Volume 2, Issue 1

Innovation Series Quarterly

https://journal.academicapress.org/aps/index.php/isq

Research Article

Real-Time AI Water Surface Obstacle Detection for Small Boat Navigation Using Stereo Vision and Edge Computing

Abdullah Atiq Arifin[1, *], Mohd Fahmi Mat Zin[2], Noor Haslyena Hassan[3]

[1] Dept. of Petrochemical Engineering, Polytechnic Tun Syed Nasir, Malaysia; abdullahatiq@ptsn.edu.my
[2] Dept. of Petrochemical Engineering, Polytechnic Tun Syed Nasir, Malaysia; fahmi@ptsn.edu.my
[3] Department of Mechanical Engineering, Polytechnic Melaka, Melaka, Malaysia; nlyena2009@gmail.com
* Correspondence: abdullahatiq1983@gmail.com;019-7557204

Keywords:
AI Water Surface Obstacle Detector
Computer vision
Jetson Orin NX
Marine navigation
YOLOv11

Abstract: *Reliable perception of water surface obstacles is essential for safe boat navigation and maritime monitoring. This project presents an AI-based water surface obstacle detection system using stereo vision and an NVIDIA Jetson Orin NX edge platform. A custom Roboflow dataset with four target classes (boat, debris, structures, aquatic plant) was developed, and a YOLOv11 model was fine-tuned using pretrained COCO weights. Evaluation on unseen samples achieved mAP@0.5 of 0.7561, precision 0.7988, recall 0.6586, and F1-score 0.7220. The results highlight strong detection performance, demonstrating potential for real-time deployment in boat navigation safety applications. Future work includes dataset expansion and sensor fusion.*

1. INTRODUCTION

Reliable real-time perception of water-surface obstacles plays a critical role in maritime safety and modern marine navigation. Autonomous maritime systems and intelligent navigation platforms increasingly depend on computer vision for situational awareness and safe decision-making in dynamic water environments. Compared to land-based perception systems, water surfaces present unique complexities including high visual noise induced by wave interference, specular glare, low visual contrast and reduced spatial resolution for small or distant objects such as floating debris and aquatic plant. These factors reduce the reliability of traditional rule-based image processing and classical feature-matching techniques which often struggle to extract meaningful object information from cluttered backgrounds (Wang et al., 2023).

Recent advances in deep learning have enabled robust computer vision-based maritime perception offering high-resolution detection capabilities viable for edge-constrained platforms (Zhang et al., 2021; Sung et al., 2020; Yang et al., 2024). Among modern detectors, models from the You Only Look Once (YOLO) family have gained prominence due to their balanced trade-off between accuracy and inference speed. However, while pretrained models achieve excellent performance on generic domains, most are trained on large-scale datasets such as COCO that do not sufficiently represent maritime-specific characteristics such as debris, floating aquatic plant or low-profile surface obstacles (Kim et al., 2022; Wang et al., 2023). This gap highlights the need for domain-specific datasets and fine-tuning approaches to improve contextual recognition performance.

To address dataset scarcity and domain mismatch challenges, recent studies employ transfer learning and augmentation pipelines—often facilitated through tools such as Roboflow—to construct maritime-specific annotated datasets. Fine-tuned models achieve significantly improved detection accuracy while requiring lower computational cost relative to training from scratch (Lin et al., 2021; Reddy & Basha., 2025). Beyond model training, efficient onboard deployment remains a key research challenge. While many studies validate performance in desktop environments, only limited work demonstrates fully embedded deployment capable of real-time field operation (Haijoub et al., 2024).

