International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

Robust Multi-Class Underwater Waste Detection via CNN and Transformer-Based Object Detection Models

Dr.P Karpagavalli, Abijith K S, Aswin babu K S, Bharathkumar M

KLN College of Engineering, Sivagangai, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: Marine pollution caused by underwater waste has become a growing environmental concern that threatens aquatic ecosystems and marine biodiversity. Detecting submerged debris in underwater environments remains a difficult task because of challenges such as low visibility, colour distortion, light absorption, and complex backgrounds. Recent advances in deep learning have made it possible to automatically detect underwater objects from visual data with improved accuracy and efficiency.

This study investigates the effectiveness of deep learning–based object detection models for multi-class underwater waste detection. A dataset containing fifteen categories of underwater waste objects was obtained from the Roboflow platform. To address common underwater imaging problems, several preprocessing techniques were applied, including red channel enhancement, white balance correction, gamma correction, contrast limited adaptive histogram equalization (CLAHE), bilateral filtering, and image normalization. Three object detection models—YOLOv8, YOLOv9, and RT-DETR—were trained and evaluated using standard detection metrics such as precision, recall, and mean Average Precision (mAP).

Experimental results show that all three models provide reliable detection performance for underwater waste objects. Among them, the RT-DETR model demonstrates higher recall while maintaining comparable precision and mAP values, indicating improved detection coverage in complex underwater scenes. These findings suggest that transformer-based detection models can offer advantages for underwater waste monitoring applications and may support the development of automated systems for marine pollution assessment and environmental conservation.

KEYWORDS: Underwater waste detection, deep learning, object detection, YOLOv8, YOLOv9, RT-DETR.

I. INTRODUCTION

Marine pollution has emerged as a significant environmental concern, with large amounts of waste such as plastics, fishing nets, and metal debris accumulating in underwater ecosystems. These pollutants persist for long periods and pose serious threats to marine life by causing entanglement, ingestion, and habitat destruction. Additionally, the breakdown of such waste contributes to microplastic formation, further impacting the marine food chain. Consequently, effective monitoring and detection of underwater debris have become essential for environmental conservation and sustainable ocean management.

However, detecting underwater waste remains a challenging task due to adverse imaging conditions. Factors such as low light, color distortion, turbidity, and scattering significantly degrade image quality, making object identification difficult. Underwater images often appear bluish or greenish due to the absorption of red wavelengths, while suspended particles reduce visibility and contrast. Traditional detection methods relying on manual inspection are time-consuming, labor-intensive, and limited in coverage, highlighting the need for automated and scalable solutions.

Recent advancements in deep learning and computer vision have enabled the development of automated object detection systems capable of analyzing complex underwater scenes. Convolutional neural network–based models such as YOLOv8 and YOLOv9 have demonstrated strong performance in real-time detection tasks by effectively capturing spatial features and detecting objects at multiple scales. In parallel, transformer-based models like RT-DETR have

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

introduced attention mechanisms that capture global contextual relationships, improving detection performance in cluttered and occluded environments commonly found underwater.

To further enhance detection accuracy, preprocessing techniques are applied to improve the visual quality of underwater images. Methods such as color correction, contrast enhancement, and noise reduction help highlight object features and improve model learning. In this study, a multi-class underwater waste detection framework is developed using a dataset of fifteen object categories. The performance of YOLOv8, YOLOv9, and RT-DETR is evaluated using metrics such as precision, recall, and mean Average Precision (mAP), with the objective of identifying the most effective approach for robust underwater waste detection.

II. LITERATURE REVIEW

The integration of deep learning into underwater object detection systems has gained considerable attention in recent years due to the increasing need for automated marine waste monitoring. Early research in this domain primarily relied on traditional computer vision techniques such as edge detection, thresholding, and colour-based segmentation to identify objects in underwater images. Although these approaches were effective in controlled environments, they struggled to perform reliably under real-world conditions characterized by poor visibility, colour distortion, and complex backgrounds. These limitations highlighted the need for more robust and adaptive detection methods capable of handling the challenges inherent in underwater imaging.

With the advancement of deep learning, convolutional neural network (CNN)–based models have significantly improved the performance of underwater object detection systems. Architectures based on the YOLO framework have been widely adopted due to their ability to achieve real-time detection with high accuracy. Models such as YOLOv8 utilize advanced feature extraction techniques and multi-scale detection strategies to identify objects of varying sizes within complex underwater scenes. These capabilities have made CNN-based detectors highly suitable for applications such as marine debris detection, underwater surveillance, and autonomous underwater vehicle navigation. However, despite their strong performance, CNN-based models primarily focus on local feature extraction, which can limit their effectiveness in capturing broader contextual relationships in cluttered environments.

