Grenze International Journal of Engineering and Technology, June Issue

Marine Animal and Trash Detection using Faster R-CNN

Anil V Turukmane[1], M. Shailaj Goud[2], D. Baladithya[3] and D. Likhita[4]
[1-4]School of Computer Science and Engineering, VIT-AP University, Inavolu, Amaravati, Vijayawada, AP
Email: anil.turukmane@vitap.ac.in, shailajgoud9908@gmail.com, balubaladithya162@gmail.com, dasarilikhita@gmail.com

***Abstract*—Marine debris has threatened the survival of aquatic ecosystems and has claimed the lives of several marine animals, thus disrupting the balance of nature. It is important to recognize and classify the underwater marine animals and trash in images with effective detection for their conservation and cleaning processes. This paper presents a deep learning-based method for detecting and classifying marine animals and trash from images using Faster R-CNN-a region-based convolutional neural network. It has a ResNet50 backbone with FPN, which is efficient for feature extraction and hence detects small and overlapped objects correctly. The model was trained on a custom dataset of annotated images of marine animals and different kinds of trash, which resulted in high precision and recall metrics. We have built an interactive web-based application using Streamlit that provides users with a visual and intuitive tool for real-time detection. The system showcases leveraging AI to marine conservation and cleaning and hence can also be used for environmental monitoring and research.**

***Index Terms*— Marine Debris Detection, Marine Conservation, Faster R-CNN, Deep Learning, Object Detection, Environmental Monitoring, Marine Ecosystem Protection, Trash Classification, Marine Animal Identification.**

I. ABBREVATIONS

R-CNN (Region-based Convolutional Neural Network), FPN (Future Pyramid Network), mAP (Mean Average Precision), loU (Intersection over Union), RPN (Region Proposal Network), GPU (Graphics Processing Unit), ROI (Region of Interest), API (Application Programming Interface).

II. INTRODUCTION

There are severe threats of accumulated marine debris in oceans and other aquatic bodies on biodiversity, ecological balance, and economic activities. Plastics, fishing gear, and other waste materials in this regard are important contributors to pollution, which highly endangers the life of marine animals and spoils their habitat. The detection of this challenge, identification, and thereby solving need efficient and automated solutions for finding and classifying marine debris. Apart from those, the conventional methods of monitoring marine debris included field observation and other classic computer vision technique-based approaches that were generally time-consuming and cost a lot while having their scope of applicability. More recently, object detection and classification enabled by progress in the AI domain have been considerably improved to achieve full automation, particularly the deep learning sub-area. The Faster R-CNN model ranked among the most accurate models in

Grenze ID: 01.GIJET.11.2.541

object detection for some time and was spread due to that into numerous applications covering various tasks on object detection. This study would be on the marine animal and trash image detection and classification using Faster R-CNN, basically for environmental conservation and sustainability. By tuning its pre-trained model of Faster R-CNN through a custom dataset, the system is bound to classify the different marine animals and kinds of trash even if the underwater environment may have poor visibility or overlapping.

We further developed a more friendly and accessible allows image uploading and seeing instantaneously the predictions of the model. This system will be useful to support researchers, environmentalists, and policy makers to identify critical areas affected by marine debris and hence prioritize cleanup operations. The rest of the paper is organized as follows: Section 2 reviews related work in the area, Section 3 describes the methodology and system design, Section 4 presents the results and analysis, and Section 5 discusses implications, limitations, and future directions for this research.

III. Literature Survey

The problem of marine debris and its impacts on ecosystems have been gaining wide attention in these years, and this has motivated a number of research efforts in the detection and monitoring of marine debris. This section reviews some key developments in the fields of marine debris detection, object detection techniques, and deep learning models. Traditional methods for identifying marine debris consist of surveys that are usually aided by aerial imagery or underwater imagery collection. While tremendously useful, quite often these usually involve a manifold of labor and provide limited scale capacity. The addition of satellite data for imaging and super-high-resolution photography has definitely been advanced accordingly, but major challenges lie on the processing aspect and analysis performance of such imagery.

