ECOSIA | International Conference
Institut Teknologi dan Bisnis Asia Malang
Indonesia, 10 December 2025

Real-Time Water Surface Obstacle Detection Using YOLOv11 and Stereo Vision on NVIDIA Jetson Orin NX

Abdullah Atiq bin Arifin[1]
Noor Haslyena binti Hassan[2]
Mohd Fadli bin Kambas[1]

[1] Affiliation : Polytechnic Tun Syed Nasir Johor, Malaysia
[2] Affiliation : Polytechnic Melaka, Malaysia
Correspondence E-mail : abdullahatiq@yahoo.com

Abstract

Introduction/Main Objectives:This research aims to investigate how digital transformation, company size, and profitability influence tax avoidance behavior in healthcare companies listed on the Indonesia Stock Exchange (IDX) between 2022 and 2024.

Background Problems: Despite post-pandemic performance growth—marked by increased revenue, asset expansion, and accelerated digitalization—the healthcare sector's average Effective Tax Rate (ETR) remains lower than the statutory corporate tax rate, indicating persistent potential for tax avoidance.

Research Methods: The study employs an associative quantitative approach using secondary data from annual financial reports. Data analysis is conducted through multiple linear regression, with the Effective Tax Rate (ETR) serving as the measure for tax avoidance.

Finding/Results: Digital transformation, company size, and profitability collectively have a significant impact on tax avoidance.Individually, digital transformation and company size show a significant effect on tax avoidance.Profitability has only a slightly significant individual effect.Digital transformation acts as a key factor in reducing tax avoidance practices.Larger companies exhibit greater scope for tax avoidance.

Conclusion: Digital transformation is an effective and significant driver in reducing tax avoidance, whereas larger company size correlates with increased potential for such practices. The findings highlight the importance of digital adoption and regulatory attention to firm scale in mitigating tax avoidance.

Keywords: AI Water Surface Obstacle Detector, Computer vision, Jetson Orin NX, Marine navigation, YOLO.

Introduction

Reliable real-time perception of water surface obstacles plays a critical role in maritime safety and marine navigation. Autonomous maritime systems and intelligent navigation platforms increasingly rely on computer vision to perceive their surroundings and support safe decision-making in dynamic environments. Compared to land-based perception systems, water surface environments pose unique challenges such as high visual noise from wave interference, glare from specular reflection and limited resolution for small or distant objects like floating debris and aquatic plants. These environmental characteristics reduce the reliability of traditional rule-based image processing or classical feature-matching techniques which often struggle to differentiate meaningful objects from background noise (Wang et al., 2023). As a result, computer vision techniques powered by deep learning have emerged as transformative solutions offering high-resolution detection capabilities adaptable to edge-constrained platforms (Zhang et al.,2021; Sung et al.,2020; Yang et al.,2024).

Among deep learning models, the You Only Look Once (YOLO) family has gained prominence due to its balance of detection accuracy and real-time inference speed. Earlier versions such as YOLOv3 and YOLOv5 have demonstrated strong performance in maritime object detection tasks including vessel classification, debris monitoring and surface anomaly detection. Recent advances in YOLO-series models have revolutionized real-time object detection for maritime applications achieving superior speed-accuracy trade-offs. For instance, YOLOv5 and YOLOv7 variants have demonstrated mAP improvements of 10-15% on water surface targets through lightweight architectures and attention mechanisms tailored for USVs (Al-Hattab et al., 2023; Li et al., 2023; Yang et al., 2024). In other hands, recent advancements such as YOLOv8 and YOLOv11 introduce transformer-based feature extraction, improved bounding box regression and enhanced training stability that enabling more robust performance in environments with high object variability and clutter (Khanam et al., 2024; Yu et al., 2024). Despite these improvements, most pretrained models rely on large-scale generalized datasets such as COCO which do not adequately represent domain-specific maritime features such as rare obstacles like variety shapes of ship, structure, aquatic plant and floating debris (Kim et al., 2022; Wang et al., 2023). This gap highlights the need for dataset customization and fine-tuning to enhance contextual recognition ‘performance.