2. RELATED WORK

The YOLO family has shown strong applicability in maritime perception tasks including vessel classification, debris detection and surface anomaly monitoring. Earlier versions such as YOLOv3 and YOLOv5 achieved successful detection performance on water surface scenes due to their favourable accuracy and speed balance. More recent iterations such as YOLOv5 and YOLOv7 demonstrated improvements in maritime benchmarking datasets, reporting mAP increases between 10–15% through lightweight architectures and attention mechanisms optimized for unmanned surface vehicle (USV) operation (Al-Hattab et al., 2023; Li et al., 2023; Yang et al., 2024). Further advancements in YOLOv8 and YOLOv11 incorporate transformer-based feature extraction modules, enhanced bounding box regression and improved training stability, resulting in more robust detection under cluttered or visually noisy environments (Khanam et al., 2024; Yu et al., 2024).

Dataset customization continues to be a prominent strategy for maritime object detection due to the limitations of general-purpose datasets. Studies have shown that transfer learning applied to domain-specific datasets with constrained sample sizes enables strong detection performance with reduced training complexity (Lin et al., 2021; Reddy & Basha., 2025). Complementarily, augmentation tools such as Roboflow provide scalable workflows for dataset curation, annotation and synthetic augmentation, further mitigating data scarcity.

Edge-based deployment remains an active research frontier. Embedded platforms such as the NVIDIA Jetson series are increasingly adopted for onboard perception due to GPU-accelerated parallel computation, enabling real-time inference under power-constrained settings (Signaroli et al., 2025). Integration with stereoscopic camera systems extends perception capabilities beyond detection to include range estimation, depth inference and collision risk assessment which are essential for USV and autonomous maritime navigation (Al-Hattab et al., 2023). Despite these advances, literature shows limited demonstrations of compact, fully embedded and ready-to-deploy maritime detection systems that operable in real outdoor conditions thus highlighting a continuing research gap.

3. SYSTEM DESIGN & METHODOLOGY

3.1 System Architecture Overview

The proposed system integrates a ZED 2i stereo camera with an NVIDIA Jetson Orin NX (16GB) embedded platform to perform onboard water surface obstacle detection. The ZED 2i provides stereo imagery for depth estimation while the Jetson platform executes YOLO-based object detection and inference without cloud dependency. The perception pipeline consists of (i) sensor acquisition, (ii) deep-learning inference, (iii) decision and (iv) output for navigation assistance. Fig. 1 illustrates the overall system architecture pipeline.

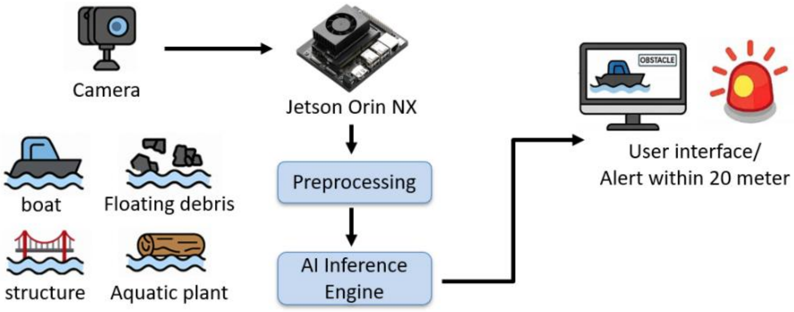


Figure 1. System Architecture for AI-based water surface obstacle detection

3.2 Dataset Development and Annotation

Due to the absence of suitable public datasets that represent real maritime environments, a custom dataset was developed for this study. Images were captured across rivers and coastal areas to ensure variation in lighting, water clarity, reflections and background complexity with additional public maritime samples added to increase object diversity. All data were uploaded to Roboflow, manually annotated and categorized into four classes: boat, floating object, structure and aquatic plant. The final dataset was exported in YOLOv11 format and split into training (70%), validation (20%), and testing (10%) sets following standard machine learning practices.

Figure 2. Example of annotated dataset with four maritime obstacle classes

3.3 *Model Training and Optimization*

The pretrained YOLOv11 model with COCO weights was selected as the baseline architecture due to its balance of detection accuracy and computational efficiency. The YOLOv11 model was fine-tuned using PyTorch-based Ultralytics Framework training pipelines with hyperparameters such as batch size, learning rate and epoch count tuned empirically to improve convergence stability. Model performance was evaluated using mAP, precision, recall and F1-score metrics. In addition, qualitative inference was conducted to assess robustness under cluttered water backgrounds and surface reflections.

