for automated fish detection and classification. Future efforts could focus on refining the Model's performance for rare species and enhancing its adaptability to diverse underwater conditions, further advancing automated marine monitoring systems.

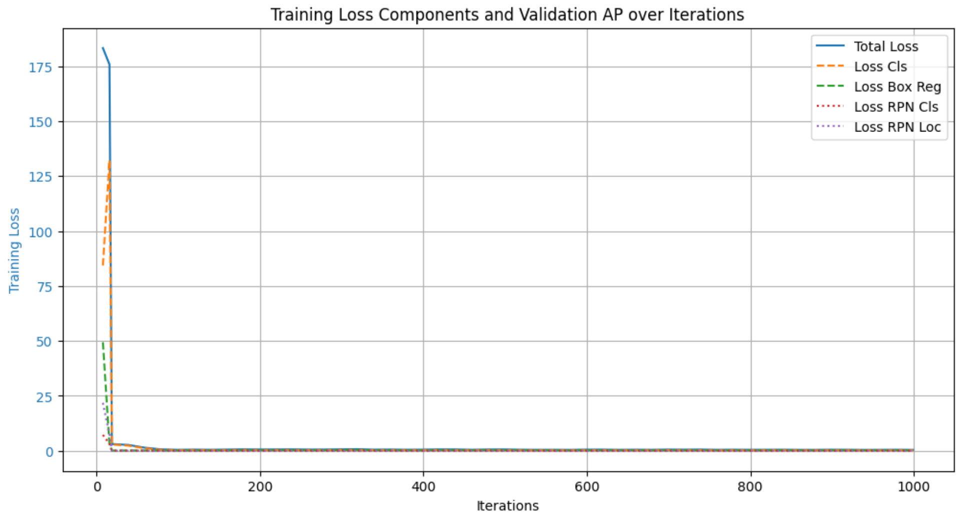

Figure 2 Training Loss of the model

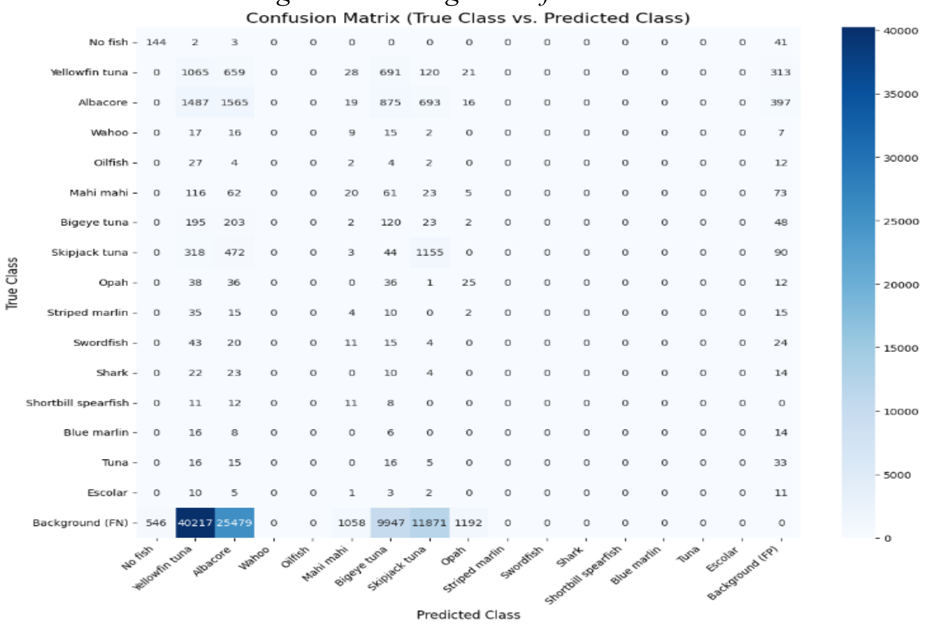

Figure 3: Confusion Matrix of Test Dataset

The training loss graph in Figure 2 shows that all loss components, including those for the Region Proposal Network (RPN) and bounding box regression, converged efficiently. The loss values dropped dramatically within the first 100 iterations and remained stable and low, indicating that the model successfully learned the features present in the training dataset without instability. The model in Figure 3 performs less well in identifying the less common fish. This suggests that the dataset may not have sufficient examples of these fish for the model to learn them properly.

Conclusion

This research demonstrates the effectiveness of the Detectron2 framework, specifically implementing a Faster R-CNN model, for fish detection and classification in challenging underwater environments using the NCFM dataset. By employing a higher input resolution of 832x832 pixels and training for 200 epochs, the Model achieved an F1 score of 86%, showcasing its robustness against challenges such as mislabeled data, low-light conditions, and low-quality images. The results highlight the potential of this robust, two-stage detection approach for real-world marine monitoring applications, offering significant benefits for fisheries management, biodiversity conservation, and ecological research. While the Model occasionally struggled with detecting small or occluded fish and classifying visually similar species, future work could address these limitations through advanced techniques, such as multi-scale training and attention mechanisms. Overall, this study underscores the transformative potential of deep learning and computer vision in revolutionising marine ecosystem monitoring and conservation efforts.

References

1. Salman, A., et al. (2020). Fish Species Classification Using Deep Learning: A Survey. *Journal of Marine Science and Engineering*.
2. Li, X., et al. (2021). Deep Learning for Underwater Image Analysis: Challenges and Opportunities. *IEEE Transactions on Neural Networks and Learning Systems*.
3. G. Sabato, G. Scardino, A. Kushabaha, M. Chirivì, A. Luparelli, and G. Scicchitano, 'Automatic seagrass banquettes detection from surveillance camera images with Detectron2', *Geografia Fisica e Dinamica Quaternaria*, vol. 45, no. 2, pp. 229–235, Dec. 2022, doi: 10.4461/GFDQ.2022.45.11.
4. Thorburn, J., et al. (2019). Applications of Computer Vision in Fisheries Monitoring. *Fisheries Research*.
5. A. Rani, D. Ortiz-Arroyo, and P. Durdevic, 'Defect Detection in Synthetic Fibre Ropes using *Detectron2* Framework', *Applied Ocean Research*, vol. 150, p. 104109, Sep. 2024, doi: 10.1016/j.apor.2024.104109.
6. NCFM (2018). The Nature Conservancy Fisheries Monitoring Dataset.
7. 'facebookresearch/detectron2: Detectron2 is a platform for object detection, segmentation and other visual recognition tasks.' Accessed: Oct. 31, 2025. [Online]. Available: https://github.com/facebookresearch/detectron2