IJRCS

ASCEE

Real-Time Obstacle Detection for Unmanned Surface Vehicle Maneuver

Anik Nur Handayani [a,1,*], Ferina Ayu Pusparani [a,2], Dyah Lestari [a,3], I Made Wirawan [a,4], Aji Prasetya Wibawa [a,5], Osamu Fukuda [b,1]

[a] Universitas Negeri Malang, Jalan Semarang 5, Malang, 65145, Indonesia
[b] Saga University, 1 Honjomachi, Saga, 840-8502, Japan
[1] aniknur.ft@um.ac.id; [2] ferina.ayu.1705366@students.um.ac.id; [3] dyah.lestari.ft@um.ac.id;
[4] made.wirawan.ft@um.ac.id; [5] aji.prasetya.ft@um.ac.id; [6] fukudao@cc.saga-u.ac.jp
* Corresponding Author

ARTICLE INFO

ABSTRACT

The rapid advancement and increasing demand for Unmanned Surface Vehicle (USV) technology have drawn considerable attention in various sectors, including commercial, research, and military, particularly in marine and shallow water applications. USVs have the potential to revolutionize monitoring systems in remote areas while reducing labor costs. One critical requirement for USVs is their ability to autonomously integrate Guidance, Navigation, and Control (GNC) technology, enabling self-reliant operation without constant human oversight. However, current study for USV shown the use of traditional method using color detection which is inadequate to detect object with unstable lighting condition. This study addresses the challenge of enabling Autonomous Surface Vehicles (ASVs) to operate with minimal human intervention by enhancing their object detection and classification capabilities. In dynamic environments, such as water surfaces, accurate and rapid object recognition is essential. To achieve this, we focus on the implementation of deep learning algorithms, including the YOLO algorithm, to empower USVs with informed navigation decision-making capabilities. Our research contributes to the field of robotics by designing an affordable USV prototype capable of independent operation characterized by precise object detection and classification. By bridging the gap between advanced visualization techniques and autonomous USV technology, we envision practical applications in remote monitoring and marine operations with object detection. This paper presents the initial phase of our research, emphasizing significance of deep learning algorithms for enhancing USV navigation and decision-making in dynamic environmental conditions, resulting in mAP of 99.51%, IoU of 87.80%, error value of the YOLOv4-tiny image processing algorithm is 0.1542.

1. Introduction

Real-time monitoring and data collection of water surfaces play a pivotal role in enabling sophisticated data analysis within unfamiliar environments. This necessitates the observation of object detection, encompassing the identification of material objects, recognition and classification of rapid-

moving objects, and determination of water levels and distances to detected objects [1]. The transition from digitization to informatization and onward to intelligence characterizes the evolution of data collection and interpretation in sampling environments.

In the realm of water environment monitoring, there is a growing role for vehicles and transportation systems embedded with artificial intelligence [2][3][4][5]. Recent advancements in machine learning, particularly deep learning approaches, have emerged as potent tools in the development of intelligent transportation systems [6][7][8]. These methodologies find application across various facets of the maritime industry, encompassing boat classification [9], object detection, collision avoidance, risk perception, and anomaly detection. Maritime surveillance and autonomous boat navigation stand out as the primary application domains. The deployment of autonomous systems holds the potential to rapidly and safely gather environmental information, offering a cost-effective alternative to human-led research vessels, especially in remote and inhospitable locations such as the Arctic [10]. Therefore, sophisticated information technology such as computer vision [11] system should support the intelligent transportation system.