Object detection has come a long way from the conventional methods in computer vision to modern methods incorporating machine learning. Feature-based early approaches like Haar cascades and the HOG feature extraction method require many computational operations and fail miserably under robust environments such as an underwater setting.While the discovery of deep learning facilitated Convolutional Neural Networks in drastically changing object detection, as feature extraction became automated. Region-based CNN introduced by Girshick et al. formed a new era of detection performance improvement of objects from images. Then Faster R-CNN with the introduction of Region Proposal Network was designed for proposing regions, doing this job faster without losing accuracy in real-time performance. Faster R-CNN has gained wider acceptance in environmental research for its robustness in the detection of small and overlapping objects. This was evidenced in some studies, like wildlife interactive web-based application using Streamlit that

model could process a wide variety of datasets on such complex tasks as distinguishing between marine animals and debris in underwater images. The underwater environment has its own challenges regarding low visibility, light scattering, and other types of marine species and objects. In these conditions, traditional object detection methods do not work well; rather, advanced models like Faster R-CNN are required. Scientists have overcome such challenges by fine-tuning the pre-trained models on domain-specific datasets and data augmentation techniques for enhancing their robustness.

There, notwithstanding a significant step-up in progress towards object detection; just fewer research works target marine animals along with trash or wastes. Most literature studies either addressed ocean and/or debris detection solely, or concentrated their efforts mainly on marine bio-diversity observation. Alongside, scarce in technique allows quite easy applications that fit real life and can potentially play at conservational strategy on a daily basis.

IV. Dataset Description

The dataset used in this work was specifically curated to support the detection and classification of marine animals and trash. It includes high-resolution images from diverse underwater scenarios to ensure robustness for a range of conditions, including low visibility, diverse lighting, and object overlap. The two main classes in the dataset are marine animals and trash, while there is also a background class for objects that do not belong to these categories. The images were garnered from open-sourced materials and adapted in this study through ROIs annotations based on bounding boxes. Therefore, each one of the annotations presents details on the particular object location about the image's class label and marine animal-trash-confidence levels. The dataset is then balanced, not over representing either class due to mixing those images to avoid biases within training. Every image in the dataset is manually labeled with bounding boxes that provide object locations and their class labels, such as "Marine Animal" or "Trash." These will be important in training the Faster R-CNN model since they

provide the ground truth needed for supervised learning. To enhance the model's ability to generalize across different environments, the dataset includes images that depict challenging situations, such as overlapping objects, cluttered backgrounds, and objects monitoring, deforestation, and identification of plastic waste. The
as rotations, flips, and brightness adjustments, were employed to further increase the variety and diversity of the training data.

This dataset provides the bedrock upon which this efficient and robust detection system was developed to enable the Faster R-CNN model to appropriately discern the bounding boxes and classify objects into various object categories across different real-world underwater applications. The quality and variety of imagery in this dataset ensure that the network is resilient-a characteristic which most presumably promises extraordinary performance in a variety of other underwater scenarios which could become challenging.

V. METHODOLOGY

The paper proposes a method based on Faster R-CNN deep learning with a proposed classification for the detection of marine animals and trash; this includes a sequence of model choice, dataset creation, training, and deployment-all guaranteeing an accurate and efficient object detection system.

A. Model Selection

Faster R-CNN was selected because it is robust for object detection and handles various complicating factors in scenarios involving overlapping objects or small object detection. This is powered by using ResNet50 backbone with a Feature Pyramid Network for feature extraction-a suitable balance between accuracy and computational efficiency. The Region Proposal Network of the model generates a set of proposed bounding boxes, which is then refined by the RoI pooling layers for object detection and classification.

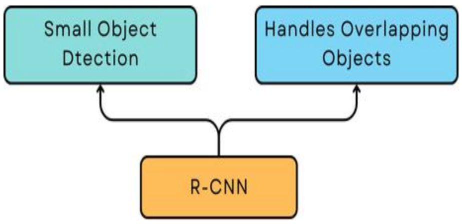


Fig. 1. Diagram of R-CNN's features

B. Dataset Preparation

The dataset is a collection of high-resolution underwater images annotated with bounding boxes and corresponding class labels, including Marine Animal and Trash. To improve the generalization capability of the model across diverse environments, data augmentation techniques, such as random rotations, flips, brightness adjustments, and scaling, have been partially obscured by underwater particles or plants. Additionally, data augmentation techniques, such performed. This enhanced the diversity of the dataset, which needed to address challenges such as those presented by the underwater conditions and occlusion.