3.4 *Embedded Deployment Pipeline*

For real-time field deployment, the prototype system was deployed on the Jetson Orin NX 16GB platform with ZED 2i stereo camera (4mm lens) for depth-based perception. To support field level implementation, the complete hardware assembly was installed within an IPX6 rated waterproof die cast enclosure and paired with an IPX6 marine LCD interface as shown in Fig. 3. Additionally, an external alarm strobe module was integrated to provide real-time warning signals during obstacle detection event. Initial deployment achieved ~30 FPS, enabling low-latency perception suitable for low-speed small craft navigation. The Jetson platform operated at an estimated 10–25 W power envelope, supporting battery-based and mobile configurations. Field deployment feasibility was validated through environmental testing including IPX56 ingress compliance confirming suitability for outdoor maritime exposure (see Fig. 4).

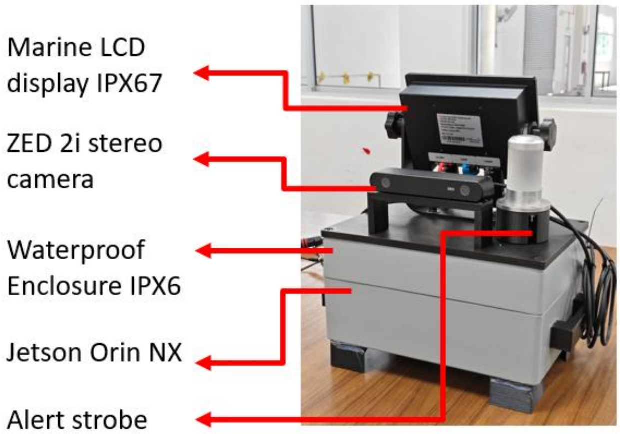


Figure 3. Prototype deployment using Jetson Orin NX

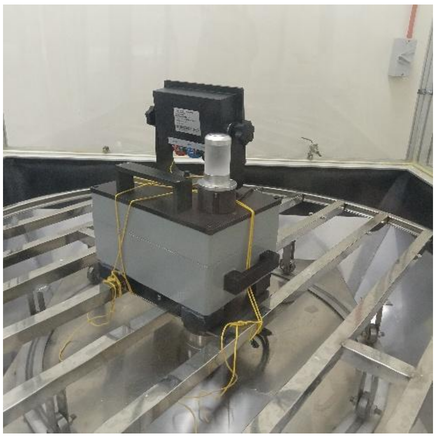


Figure 4. IPX56 ingress test for embedded hardware prototype

4. RESULT AND DISCUSSION

The fine-tuned YOLOv11 model successfully improved maritime obstacle detection performance compared to it pretrained baseline demonstrating strong quantitative results on unseen test samples. The model achieved **mAP@0.5 = 0.7561**, precision **0.7988**, recall **0.6586**, and F1-score **0.7220**, indicating effective adaptation to aquatic scene characteristics. *Table 1* summarizes performance metrics for both validation and test datasets, confirming consistent behaviour across evaluation splits.

Table 1. Performance metrics for validation and test datasets

Metric	Validation	Test
mAP@0.5	0.7431	0.7561
mAP@0.5	0.7431	0.7561
Precision	0.8088	0.7988
Recall	0.6997	0.6586
F1-score	0.7503	0.722
Optimal confidence threshold	0.446	0.446

4.1 Confusion Matrix Interpretation

The normalized confusion matrix (Fig. 5) demonstrates that the model performs well across multiple classes with particularly strong detection accuracy for plant and objects achieving true classification rates of 0.87 and 0.83, respectively. The boat class also exhibits a relatively high correct prediction rate of 0.75 indicating stable generalization toward dynamic floating obstacles.