The role of computer vision is pivotal in analyzing unfamiliar environmental features, bridging the realms of robotics and object recognition detection [12]. Computer vision is presently employed to enhance image quality and meet the demands of human visual limitations. It empowers computers to detect objects within images and determine their precise coordinates, finding diverse applications both within and beyond the realm of computer science. Traditional computer vision methods frequently utilize color as a means of object detection, employing algorithms like color correction [13][14], Hue Saturation Value (HSV) [15][16][17][18][19], Hue, Saturation, and Lightness (HSL) [20][21][22], Color Tracking Methods [23], and RGB [24][25]. Within the domain of Unmanned Surface Vehicles (USVs) [4][26][27][28][29][30][31][32], object recognition detection has predominantly relied on the HSV filter as the primary algorithm. However, color-based detection is susceptible to variations in lighting conditions, necessitating a robust solution [33]. This weakness may interrupt the system in recognizing the object while the environment lighting is unstable. Thus, the use of deep learning method use to enhance object detection precision and speed.

Few of the traditional deep learning method have been established such as SIFT, SVM and HOG to extract feature from images then classifying it. However, target detection algorithm based on convolutional neural network impoved to be significantly better in efficiency and accuracy, one of which YOLO are one of best develop object detection algorithm [34]. YOLO offers remarkable accuracy [28][35]-[39], even in the presence of lighting noise [26][32]. The study focuses on real-time obstacle detection, a critical component of USV navigation.

This paper presents a novel solution, leveraging the You Only Look Once (YOLO) algorithm for object recognition on water surfaces, specifically deployed within USVs. In achieving this goal we develops real-time obstacle detection for the USV maneuver. The design and implementation of obstacle detection determine the USV directional navigation. The navigation system is a crucial direction guide information that determines the coordinate location and confidence value of the object that moves the USV actuator to take action based on the detected object.

The design and implementation of this obstacle detection system play a pivotal role in guiding the USV's directional navigation, providing essential coordinate location and confidence values for detected objects. The main contribution benefit from this paper are as follows: (1) YOLOv4-tiny network was applied as the main algorithm of the object detection which greatly improve the accuracy and detection speed of the designated object. (2) Improve dual-hull trajectories and wave resistant using combined IMU in navigating in a dynamic environment.c

In Section 2, we delve into the proposed system design, introducing the YOLO computer vision algorithm [34], the used dataset, USV motor and hull robot design, and the navigation system within the designated testing area. Section 3 presents results related to the detection of designated objects in varying light intensities and discuss the implications of these results for potential applications as an unmanned monitoring tool for water surface areas. Finally, in Section 4, we outline future research directions and areas for further development.

2. Method

2.1. Object Selection

The objects used in this study will be classified into three different classifications using the YOLO detection algorithm labelled as: Boat 1, Boat 2, and obstacle as in Table 1. Color choosen in this study are based on IALA Maritime Buoyage System that shows different color legend for water depth information [40] shown in Fig. 1.

Table 1. Designated object for research purpose

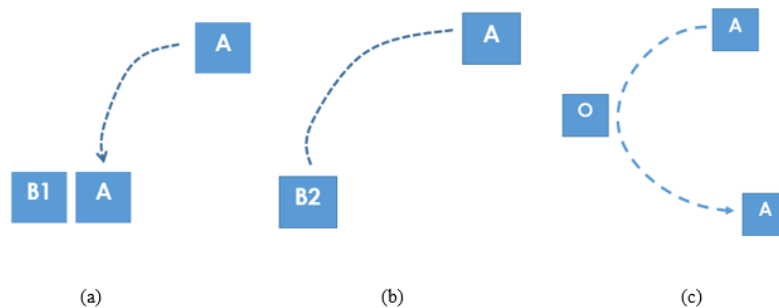| Object	Labelled Object	Object Description
	A	USV
	Boat 1 (B1)	Boat 1 acts as a foreign Boat, if Boat 2 in frame, then USV approached Boat 1
	Boat 2 (B2)	Boat 2 acts as a local Boat, if Boat 1 in frame, then USV avoided Boat 2
	Obstacle (O)	Obstacle acts as a buoy in the sea to indicate the depth of the sea along with the legend of the sea limit.