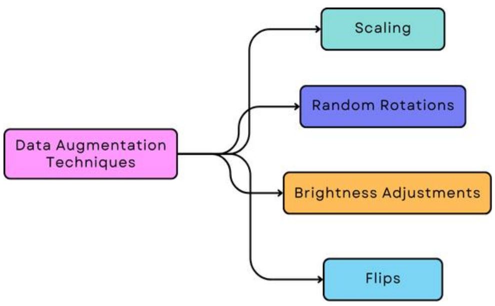


Fig. 2. Data Augmentation Techniques

C. Model Training

The Faster R-CNN model was trained by the method of supervised learning on the prepared dataset. The loss to be minimized during training will include a classification loss for identifying the object class and regression loss to refine bounding box predictions. It fine-tunes a pre-trained ResNet50 network, reusing a previously trained model by transfer learning that enables faster convergence and higher accuracy. Other hyper parameters: Learning rate and batch size were optimized by experimentation to arrive at the best results.

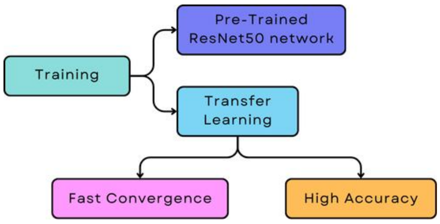


Fig. 3. Flow Diagram of Training using ResNet50 network

D. Inference and Visualization

During inference, input images are preprocessed and run through the trained model to output bounding box coordinates, class labels, and confidence scores. To get a much better view, these results were visualized using a Python pipeline, which was constructed using OpenCV by overlaying bounding boxes and class labels onto the original image. Class-specific colors along with annotations were used for intuitive interpretation.

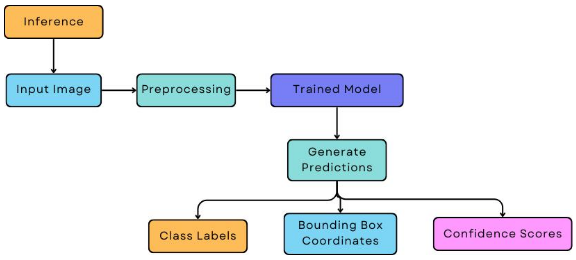


Fig.4. Flow Diagram of Inference that goes under preprocessing

E. Deployment

A web application with logic built on Streamlit for the system, which should be accessible and user-friendly; a user is able to upload images there, and for these images, real-time predictions will appear. That model receives the uploaded images witted after some preprocessing and makes inferences using an already trained Faster R-CNN. It shows the result-an image with annotations and object detection details-outside the application.

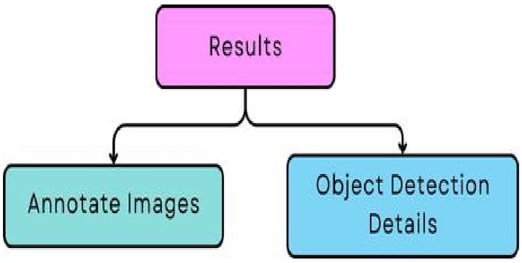


Fig. 5. Flow Diagram of Results

VI. Modules Used

In general, the following Python modules and libraries have been used throughout the development process of the marine animal and trash detection system. These modules were selected based on their compatibility, functionality, and efficiency in the execution of key tasks that involve deep learning, image processing, and web application development.

A. PyTorch

PyTorch is one of the most liked deep learning frameworks; hence, it is used here for developing and training the model. Its dynamically computed computation graph and extensive support for pre-trained models provide an ideal environment for the implementation of Faster R-CNN. The PyTorchtorchvision module provided the pre-trained models of Faster R-CNN and tools for image transformations.

B. Torchvision

Torchvision was used for loading the Faster R-CNN model and applying the image transformations necessary for the preprocessing. In addition, it includes utilities to be used with datasets to manipulate data to allow easy interaction with the deep learning pipeline.

C. NumPy

Meanwhile, numerical computations like processing the output of the model, handling bounding box coordinates, and array-based operations were done using NumPy. Its efficiency in handling multi-dimensional arrays served very well in handling image data.

D. OpenCV

OpenCV is an open-source library of computer vision, which will be used for processing and visualizing images in this work. It would be able to draw the bounding box, label overlays on top of images, and color space conversions to show the detection results in a more fashionable way.

