DEEP LEARNING-BASED MICROPLASTIC DETECTION AND SEGMENTATION

Abhin P. T., Omal S.*, Abhishek Mohan K., Abhishek P., Ritto Thomas, Remya K. Sasi

Department of Computer Science, Cochin University of Science and Technology, Kochi, 682022, Kerala, India

Email: omammu311@gmail.com*

ABSTRACT

Microplastics are an increasingly serious threat to aquatic ecosystems, affecting human health and marine biodiversity through bioaccumulation and contamination in water supplies. These microscopic plastic particles, which are the result of the breakdown of larger plastic debris released from wastewater and industrial discharge, enter aquatic environments and are ingested by organisms that accumulate toxic chemicals higher up the food chain. Conventional methods for measuring and detecting microplastics are costly, time-consuming, and not scalable for widespread monitoring, necessitating the development of innovative and automated solutions. This paper presents a novel deep learning-based method to overcome the limitations mentioned above by using state-of-the-art models like Faster R-CNN, U-Net, ResNet-50, and YOLOv8, which were trained on a carefully chosen dataset of microplastic images. While YOLOv8 offers better real-time detection with a mean Average Precision (mAP) of 78.4% at IoU=0.50 and a fast inference time of 0.7 milliseconds per image, U-Net excels at accurate pixel-level identification with the highest segmentation accuracy of 96.03%. To effectively identify and classify microplastics across diverse environmental conditions, a multi-model approach provides both flexibility and thoroughness. Each model brings unique strengths to the task—Faster R-CNN excels at detecting medium-sized microplastic particles, while ResNet-50 offers robust classification capabilities, ensuring accurate differentiation between plastic types and other debris. By combining these models, researchers can achieve more reliable and comprehensive results, adapting to varying particle sizes and complex backgrounds often encountered in real-world aquatic environments. This integrated method enhances detection accuracy, making it a practical solution for large-scale environmental monitoring efforts. This breakthrough supports conservation efforts and tackles the urgent environmental and public health issues caused by microplastic pollution, opening the door for more effective policies and interventions to protect aquatic life and human well-being. This AI-powered framework offers a scalable, high-precision alternative for tracking microplastic pollution, moving beyond outdated manual methods. By integrating cutting-edge deep learning techniques, researchers and environmental agencies can conduct large-scale, real-time assessments of aquatic ecosystems.

Keywords: *Microplastic Detection, Deep Learning, U-Net, YOLOv8, Environmental Monitoring, Aquatic Ecosystems, Image Segmentation.*

1. INTRODUCTION

Microplastics, which are smaller plastic particles less than 5 mm in diameter, are widespread pollutants that enter aquatic ecosystems and cause marine biodiversity and severe threats (Galloway & Lewis, 2016) to human health through invasion of food chains and water supplies. Studies have revealed that marine organisms ingest microplastics and act as carriers for toxic chemical pollutants, magnifying their ecological and toxicological effects (Wright et al., 2013). The rising levels of microplastics in the environment are a significant concern globally as they are a result of the breakdown of larger plastic waste, as well as direct incorporation into consumer products.

It is tough to deal with the microplastic problem because conventional detection technologies fall short. Such methods as spectroscopy and visual examination in a microscope are labor-consuming, costly, and not scalable regarding the surveillance of large water surfaces (Walkinshaw et al., 2022). Additionally, these methods are also limited by the need for specific apparatus and highly qualified staff, which makes them unapplicable to full-fledged environmental evaluation.

Recent developments in machine learning and computer vision have enabled the application of computer vision for environmental problems, e.g., detecting microplastics. This paper employs cutting-edge deep learning models to introduce a scalable and automated scheme for microplastic detection and segmentation. Microplastic contamination has emerged as one of our most urgent environmental issues, and scientists are now looking to artificial intelligence for improved solutions. Deep learning, specifically convolutional neural networks (CNNs), has changed the way we process images - from detecting tumours in medical images to locating objects in satellite images (LeCun et al., 2015). We are testing some of these strong models here - Faster R-CNN, U-Net, ResNet-50, and YOLOv8 - each of which has its strengths for identifying microplastics with accuracy and efficiency.

The work is based on the marked-up microplastic images. These are not simple images - every one includes detailed annotations such as bounding boxes and segmentation masks that instruct our AI algorithms what to search for. This is a rich dataset that enables the models to learn how to identify microplastics in all manner of real-world environments, from dirty river water to ocean surfaces. What we are creating is not a mere research endeavour - it is a working tool that environmental authorities can utilize to monitor water quality in real time, taking us far beyond the high-cost means of the past.