To address the dataset scarcity challenge, recent studies have incorporated transfer learning and dataset augmentation tools such as Roboflow to construct annotated datasets tailored to specific detection environments. Fine-tuning pretrained models on domain-specific datasets has been shown to significantly improve detection accuracy while requiring limited computational cost compared to training from scratch (Lin et al., 2021; Reddy & Basha., 2025). However, deployment of such models often remains limited to high-performance desktop environments and only a few studies demonstrate fully embedded implementation suitable for real-time field deployment (Haijoub et al., 2024).

Hardware acceleration plays a critical role in the feasibility of onboard AI deployment for maritime navigation. The NVIDIA Jetson platform is increasingly adopted for edge inference because it provides GPU-accelerated parallel processing capabilities with a relatively low power footprint and enabling real-time object detection in resource-constrained environments (Signaroli et al., 2025). When integrated with stereoscopic cameras, embedded perception systems have the potential not only to detect objects but also to infer spatial relationships such as distance, object depth and potential collision risk (Al-Hattab et al., 2023). This capability is essential in maritime navigation, where floating objects, drifting debris or static obstacles such as piers and rocky structures may pose direct hazards.

Given the technological landscape and ongoing research needs, this study focuses on developing an AI-driven water surface obstacle detection system using YOLOv11 fine-tuned with a custom maritime dataset. The dataset was prepared using Roboflow and annotated into four key classes relevant to operational maritime environments: boat, floating debris, structure, and aquatic plant. The model was trained using a transfer learning strategy based on pretrained COCO weights and deployed on a Jetson Orin NX platform integrated with a stereo vision system. The primary objective of this study is to evaluate the performance and feasibility of the embedded system for real-time obstacle detection in varying field conditions. The approach aims to bridge the gap between controlled research environments and deployable real-world maritime applications while contributing toward safer autonomous navigation systems for inland waterways, coastal zones and marine monitoring operations.

Methodology

Dataset Preparation and Model Fine-Tuning

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 1 Image of labeled data set with class

Source : Author’s Data, 2025

The pretrained YOLOv11 model with COCO weights was selected as the baseline architecture due to its balance of detection accuracy and computational efficiency. Figure 1 show label trained image. Training was conducted in Python 3 using the PyTorch-based Ultralytics Framework.

Hardware Configuration

For real-time field deployment, the prototype systemwas deployed on an NVIDIA 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. 2. Additionally, an external alarm strobe module was integrated to provide real-time warning signals during obstacle detection event.

System Architecture

The system architecture comprises four primary layers: sensing, processing, decision and output. The sensing layer acquires real-time RGB and depth data using a ZED 2i stereo camera. The data is processed by an Jetson Orin NX running the optimized fine-tuned YOLOv11 model. In the decision layer, detections are evaluated based on bounding box confidence, distance estimation and predefined safety thresholds. When confidence and proximity criteria are met within 20 meter, the output layer triggers visual and alert notifications.