To overcome these limitations, recent studies have explored improved detection architectures and alternative modeling approaches. Enhanced CNN-based models such as YOLOv9 introduce advanced feature aggregation mechanisms and optimized gradient flow to improve detection accuracy and efficiency. At the same time, transformer-based architectures have emerged as a powerful alternative for visual recognition tasks. These models leverage attention mechanisms to capture global contextual information across the entire image, enabling better understanding of complex visual patterns. Models such as RT-DETR combine transformer-based feature learning with efficient detection pipelines, offering competitive real-time performance while improving robustness in scenarios involving occlusion and background interference.

In addition to model architecture, the role of preprocessing techniques has been extensively discussed in the literature as a critical factor in improving detection performance. Underwater images often suffer from low contrast, noise, and colour degradation, which can negatively impact model training and inference. Techniques such as white balance correction, gamma adjustment, contrast enhancement, and filtering are commonly applied to improve image quality and highlight relevant features. Furthermore, recent studies emphasize the importance of comprehensive evaluation frameworks that compare multiple detection models on multi-class underwater datasets. Despite these advancements, challenges such as class imbalance, small object detection, and variability in underwater conditions continue to motivate ongoing research in this field.

III. RESEARCH METHODOLOGY

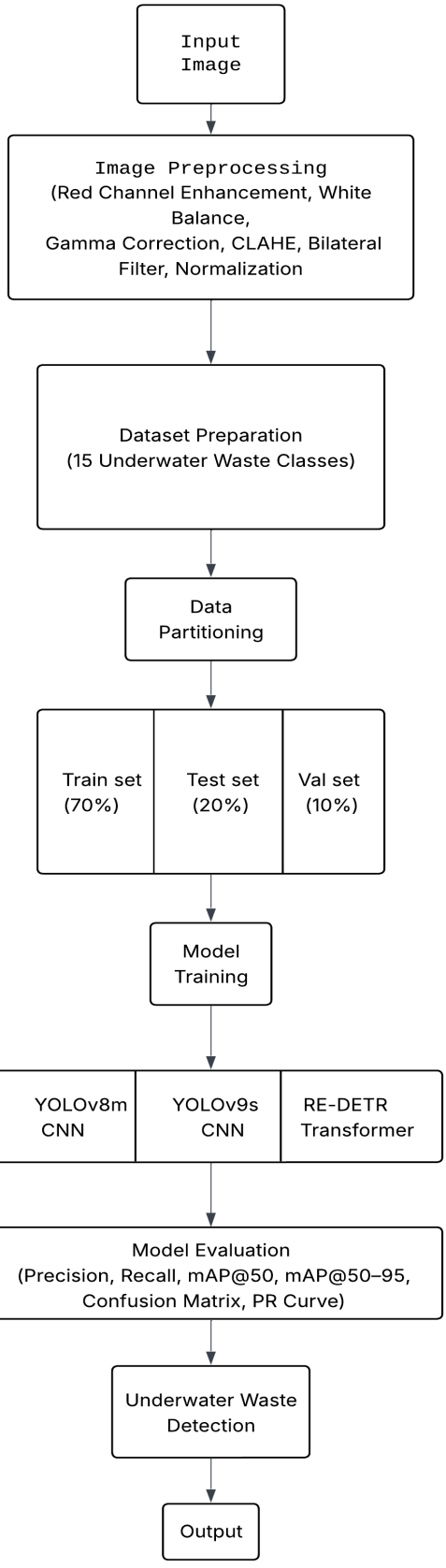
This study follows a structured methodology to develop and evaluate a deep learning–based framework for multi-class underwater waste detection. The methodology is designed to ensure consistent training conditions and a fair comparison between convolutional neural network–based and transformer-based object detection models.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105



International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

Data Collection and Preparation

The dataset used in this study was obtained from the Roboflow and consists of annotated underwater images representing fifteen categories of marine waste. The dataset includes objects such as plastic bags, bottles, nets, gloves, and tires. All images were resized to a uniform resolution of 640 × 640 pixels to maintain consistency across the training process.

To enable systematic training and evaluation, the dataset was divided into three subsets: training (70%), validation (10%), and testing (20%). The training set is used for model learning, the validation set is used for tuning and monitoring performance during training, and the test set is used for final evaluation.

