

## Review of Underwater Object Detection Using YOLO: Advances, Challenges, and Future Directions

Shaymaa K. Hussein<sup>1</sup>, Zainab N. Nemer<sup>2</sup>

<sup>1,2</sup>Department of Computer Science, College of Computer Science and Information Technology, University of Basrah, Iraq

### Article Info

#### Article history:

Received May 4, 2025  
Revised Jun 29, 2025  
Accepted July 7, 2025

#### Keywords:

Underwater Object Detection  
YOLO (You Only Look Once)  
Attention Mechanism  
Underwater Datasets  
Deep learning

### ABSTRACT

Underwater object detection is critical for environmental monitoring, maritime security, and rescue operations, yet it faces challenges such as light scattering, color distortion, and low visibility. This paper presents a comprehensive review of YOLO (You Only Look Once) algorithms and their integration with attention mechanisms to address these challenges. We systematically analyze the evolution of YOLO models—from YOLOv1 to YOLOv11—highlighting key architectural advancements, including anchor-free detection, multi-scale feature fusion, and attention modules like CBAM and SimAM. These innovations enhance detection accuracy in underwater environments, where small, occluded objects and dynamic backgrounds degrade performance.

We evaluate YOLO variants on underwater datasets (e.g., URPC, SUIM, RUIE), comparing metrics such as mean Average Precision (mAP), inference speed (FPS), and computational complexity. Attention mechanisms, including spatial, channel, and self-attention, are shown to improve feature discrimination, achieving up to a 25% reduction in false positives. Challenges such as limited annotated data and real-time processing constraints are discussed, along with solutions like semi-supervised learning and synthetic data augmentation.

Based on our findings, YOLOv8 and YOLOv9 models integrating attention mechanisms provide the best trade-offs between accuracy and efficiency for underwater detection. These suggest other directions for future research such as novel lightweight attention designs and multi-sensor fusion to give even