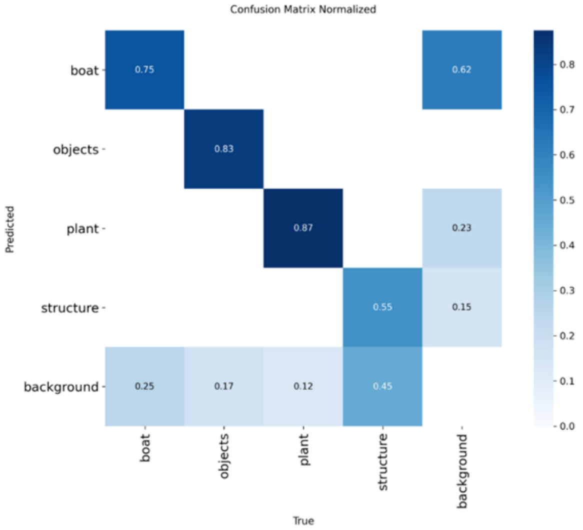


Figure 5. Confusion matrix normalized

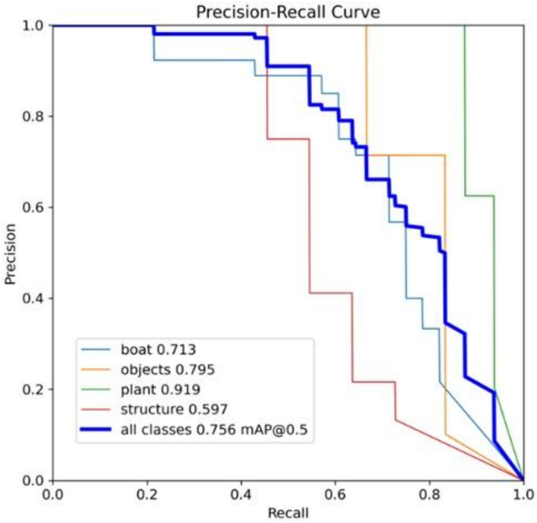


Figure 6. Precision-Recall curve

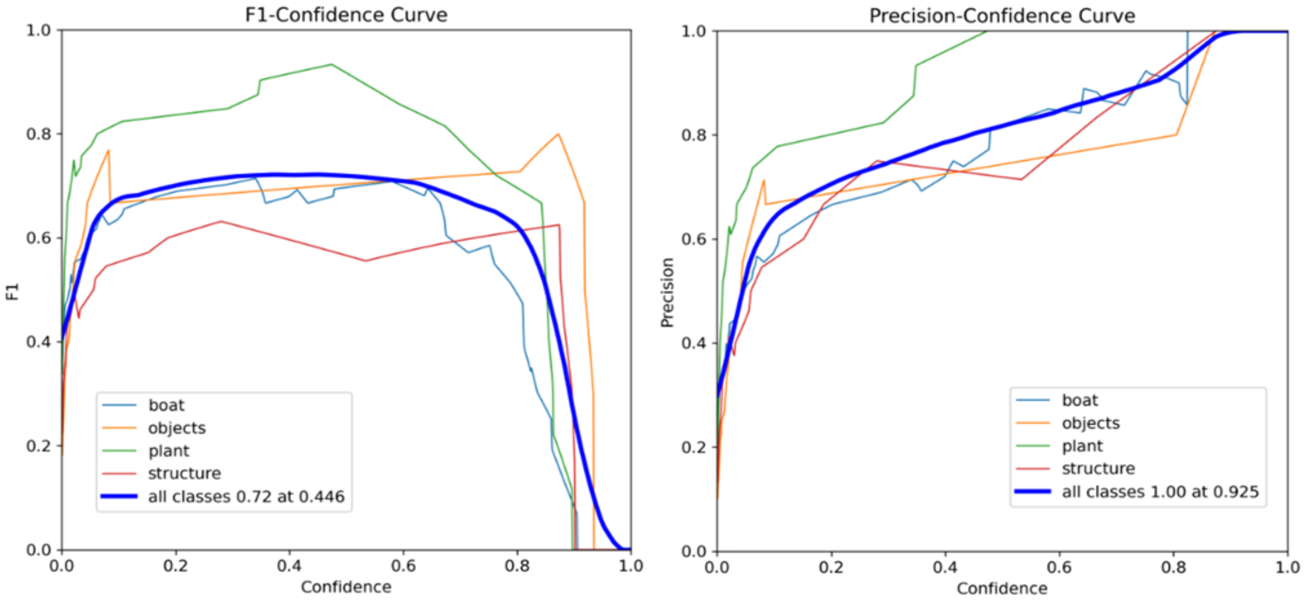
In contrast, the structure class shows a lower correct prediction value of 0.55, highlighting difficulty in differentiating man-made structures (e.g., bridges, piers, retaining walls) from the background. This challenge is consistent with findings in other marine vision studies where background–structure similarity and occlusion reduce detection precision (Kim et al., 2022). The most frequent misclassification occurred between structure (0.45) and boat (0.25) likely due to environmental factors such as reflections, water turbulence and illumination variation which have been extensively identified as key failure points in aquatic perception systems. These observations confirm that aquatic environments introduce higher visual ambiguity compared to terrestrial datasets.