Our objective is to develop an intelligent system capable of automatically monitoring microplastic pollution at scale. The conventional methods just cannot cope - they are too labour-intensive, too expensive, and too narrow in scope. By integrating the newest AI models, we are building something unique: a system that operates fast, accurately, and can handle everything from lab samples to open ocean environments. Faster R-CNN assists us in finding particles precisely, U-Net is best at defining their exact contours, ResNet-50 identifies them accurately, and YOLOv8 provides super-fast detection when speed is most critical.

Architectures such as Faster R-CNN have revolutionized the way computers interpret images since their advent (Ren et al., 2017), and U-Net's ingenious encoder-decoder architecture has earned it a place in medical imaging where each pixel counts (Ronneberger et al., 2015). Today, we are using these very developments to safeguard our waterways.

What is so exciting about this is how it unites bleeding-edge AI with pressing environmental imperatives. We are developing tools that potentially transform the way we track and fight microplastic pollution. By training these models on broad, thoughtfully annotated datasets, we are constructing systems that environmental teams might use in the real world. It is about converting AI research into real-world results, assisting in the preservation of marine ecosystems, and ultimately safeguarding human health from this imperceptible menace. The latest version of the YOLO series, YOLOv8, is ideal for applications that require both high accuracy and fast processing time since it achieves the balance between inference speed and detection accuracy. YOLOv8 is used mostly for real-time uses such as tracking environmental contamination or monitoring marine litter because it can have high mean Average Precision (mAP) scores at low latency. Image processing without compromising detection capability makes YOLOv8 a possible candidate for large-scale monitoring of aquatic ecosystem microplastic pollution (Yaseen, 2024).

Even though deep learning based segmentation and object detection methods have improved considerably, there is limited understanding of applying these models to microplastic detection. Microplastics are becoming an increasing issue in aquatic environments. Conventional detection techniques, including spectroscopy and microscopy, demand significant time and effort, particularly regarding their capacity for real-time estimation and management of large datasets. Hence, there is significant potential to solve these issues by scaling and automating the detection of microplastics based on deep learning algorithms. This study will address this research gap through a comparative evaluation and assessment of the performance of Faster R-CNN, U-Net, and YOLOv8 on a carefully curated dataset of microplastics. For setting the foundation of developing automated environmental monitoring devices, performance standards for the detection of microplastics and segmentation processes are to be determined.

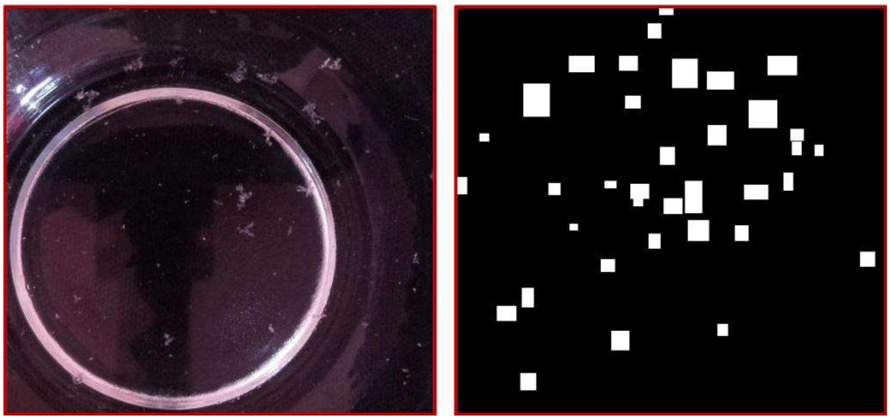

Fig. 1 Sample of Microplastic dataset for computer vision and corresponding binary mask.

2. MATERIALS AND METHODS

Annotated images of microplastics collected from marine environments and annotated with pixel-level segmentation masks and bounding boxes constitute the dataset employed in this work. The images were converted into the widely used object detection task format called COCO (Common Objects in Context). The region of interest (ROIs) here refers to the microplastic particles, which are marked by the bounding boxes in the images. These ROIs were then converted to binary segmentation masks usable in pixel-level segmentation operations (Momeni, n.d.).