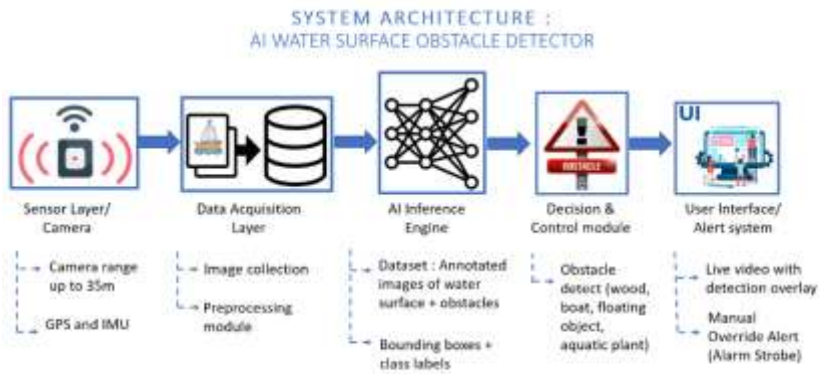


Figure 2 Prototype setup

Source : Author's Data, 2025

Figure 3 System architecture

Source : Author's Data, 2025

The output/alert layer consists of a marine-grade IPX6 LCD module for real-time visualization and a high-intensity alarm strobe module for hazard warning. All components including the Jetson module and camera interface are housed within an IPX6-rated waterproof enclosure to ensure durability in outdoor and marine environments during field deployment. Figure 3 shows the system architecture for present work.

Evaluation Protocol

Model performance was evaluated on unseen test set (38 images) using COCO-standard metrics: mAP@0.5 (primary), mAP@0.5:0.95, Precision-Recall curves, F1-score and per-class Average Precision. Confusion matrices visualized false positive/negative patterns across confidence thresholds (0.1-0.9).

Results and Discussion

The performance of the fine-tuned YOLOv11 model was evaluated using multiple metrics, including mean Average Precision (mAP), precision–recall trends, confusion matrix analysis, and confidence-based scoring. Table 1 summarizes the results for both validation and test sets. Qualitative assessment using field trial images was also conducted to verify real-world inference performance. The following section presents and discusses these evaluation outcomes.

Table 1 Model Performance Summary

Metric	Validation	Test
mAP@0.5	0.7431	0.7561
mAP@0.5:0.95	0.424	0.4387
Precision	0.8088	0.7988
Recall	0.6997	0.6586
F1-score	0.7503	0.722
Optimal confidence threshold	0.446	0.446

Source : Author's Data, 2025

Confusion Matrix Interpretation

The normalized confusion matrix (Fig. 4) 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.