Fig. 1. Block diagram of USV navigation decision (a) A approach at B1 (b) A stay beside B2 (c) A avoided O

2.2. Proposed Method

Data labeling provides a label to the dataset object that carries out the data training process of the YOLO Algorithm. This operation uses interlinked terms such as object classification, object localization, and recognition. Object classification an object assigned as a class in the existing images assigned as labels. Object Localization [41] is a process of locating the object's position in respect of x,y coordinates in the images or video assigned as a bounding box. The combination of both process object localization and classification is called Object detection. Image labeling is an image annotation process for class classification in a text file extension form. The results of the labeled information include class, center coordinates of the bounding box (x,y), and the height and width of the bounding box (h,w)—the format used in obtaining object position in an image used YOLO format. Fig. 2 describes conversing labeling square into YOLO format in a text file. Images can have 2 or more classes in one image, adding more than one class in the algorithm containing the class of each annotated object. (Coherence with X,Y revise briefly explain X,Y Formula).
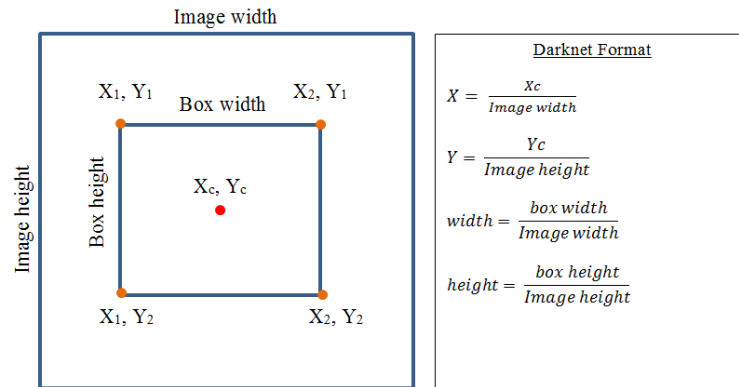
Fig. 2. Bounding box coordinate calculation

The total image retrieval dataset is 3087 images. Based on [42]-[44] the best training and test data ratio used is 80% and 20%. The labeling process provides two formats, namely Pascal VOC with .xml file output and YOLO with .txt file output. In this research, the YOLO format in LabelImg [45] will be used to generate labeling information. The YOLO algorithm uses the results of labeling information to process the training stage. Before the training session, settings are configured in the Pre-Training Stage. Table 2 shows the object in the test data with 671 images. Ground truth is labeled image data to be trained with the YOLO algorithm.

Table 2. Total ground-truth in a dataset

Total Data	Types of Data	Ground-Truth
	Red Boat	308
671	Yellow Boat	219
	Obstacle	180

2.3. YOLO Algorithm Configuration

Configuring the YOLO network requires an initial set of the configuration file and weight in the beginning of the pre-training process. The pre-trained set of weights used is called "yolov4-tiny.weights". The initial weights have been trained through the ImageNet dataset and contain an entire network structure of the YOLO algorithm. All the general parameters are listed in Table 3. Width, and height of the images are set to 416×416, keeping the training time length low. Filters in the configuration are adjusted accordingly to the number of classes used in the dataset. Max Batches is set as how many iteration running through the training process as referred to the article published by Redmon and Fahardi (2018) [46] output layer of the filter follows the equation shown in Table 3.

Table 3. Configuration of YOLO parameter

Parameter	Equations	Configurated Values
Classes	N	3
Batches	-	64
Max Batches	$n \times 2000$	6000
Subdivisions	-	32
Width and Height	$S \times 32$	416×416
Steps	$80\% \times Max\ Batches$	4800
	$90\% \times Max\ Batches$	5400
Filters	$(n + 5\ ) \times 3$	24
Random	-	1