4.2 Precision-Recall and Confidence-Based Behavior

The precision–recall curve demonstrated stable behavior across most classes with the overall testing mAP@0.5 reaching 0.7561 indicating high object-level consistency as shown in Fig.6. The F1-Confidence curve (see Fig. 7a) identified the optimal operational confidence threshold at 0.446 with an

F1-Score of 0.72 represent the best trade-off between precision and recall when deployed on the NVIDIA Jetson Orin NX.

Conversely, Fig.7(b) show the precision-confidence curve peaked at 1.00 precision at 0.925 confidence confirming that higher confidence thresholds suppress false positives but at the cost of increased missed detections as. Similarly, the recall-confidence curve demonstrated a gradual decline under increasing thresholds validating that the model becomes more conservative as confidence filtering increases. These behaviors support the selection of a dynamic confidence threshold strategy during real-time deployment particularly when environmental visibility changes.



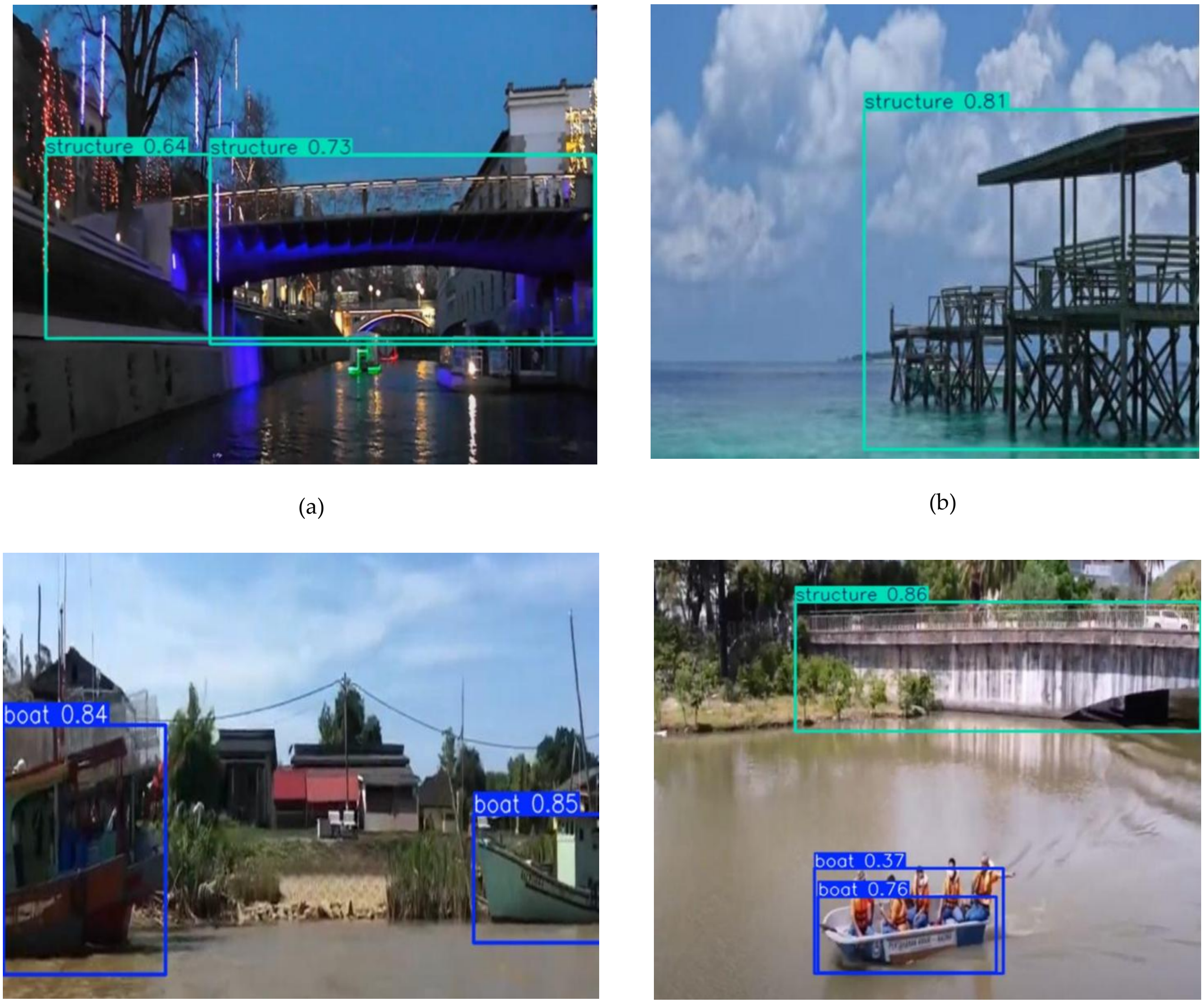
(a) (b)

Figure 7. Confidence curve a) F1-Confidence and b) Precision-Confidence

4.3 Qualitative Detection Results

Inference sample in Fig. 8(a-d) demonstrated that the model successfully identified multiple water-borne and structural features including moving boats at different distance and scales, aquatic plant, stationary marine navigational buoy, bridge, docks, and pier structure. Detection confidence remained stable even under challenging conditions such as reflections, shadows and variable backgrounds. Minor duplication in bounding boxes occurred in dense spatial regions especially where multiple objects appeared close to one another or were partially occluded.