Image Preprocessing

Underwater images often suffer from poor visibility due to light absorption and scattering effects. To address these challenges, a preprocessing pipeline was applied before model training. The preprocessing steps include red channel enhancement, white balance correction, gamma correction, contrast limited adaptive histogram equalization (CLAHE), bilateral filtering, and normalization. These techniques improve image clarity, enhance contrast, and reduce noise, enabling the models to learn more discriminative features.

Model Selection and Training

Three object detection models were selected for evaluation: YOLOv8, YOLOv9, and RT-DETR. These models represent both CNN-based and transformer-based detection approaches.

All models were trained using identical input configurations to ensure a fair comparison. The training parameters include an input image size of 640 × 640 pixels, batch size of 16, learning rate of 0.01, and training duration of 40 epochs. The models were trained using annotated bounding box labels to learn object localization and classification.

Performance Evaluation

The performance of the trained models was evaluated using standard object detection metrics. Precision and recall were used to measure detection accuracy and object coverage, respectively. Mean Average Precision (mAP@50 and mAP@50–95) was used to assess overall detection performance across different intersection over union thresholds.

In addition to numerical evaluation, confusion matrices were used to analyze class-wise detection performance, and qualitative analysis was performed using sample detection outputs to verify the effectiveness of the models in real-world underwater conditions.

Comparative Analysis

A comparative analysis was conducted to evaluate the strengths and limitations of each model. The CNN-based models (YOLOv8 and YOLOv9) were analyzed in terms of precision and real-time performance, while the transformer-based RT-DETR model was evaluated for its ability to capture global contextual information and improve detection coverage.

This methodology ensures a comprehensive evaluation of different object detection architectures and provides a reliable basis for identifying the most suitable model for underwater waste detection.

IV. RESULTS AND DISCUSSION

This section analyzes the performance of the evaluated object detection models for multi-class underwater waste detection. The models considered include YOLOv8, YOLOv9, and RT-DETR. The evaluation is based on both quantitative metrics and qualitative observations.

A. Quantitative Performance Analysis

The overall performance of the models was evaluated using precision, recall, and mean Average Precision (mAP). The results indicate that all three models achieve competitive performance in detecting underwater waste objects.

YOLOv8 achieves the highest precision among the evaluated models, indicating that it produces fewer false positive detections. This makes it reliable for applications where prediction accuracy is critical. YOLOv9 demonstrates balanced performance with moderate improvements in recall compared to YOLOv8, suggesting better detection capability for certain object classes.

The RT-DETR model achieves the highest recall value, indicating that it is able to detect a larger proportion of actual objects present in underwater images. This is particularly important in environmental monitoring scenarios where missing objects can lead to incomplete analysis. Although its mAP@50–95 is slightly lower than YOLOv8, the difference is marginal, and the model maintains competitive overall performance.

International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

B. Class-wise Performance Analysis

The confusion matrix analysis provides insights into the performance of the models across individual object categories. Classes such as **cellphone**, **net**, **plastic bag**, and **sunglasses** exhibit high detection accuracy, as indicated by strong diagonal values. These objects are easier to detect due to their distinct visual characteristics. On the other hand, certain classes such as **metal**, **plastic**, and **rod** show relatively lower detection accuracy. This is mainly due to their visual similarity with surrounding objects and the challenging conditions of underwater environments, including low contrast and occlusion. Despite these challenges, the overall class-wise performance remains consistent across the majority of categories.

C. Qualitative Analysis

The qualitative results obtained from the detection outputs demonstrate the practical effectiveness of the models. The models successfully detect multiple underwater waste objects within complex scenes, including objects such as tires, plastic debris, and bottles. The predicted bounding boxes accurately localize the objects, and the associated confidence scores indicate reliable classification. These results confirm that the proposed framework is capable of handling real-world underwater conditions, including variations in lighting, background clutter, and object appearance.

D. Comparative Discussion

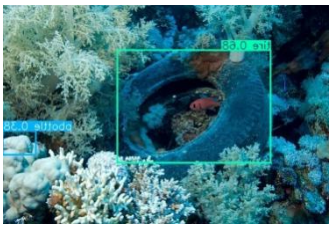
The comparative analysis highlights the strengths of both CNN-based and transformer-based detection approaches. YOLOv8 provides high precision and efficient real-time detection, making it suitable for applications requiring fast and accurate predictions. YOLOv9 introduces architectural improvements that enhance feature aggregation and improve detection performance across multiple object scales. The transformer-based RT-DETR model demonstrates improved detection coverage due to its ability to capture global contextual information using attention mechanisms. This enables the model to detect objects that may be partially occluded or embedded within complex backgrounds.