E. Streamlit

Streamlit has been used for developing a web application user-friendly, capable of uploading an image and showing detection results really quick. Simplicity and quick generation of interactive web interfaces, this was very ideal to deploy.

F. PIL (Pillow)

Pillow was employed for handling image input and conversion. It facilitated loading images in different formats, resizing them, and converting them into the format required for model inference.

G. Matplotlib

Matplotlib was potentially used for generating visualizations during the development phase, such as plotting detection results and model performance metrics.

VII. IMPLEMENTATION

In developing the marine animal and trash detection system, one went through setting up the model and creating a user-friendly web application from scratch. Thus, object detection was done via Faster R-CNN, comprising a workstream with respect to inputting images, prediction, and presentation.

A. Model Development

This model utilizes the Faster R-CNN model, pre-trained on COCO, using a ResNet50 backbone. Fine-tuning the model has been performed in order to classify the model into two specified classes: "Marine Animal" and "Trash." In this process, PyTorch was used, along with Torchvision, for loading the pre-trained model and adjusting the classification head in accordance with custom classes. The model was optimized in its RoI pooling layer and RPN to detect small and overlapping objects in underwater images. Transfer learning was applied to leverage the pre-trained weights, reducing computational resources and time required for training.

B. Dataset Preparation

A high-resolution underwater image, therefore, constitutes the dataset to which the ground truth, in the form of bounding box and class label information, is applied. These bounding box and class labels were to serve as supervision while training. Data augmentation including rotation, flip, changing brightness, and scaling was considered for better generalization and robustness of the model.

C. Training Process

The model was trained by using a multi-class loss function, incorporating the losses from classification and regression for bounding box predictions. Training includes:

- Hyperparameter Optimization: The experiments are done on the learning rates, batch sizes, and the number of epochs for finding the best performance.
- Validation: Checking the model's performance on the validation data from time to time to keep a check on overfitting and for tuning hyperparameters.
- Hardware Utilization: Training was done on our GPUs, allowing faster computation via parallel processing.

D. Inference Pipeline

During inference, input images were pre-processed using Torchvision's transformation utilities to match the model's input requirements. The pre-processed images were fed into the trained model to generate:

- Bounding box coordinates of the detected object.
- Class labels indicating whether "Marine Animal" or "Trash.
- Confidence scores for every prediction.

Therefore, post-processing was performed on the output in order to filter predictions according to a certain threshold of confidence, say 50%.

E. Visualization and Interpretation

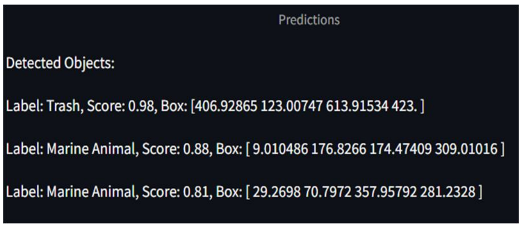
These detections were represented in OpenCV with bounding boxes on detected objects in class-specific color coding-orange for marine animals and yellow for trash, along with their confidence score labelled on the detected box. Such intuitive visualization easily interpreted the results to the non-technical users.

F. Testing and Evaluation

The final system has been tested on unseen data to check its performance for the task in object detection and classifying different marine animals and trash. Then, precision, recall, and mAP measures are calculated to measure the accuracy and robustness of the model.

Fig. 6. Input Image of a marine animal along with trash

Fig. 7. Visualization and Interpretation of the Input Image



Predictions

Detected Objects:

Label: Trash, Score: 0.98, Box: [406.92865 123.00747 613.91534 423.]

Label: Marine Animal, Score: 0.88, Box: [9.010486 176.8266 174.47409 309.01016]

Label: Marine Animal, Score: 0.81, Box: [29.2698 70.7972 357.95792 281.2328]

Fig. 8. Predictions of the Processed image along with readings

VIII. RESULTS

The results obtained from the marine animal and trash detection system proved the efficiency of the Faster R-

CNN model in recognizing and classifying objects in an underwater environment with good accuracy. The performance was measured in terms of precision, recall, and mean Average Precision; it turned out to be very promising for all scenarios.