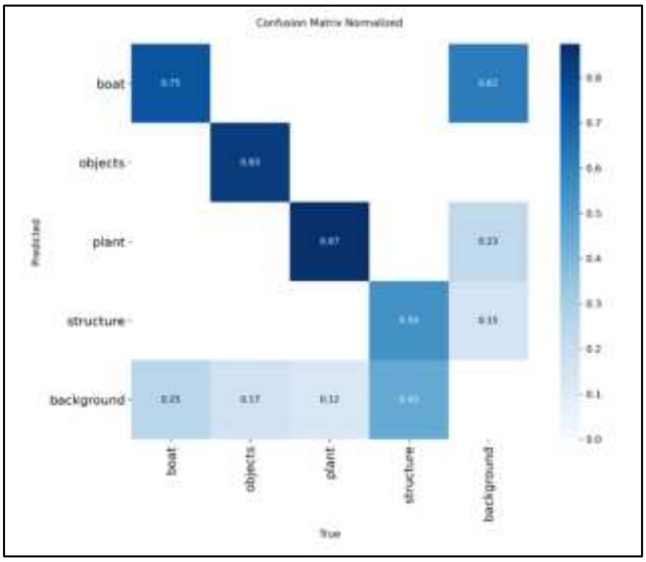


Figure 4 Confusion matrix normalized

Source : Author's Data, 2025

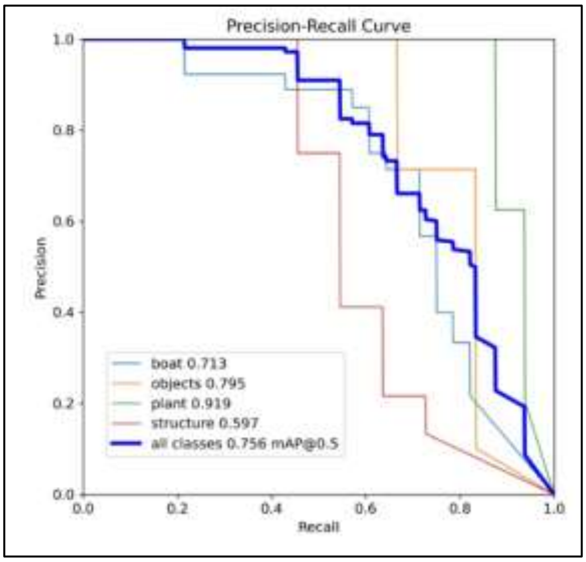


Figure 5 Precision-Recall curve

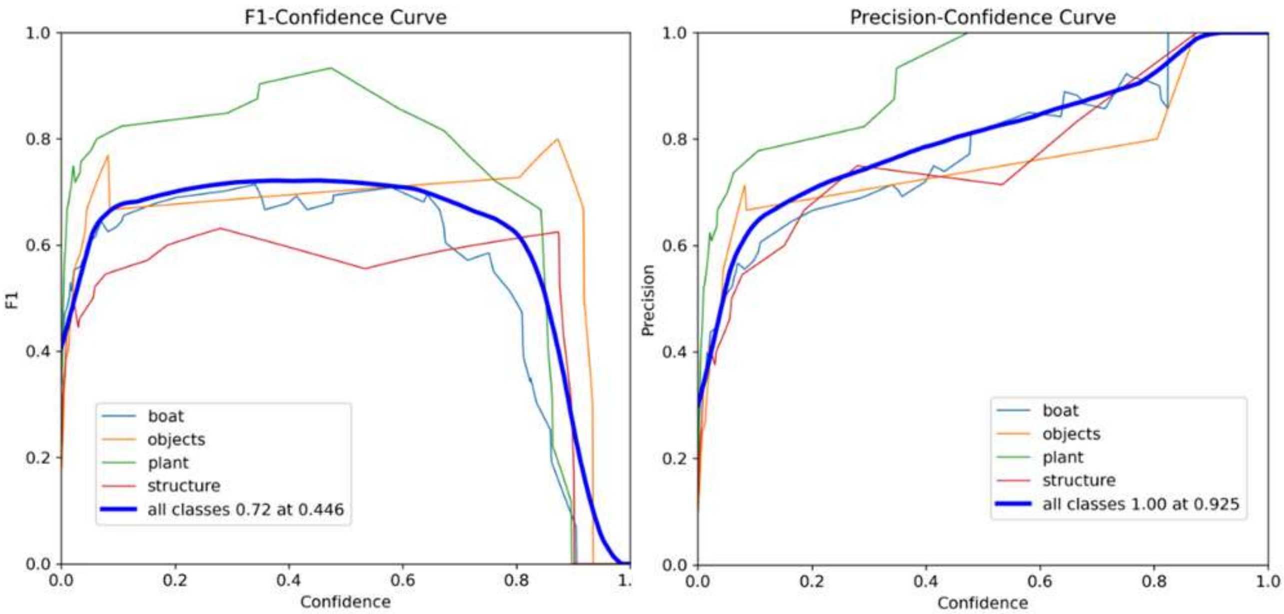
Source : Author's Data, 2025

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.

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. 5. The F1-Confidence curve (see Fig. 6a) 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.6(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.