Overall, the results indicate that while YOLO-based models provide strong precision and efficiency, RT-DETR offers improved recall and robustness in underwater environments. This makes RT-DETR a suitable choice for applications that require comprehensive detection of underwater waste objects.

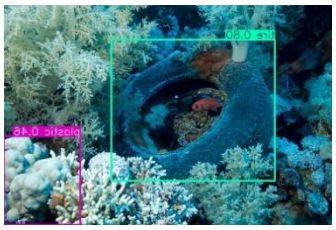
Model	Precision	Recall	mAP@50	mAP@50–95
YOLOv8	**0.847**	0.707	**0.795**	**0.527**
YOLOv9	0.802	0.725	0.793	0.519
RT-DETR	0.846	**0.765**	**0.795**	0.515

Table 1

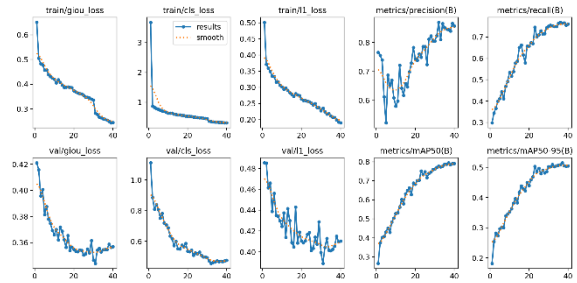
YOLO v9s
Fig 1



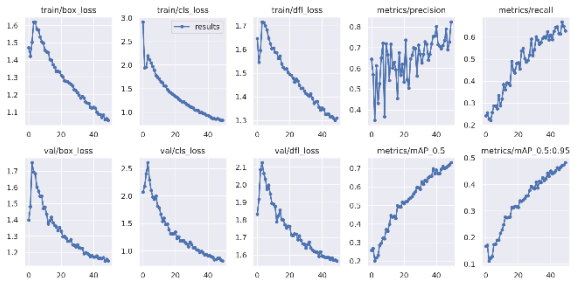
YOLOv8m
Fig 2



RE-DETR
Fig 3



RE-DETR



YOLO v9s

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

V. CONCLUSION

This study presented a deep learning–based framework for multi-class underwater waste detection using both convolutional neural network and transformer-based object detection models. A preprocessing pipeline was applied to enhance underwater image quality and improve feature visibility under challenging environmental conditions.

The experimental results demonstrate that all evaluated models—YOLOv8, YOLOv9, and RT-DETR—achieve competitive performance in detecting underwater waste objects. YOLOv8 provides high precision and efficient real-time detection, while YOLOv9 offers balanced performance with improved feature representation.

The transformer-based RT-DETR model demonstrates superior recall, indicating improved detection coverage in complex underwater environments. This suggests that transformer-based architectures are more effective in capturing global contextual information, enabling better identification of partially occluded and visually challenging objects.

Overall, the findings indicate that RT-DETR is a promising approach for underwater waste detection, offering robust performance and improved detection capability. The proposed framework can contribute to the development of automated systems for marine pollution monitoring and environmental conservation. Future work may focus on improving detection accuracy for challenging object classes and optimizing the models for real-time deployment in underwater robotic systems.

VI. FUTURE WORK

1. **Improve detection of challenging classes** such as metal, plastic, and rod using advanced data augmentation and class balancing techniques.
2. **Expand the dataset** with more diverse underwater scenes to improve model generalization and robustness.
3. **Optimize models for real-time deployment** on embedded systems and autonomous underwater vehicles (AUVs).
4. **Integrate ensemble learning techniques** by combining YOLO and transformer models to improve overall detection performance.
5. **Incorporate temporal information** using video-based detection for continuous underwater monitoring.
6. **Enhance preprocessing pipeline** using advanced image restoration and dehazing techniques.
7. **Explore lightweight transformer models** to reduce computational complexity while maintaining accuracy.
8. **Develop a real-time monitoring system** for large-scale marine pollution detection and analysis.