A. Model Performance

The accuracy of the trained Faster R-CNN model for marine animal and trash detection was very good. The measures of evaluation are as follows:

- Precision: It detected the objects with a high level of accuracy, minimal false positives; hence, objects of dependably classifying marine animals from trash were detected.
- Recall: the system was able to detect most of the objects of interest, even in poor visibility or overlapping conditions.
- Mean Average Precision (mAP): The mAP score reflected the overall effectiveness of the model in localizing and classifying objects accurately.

B. Qualitative Analysis

Then, detection results were visually inspected for model robustness and interpretability. The system was able to identify small objects, such as debris pieces, and correctly differentiated marine animals from trash in cluttered underwater scenes. The bounding boxes were well-aligned with the objects, and confidence scores provided users with clear indications of detection reliability. Color-coded visualization-for example, orange for marine animals and yellow for trash-enhanced the interpretability of the results.

IX. CONCLUSION

These results have confirmed that Faster R-CNN works for underwater object detection, wherein the system could achieve high accuracy and usability. The identified challenges further create room for improvements by increasing training data, enhancing the techniques for image preprocessing, and fine-tuning of the model on specific underwater conditions. In general, the system has a great potential to contribute to marine conservation by automating debris detection in water bodies.

REFERENCES

[1] Faisal, Muhammad, SushovanChaudhury, K. SakthidasanSankaran, S. Raghavendra, R. Jothi Chitra, Malathi Eswaran, and Rajasekhar Boddu. "Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation." Mathematical Problems in Engineering 2022, no. 1 (2022): 3639222.

[2] Faisal, Muhammad, SushovanChaudhury, K. SakthidasanSankaran, S. Raghavendra, R. Jothi Chitra, Malathi Eswaran, and Rajasekhar Boddu. "Research Article Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation." (2022).

[3] Jain, Ritik, SarikaZaware, NirajKacholia, Het Bhalala, and OamJagtap. "Advancing Underwater Trash Detection: Harnessing Mask R-CNN, YOLOv8, EfficientDet-D0 and YOLACT." In 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS), pp. 1314-1325. IEEE, 2024.

[4] Roshni, Nishat Mahmud, MeshkatulArefin, KaziMasfiqulAlam Joy, Marjanul Hassan, and Shashwata Karmakar. "Trash detection in an aquatic environment." PhD diss., Brac University, 2024.

[5] Fulton, Michael, Jungseok Hong, MdJahidul Islam, and JunaedSattar. "Robotic detection of marine litter using deep visual detection models." In 2019 international conference on robotics and automation (ICRA), pp. 5752-5758. IEEE, 2019.

[6] Mathangi, Mounica. "Detection of underwater trash objectsusing deep learning algorithms." (2023).

[7] Wahlig, Victoria, and Ricardo A. Gonzales. "Enhancing marine debris identification with convolutional neural networks." Journal of Emerging Investigators 7 (2024).

[8] Ren, Chengjuan, Hyunjun Jung, Sukhoon Lee, and Dongwon Jeong. "Coastal waste detection based on deep convolutional neural networks." Sensors 21, no. 21 (2021): 7269.

[9] Hong, Jungseok, Michael Fulton, and JunaedSattar. "Trashcan: A semantically-segmented dataset towards visual detection of marine debris." arXiv preprint arXiv:2007.08097 (2020).

[10] Bai, Jing, and Shujie Yang. "Detection Method of Marine Floating Garbage based on Improved Faster R-CNN." Highlights in Science, Engineering and Technology 15 (2022): 120-127.

[11] Walia, Jaskaran Singh, Kesar Mehta, Shivram Harshavardhana, and Nandini Tyagi. "Deep learning Innovations for Underwater Waste Detection: An In-Depth Analysis." arXiv preprint arXiv:2405.18299 (2024).

[12] Dias, Kimbrel, Sadaf Ansari, and Ameeta Amonkar. "Marine Trash Detection Using Deep Learning Models." In Computing and Communications Engineering in Real-Time Application Development, pp. 231-242. Apple Academic Press, 2022.

[13] Corrigan, Brendan Chongzhi, Zhi Yung Tay, and Dimitrios Konovessis. "Real-time instance segmentation for detection of underwater litter as a plastic source." Journal of Marine Science and Engineering 11, no. 8 (2023): 1532.