Overall, the fine-tuned YOLOv11 model demonstrated strong potential for real-time water surface obstacle detection, achieving reliable mAP, F1-Score and precision-recall dynamics. While class confusion was observed between background and structure categories, the integration of stereo depth sensing and adaptive thresholding strategies is expected to further enhance reliability during field deployment. The results confirm that the model is suitable for embedded execution on the Jetson Orin NX platform and capable of supporting maritime navigation and safety-alerting applications in inland waterways and near-shore environments.



(a) (b)

(c) (d)

Figure 8. Detection of a) brige and structure, b) pier structure, c) boat and d) moving boats and static structure

5. CONCLUSION

This study demonstrates that YOLOv11 is a viable real-time perception model for water surface obstacle detection. The system achieved strong results with mAP@0.5 = 0.7561 and demonstrated stable real-time performance during field deployment. Qualitative testing confirmed that the model was capable of detecting critical waterway obstacle and hazards under varying environmental condition thus show its potential for integration into maritime navigation system, environmental monitoring and early-warning collision-avoidance systems. While performance was strong for the plant, boat and objects classes, structural objects showed higher misclassification due to visual similarity with background regions and environmental reflections. Future work will address these limitations through expanding the dataset including integration of multimodal sensing such as LiDAR and thermal infrared imaging.