REFERENCES

1. K. Samanth, R. Ramyashree, B. N. Anoop, and S. Raghavendra, “A Comprehensive Study on Underwater Object Detection Using Deep Neural Networks,” *IEEE Access*, vol. 13, pp. 99446–99464, 2025, doi: 10.1109/ACCESS.2025.3577239.
2. S. Wu, P. Luo, Y. Song, and G. Jiang, “Underwater Image Enhancement and Trash Detection Using Deep Learning,” in *Proc. IEEE*, 2025.
3. Y. Liu, X. Fu, X. Ding, Y. Huang, and J. Paisley, “A Benchmark Dataset and Learning Pipeline for Underwater Image Enhancement,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 801–1005, 2021.
4. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
5. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” *arXiv preprint arXiv:2004.10934*, 2020.
6. C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, “YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information,” *arXiv preprint arXiv:2402.13616*, 2024.
7. G. Jocher et al., “Ultralytics YOLOv8,” GitHub Repository, 2023. [Online]. Available: https://github.com/ultralytics/ultralytics
8. N. Carion et al., “End-to-End Object Detection with Transformers,” in *Proc. European Conf. Computer Vision (ECCV)*, 2020, pp. 213–229.
9. W. Liu et al., “SSD: Single Shot MultiBox Detector,” in *Proc. European Conf. Computer Vision (ECCV)*, 2016, pp. 21–37.
10. S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

11. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2017, pp. 2961–2969.
12. T.-Y. Lin et al., "Feature Pyramid Networks for Object Detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117–2125.
13. Z. Wang et al., "DETR: End-to-End Object Detection with Transformers," in *Proc. European Conf. Computer Vision (ECCV)*, 2020, pp. 213–229.
14. J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object Detection via Region-Based Fully Convolutional Networks," in *Proc. Advances in Neural Information Processing Systems*, 2016, pp. 379–387.
15. [15] M. Islam, M. R. Islam, and J. Sattar, "Fast Underwater Image Enhancement for Improved Visual Perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.
16. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
17. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
18. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
19. S.Tamilselvi, R.Prakash, C.Nagarajan,"Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
20. S.Tamilselvi, R.Prakash, C.Nagarajan," Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428
21. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
22. Vimal, V. R., John Justin Thangaraj, S., Narayanan, L. K., Alagu Thangam, S., Loganayagi, S., & Balakrishnan, S. (2025, April). Enhanced Phishing Detection and Classification Using an Ensemble Machine Learning Approach for URL Analysis. In *International Conference on Information and Communication Technology for Intelligent Systems* (pp. 229-239). Springer Nature Singapore.
23. Mathew, A. (2021). Obfuscation Techniques for Magecart Detection and Prevention. International Journal of Computer Science and Mobile Computing, 10(2), 39-44.
24. Soundappan, S. J. (2026). Building Trustworthy AI: Explainability and Security in Modern Cloud-Native Data-Driven Ecosystem Platforms. *International Journal of Engineering & Extended Technologies Research (IJEETR), 8(2), 570-579.*
25. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.
26. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
27. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
28. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
29. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721
30. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
31. D. Akkaynak and T. Treibitz, "Sea-thru: A Method for Removing Water From Underwater Images," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 1682–1691.

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802105

32. C. Li, J. Guo, and C. Guo, "Emerging From Water: Underwater Image Color Correction Based on Weakly Supervised Learning," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 323–327, 2018.
33. X. Fu, P. Zhuang, Y. Huang, X. Ding, and J. Paisley, "A Retinex-Based Enhancing Approach for Single Underwater Image," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2014, pp. 4572–4576.
34. P. Panetta, C. Gao, and S. Agaian, "Human Visual System-Based Image Enhancement and Logarithmic Contrast Measure," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 38, no. 1, pp. 174–188, 2008.
35. Z. Zhang et al., "Underwater Object Detection Based on Improved YOLO Network," *IEEE Access*, vol. 9, pp. 123456–123468, 2021.
36. S. Mandal, S. Gupta, and A. Banerjee, "Deep Learning-Based Marine Debris Detection in Underwater Images," *IEEE Access*, vol. 10, pp. 45678–45689, 2022.
37. M. Pedersen et al., "Detection of Marine Litter Using Deep Neural Networks," *IEEE Access*, vol. 8, pp. 102933–102945, 2020.
38. R. Li, X. Zhang, and Y. Wang, "Deep Learning-Based Marine Debris Detection Using Underwater Imagery," *IEEE Journal of Oceanic Engineering*, vol. 46, no. 4, pp. 1234–1245, 2021.
39. H. Zhang, Y. Liu, and X. Ding, "Underwater Object Detection via Multi-Scale Deep Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 9, pp. 1543–1547, 2021.
40. J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3431–3440.