Bounding boxes were transformed into binary masks per image by using white pixels (255) to fill the interior of the box and black (0) for the background. This transformation relies on compatibility with segmentation models such as the U-Net, which require pixel-wise labels to be trained effectively. The data set was split into training (70%), validation (20%), and testing (10%) to ensure the models would generalize and not overfit. This was in adherence to standard operating procedures. The training and validation sets were utilized to fine-tune the model, while the test set was reserved to assess the performance of the final model. Various data augmentation methods, such as rotation, flipping, and colour jittering, were applied to enhance the diversity of the dataset and enhance the model's robustness.

Four deep learning models mentioned in this study are selected based on their unique strengths in the task of microplastic segmentation and detection. To determine if these models can detect and segment microplastic particles in water bodies, the prepared microplastic dataset was utilized to train and improve the Faster R-CNN, U-Net, ResNet-50, and YOLOv8 models.

a) Faster R-CNN: A widely used model for object detection tasks, Faster R-CNN has great accuracy and efficiency. It utilizes a Fast R-CNN detector to classify the candidate object regions generated by a Region Proposal Network (RPN). The model has a batch size of two, using a pre-trained ResNet-50 backbone on the

COCO dataset and fine-tuned on the microplastic dataset for 2000 iterations. Precision in region proposal generation was the primary aim so that microplastics could be found in the images. For object detection in challenging environmental conditions, Faster R-CNN is ideal since it is capable of generating region proposals and classifying them within one framework (Ren et al., 2017).

b) ResNet-50: A residual deep network by the name ResNet-50 is renowned for its capability to train highly complex architectures without exposing itself to vanishing gradients. To improve performance, the task of segmentation was fine-tuned, focusing on the identification of key features in the input images. We can take advantage of the power of ResNet-50 in microplastic detection here, which is famous for its performance on big datasets (He et al., 2016).

c) U-Net: It is famous for its performance in pixel-level segmentation tasks. It was designed specifically for biomedical image segmentation. It uses a convolutional encoder-decoder structure with skip connections to get high-level features without sacrificing spatial resolution. In this paper, we used binary cross-entropy loss and Adam optimizer to train U-Net for 25 iterations with a batch size of 8. The encoder downsamples the input image to get contextual information, and the decoder reconstructs the image pixel by pixel. U-Net is specifically best suited to segment small objects, such as microplastics, in complex images because of this architecture. Two such areas where the application of U-Net has been effective are medical image segmentation and environmental monitoring (Ronneberger et al., 2015).

d) YOLOv8: Trained to perform object detection in real time, YOLOv8 is the latest model from the YOLO (You Only Look Once) family. YOLOv8 was trained for 10 epochs at an input image resolution of 640 x 640 pixels and an optimized batch size of 16. The model was specifically selected due to its fast inference time and high mean Average Precision (mAP) score. YOLOv8's efficient real-time object detection with low latency is needed for high-rate monitoring tasks such as detecting microplastics in water bodies. It provides the optimal speed-accuracy trade-off for continuous environmental monitoring (Yaseen, 2024).

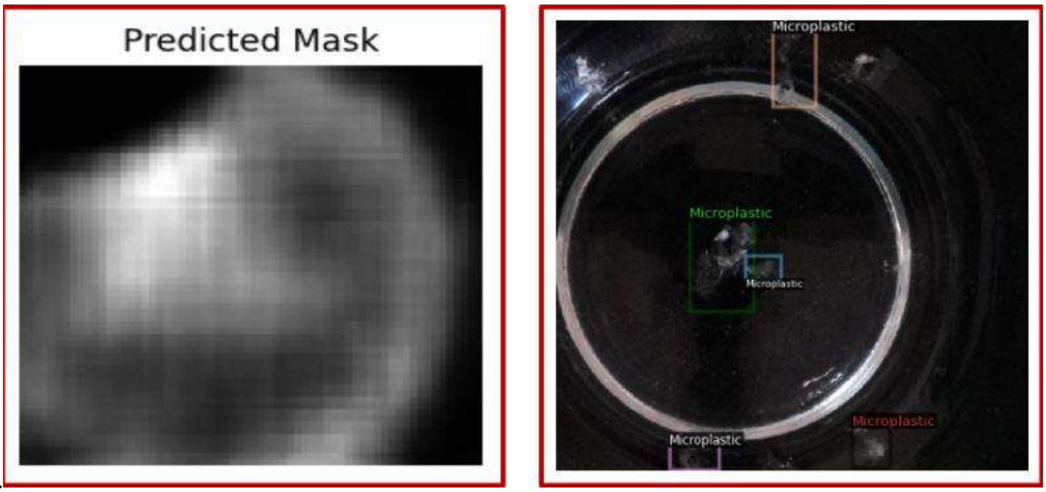

Fig. 2 Comparison of Faster R-CNN and ResNet-50 performance metrics for microplastic detection.