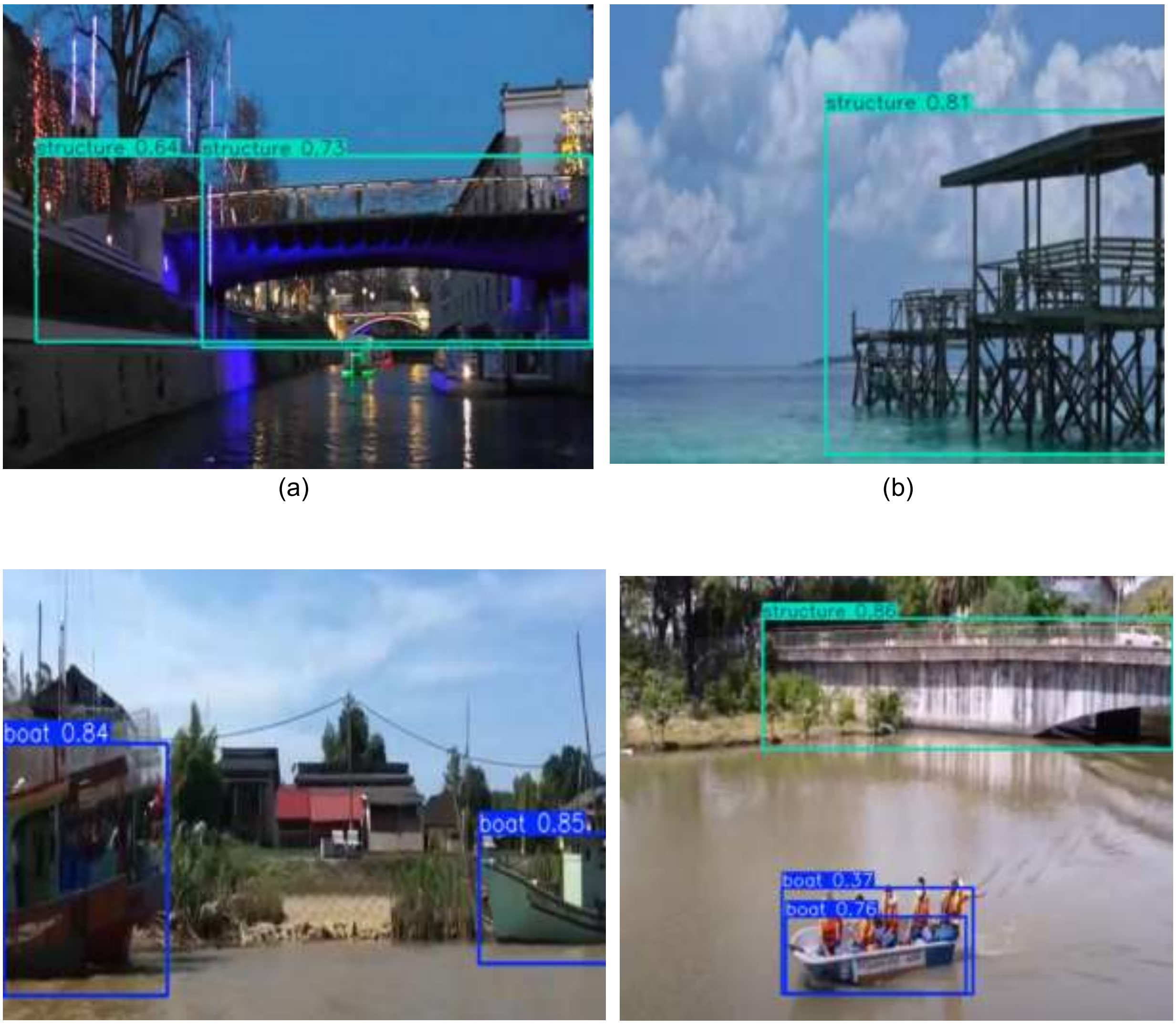
(a) (b)

Figure 6 Confidence curve a) F1-Confidence and b) Precision-Confidence

Source : Author's Data, 2025

Qualitative Detection Results

Inference sample in Fig. 7(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.



(a) (b)

(c) (d)

Figure 7 Detection of a)brige and structure, b)pier structure, c) boat and d) moving boats and static structure

Source : Author’s Data, 2025

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.

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.

Acknowledgement

The authors extend sincere appreciation to the T-AIGS research team of Politeknik Tun Syed Nasir, Johor, and the Centre for Unmanned Technologies (CUTe), International Islamic University Malaysia (IIUM), Selangor, for their collaboration, technical assistance, and continuous support throughout this project. The authors also gratefully acknowledge the funding support from Jabatan Pengajian Politeknik dan Kolej Komuniti (JPPKK) under the TVET-Applied Innovation Grant Scheme (T-AIGS), which made the development, testing and validation of this research possible.

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

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

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