Acknowledgments: The authors acknowledge the support from the T-AIGS research team, Politeknik Tun Syed Nasir, CUTe IIUM, and the funding support from JPPKK under the T-AIGS Grant Scheme.

REFERENCES

Al-Hattab, Y. A., Abidin, Z. Z., Faizabadi, A. R., Zaki, H. F. M., & Ibarahim, A. I. (2023). Integration of stereo vision and MOOS-IvP for enhanced obstacle detection in USVs. *IEEE Access*, *11*,pp 128932-128956. Doi: 10.1109/ACCESS.2023.3332032

Haijoub, A., Hatim, A., Guerrero-Gonzalez, A., Arioua, M., & Chougdali, K. (2024). Enhanced YOLOv8 Ship Detection Empower Unmanned Surface Vehicles for Advanced Maritime Surveillance. *Journal of Imaging*, *10*(12), 303. https://doi.org/10.3390/jimaging10120303

Khanam, R., & Hussain, M. (2024). YOLOv11: An Overview of the Key Architectural Enhancements. *arXiv:2410.17725v1*

Kim, J.-H., Kim, N., Park, Y. W., & Won, C. S. (2022). Object Detection and Classification Based on YOLO-V5 with Improved Maritime Dataset. *Journal of Marine Science and Engineering*, *10*(3), 377. https://doi.org/10.3390/jmse10030377

Li, Y., Li, Y., Jiang, Z. & Wang, H. (2023). Real-time Detection of Surface Floating Garbage Based on Improved YOLOv7. Intelligent Robotics and Application, ICIRA 2023 proceeding, Part VI, *pp. 573-582*. https://doi.org/10.1007/978-981-99-6480-2_47

Lin, F., Hou, T., Jin, Q., & You, A. (2021). Improved YOLO Based Detection Algorithm for Floating Debris in Waterway. *Entropy*, *23*(9), 1111. https://doi.org/10.3390/e2309111

Reddy, K. G. & Basha, S. S. (2025). Real Time Object Identification : A Study on COCO Dataset. *Advance in Engineering Research* 257, 860-870.

Signaroli, M., Lana, A., Cutolo, E., Alos, J., & Gonzalez-Cid, Y. (2025). Real-time tracking of recreational boats in coastal areas using deep learning. *Ocean and Coastal Management*, 267(2025) 107762, https://doi.org/10.1016/j.ocecoaman.2025.107762

Sung, L., Myung, I. & M, O. (2020). Image-based ship detection using deep learning. *Ocean Systems Engineering*, 4(40), 415-434. https://doi.org/10.12989/ose.2020.10.4.415

Wang, L., Xiao, Y., Zhang, B., Liu, R., & Zhao, B. (2023). Water Surface Targets Detection Based on the Fusion of Vision and LiDAR. *Sensors*, 23(4),1768. **https://doi.org/10.3390/s23041768**

Yang, D., Solihin, M. I., Zhao, Y., Li, W., Cai, B. & Chen, C. (2024). A streamlined approach for intelligent ship object detection using EL-YOLO algorithm. *Scientific Reports*, *14*, Article 15234. **https://doi.org/10.1038/s41598-024-64225-y**

Yu, C., Yin, H., Rong, C., Zhao, J., Liang, X., Li, R., & Mo, X. (2024). YOLO-MRS: An efficient deep learning-based maritime object detection method for unmanned surface vehicles. *Applied Ocean Research*, *153*, Article 104240. https://doi.org/10.1016/j.apor.2024.104240

Zhang, L., Wei, Y., Wang, H., Shao, Y., & Shen, J. (2021). Real-Time Detection of River Surface Floating Object Based on Improved RefineDet. IEEE Access, vol. 9, pp. 81147-81160. doi: 10.1109/ACCESS.2021.3085348