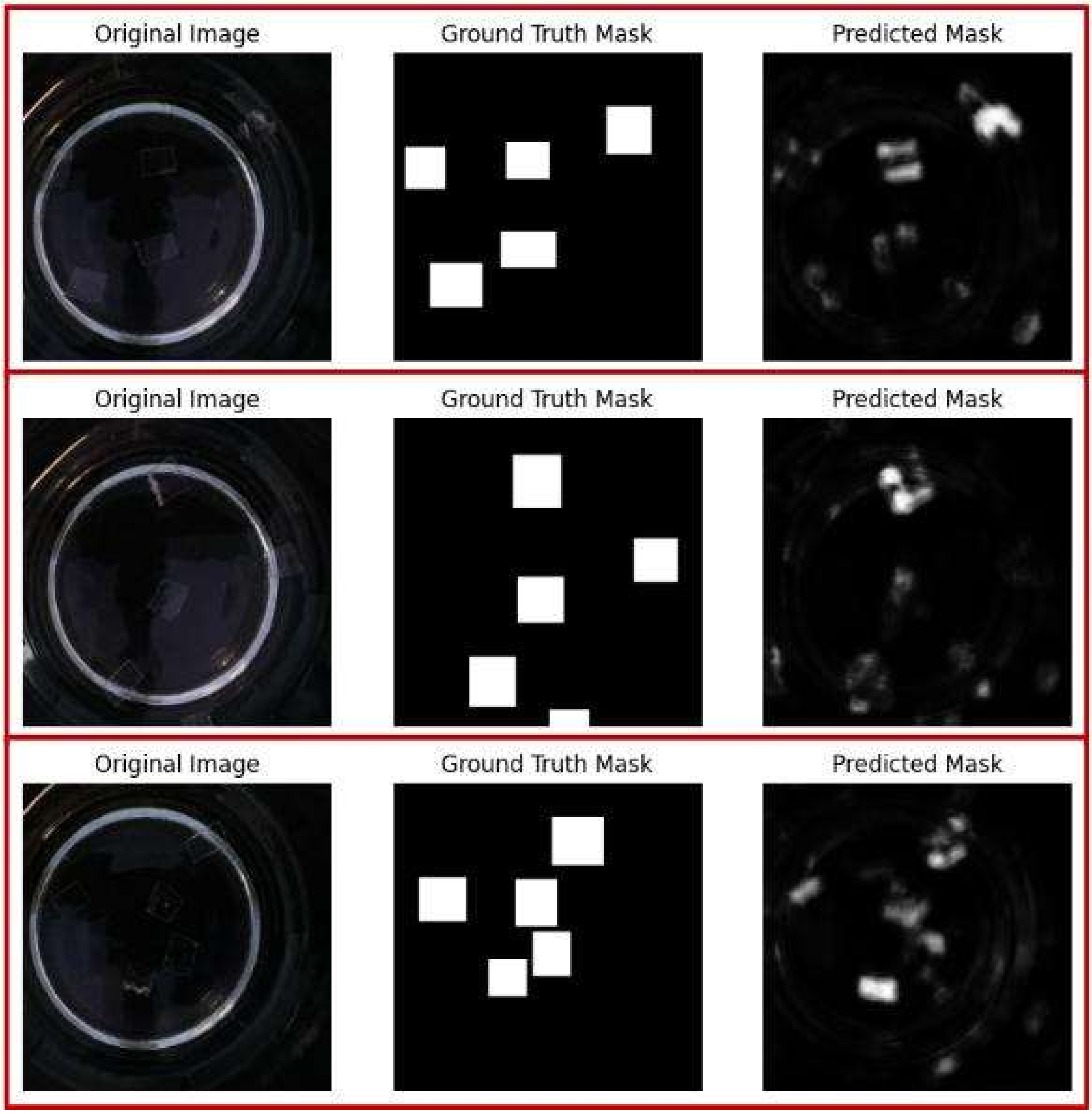

Fig. 3 UNet Predictions

Different optimization algorithms tailored for each architecture were utilized to train the models. We used a stochastic gradient descent (SGD) optimizer with momentum and weight decay for both Faster R-CNN and YOLOv8, which helped to boost convergence without overfitting. In both U-Net and ResNet-50, we used the Adam optimizer because of its learning rate adaptability and better training capabilities of deep networks. Early stopping was applied in all cases to prevent overfitting and select the best model based on validation set performance.

All models are trained on GPUs for faster and better training. To determine the optimal combination for each model, hyperparameter tuning was performed through a grid search over learning rates, batch sizes, and regularization terms. To ensure the results were robust and not biased by a certain train-test split, cross-validation was performed. To provide a comprehensive measurement of the performance of each model in microplastic identification, the performance of the models was finally measured using a range of metrics, including segmentation accuracy, Average Precision (AP), and inference time.

3. RESULTS AND DISCUSSION

The performance of the evaluated models indicated remarkable improvements in microplastics detection and segmentation, overcoming key limitations of traditional approaches. The U-Net proved effective for complex, pixel-level segmentation tasks by reflecting the maximum segmentation accuracy of 96.03%. Even under such unfavourable conditions, its biomedical-applications-oriented architecture proved to be proficient in the detection of microplastic particles. Through its region proposal network, Faster R-CNN detected medium-sized aggregations of microplastics at an Average Precision (AP) of 30.5% (IoU=0.50:0.95). While the same performance for ResNet-50 when normalized for segmentation was achieved, its high computational requirements made it fail in performing optimally during real-time inference.

Table 1 Result Comparisons for Faster R-CNN, UNet, ResNet-50, and YOLOv8

Model	Task	AP (IoU=0.50: 0.95)	AP50	AP75	Validation Loss	Validation Accuracy
Faster R-CNN	Object Detection	30.47%	68.37%	23.37%	-	-
U-Net	Semantic Segmentation	-	-	-	0.1161	96.03%
ResNet-50	Classification	-	-	-	0.2091	94.02%
YOLOv8	Object Detection	37.6%	78.4%	32.1%	-	-

With a mean Average Precision (mAP) of 78.4% and an inference time of 0.7 milliseconds per image, YOLOv8 was the top-performing model to be used in real-time applications. This means that it can potentially be utilized for large-scale environmental monitoring where rapid detection is needed.

4. CONCLUSION

A robust basis for the automatic identification and segmentation of microplastics with deep learning models was effectively developed and tested by the project. The results show the compromises between models: While YOLOv8 offers unparalleled speed and efficiency, thereby suitable for real-time use in large-scale monitoring applications, U-Net offers high accuracy and precise segmentation, thus best suited for applications needing precise analysis. The project provides an expansive solution to the management of microplastic pollution through the integration of these models within an automated system. Besides addressing the disadvantages of traditional methods, the framework provides decision-makers and scientists with tools for monitoring the environment beforehand. The findings outline the importance of selecting models based on specific application requirements, paving the way for efforts in marine ecosystem preservation, public health protection, and conformity with global sustainability goals.

ACKNOWLEDGMENT

We would like to extend our deepest gratitude to Dr. Remya K. Sasi, our supervisor, for her support, assistance, and encouragement throughout this research work. We would also like to thank our family and friends, whose constant encouragement has been our source of inspiration throughout this course. Lastly, we appreciate our classmates for their teamwork, which made our vision clearer and our determination to be innovative and serious in doing the project stronger. We thank all the people who helped in the creation of the project for a learning and successful experience.

REFERENCES

Galloway, T. S., & Lewis, C. N. (2016). Marine microplastics spell big problems for future generations. *Proceedings of the National Academy of Sciences*, *113*(9), 2331–2333. https://doi.org/10.1073/pnas.1600715113

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. Retrieved from https://doi.org/10.1109/cvpr.2016.90

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Momeni, M. (n.d.). Microplastic dataset for computer vision. Retrieved May 12, 2025, from Kaggle website: https://www.kaggle.com/datasets/imtkaggleteam/microplastic-dataset-for-computer-vision

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/tpami.2016.2577031

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science* (pp. 234–241). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-319-24574-4_28

Walkinshaw, C., Tolhurst, T. J., Lindeque, P. K., Thompson, R., & Cole, M. (2022). Detection and characterisation of microplastics and microfibres in fishmeal and soybean meal. *Marine Pollution Bulletin*, *185*, 114189. https://doi.org/10.1016/j.marpolbul.2022.114189

Wright, S. L., Thompson, R. C., & Galloway, T. S. (2013). The physical impacts of microplastics on marine organisms: A review. *Environmental Pollution*, *178*, 483–492. https://doi.org/10.1016/j.envpol.2013.02.031

Yaseen, M. (2024, August 28). What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector. Retrieved May 12, 2025, from arXiv.org website: https://arxiv.org/abs/2408.15857