Fig. 3 shows the general object detection system. Object detection has two processes: feature extraction, classification and localization of an object based on CNN (Convolutional Neural Network). The method used is the development of MultiBox with the addition of layers performing parallel processes with regression boxes and object classification [23].
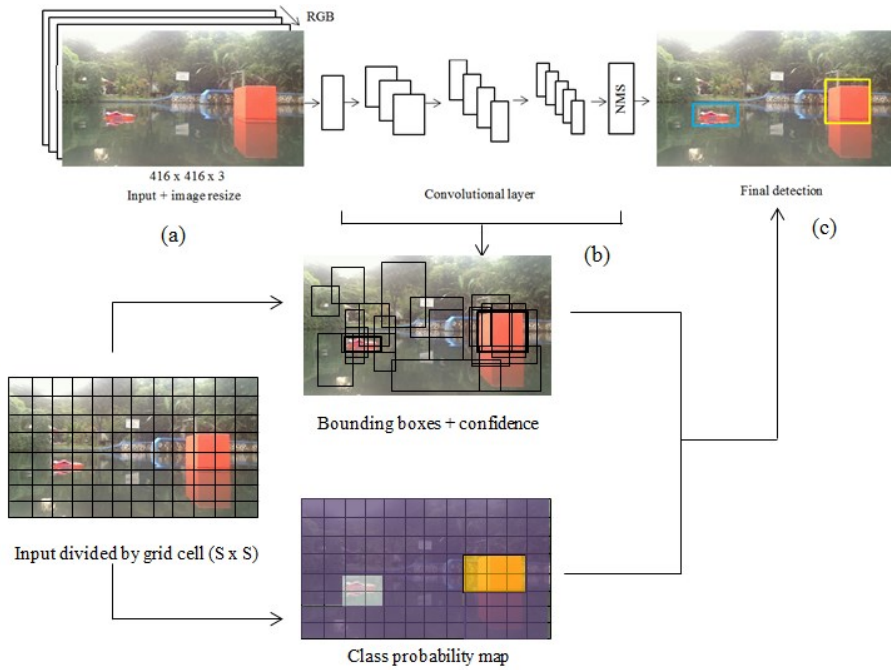
Fig. 3. (a) Raw images used as an input are resized to 416 x 416 x 3, (b) Convolutional layer process, (c) Final detection results

The input image can have an uncertain size. The algorithm has a resize feature according to the pre-training parameters set. The resized pixel size in the image is not limited to how many measurements as long as the value in the division is 32. In image detection, the image will enter the resize first and then be divided into S x S grid cells. Each grid cell has the task of predicting the bounding box. Each bounding box contains object information of x,y coordinates, height (h), width (w), and the predicted value of the object (confidence score). The threshold is given to eliminate bounding boxes that have low accuracy values. Each grid cell will predict a probability class C that predicts one object class only despite the number of bounding boxes in the grid cell. The confidence score of the detected object uses a formula by multiplying the conditional class C with the confidence score of each bounding box generated using the equation (1) [47].

$$Pr(Classi|Object) \times Pr(Object) \times IOU_{pred}^{truth} = Pr(Classi) \times IOU_{pred}^{truth} \qquad (1)$$

2.4. Testing Environment

In providing a USV that serves as a monitoring robot, this prototype will be implemented in a pool with a depth of 4 meters. The resulting data was recorded through live feed camera and connected through a remote monitor using a cross-platform screen-sharing system, enabling users to monitor the surrounding area remotely and efficiently.

The lighting environment within the recreational pool may vary depending on the sunlight throughout the day. Environment used for the testing condition are outdoor which deliberately effect how object detection capture the detected object. Tested have been done within three different daytime: (1) Morning (07:00 – 10.00). (2) Afternoon (10:00-13:00). (3) Evening (13:00-16:00).

2.5. USV Design

High-speed USV is crucial in determining research performance, especially in the preliminary design stages. Preliminary design determines specific hull types used and performance comparison between different hull types. USV design should overlook the following components [48]: viscous resistance, wave making resistance, and body form resistance which consists of pressure drag. The ability to move the USV with many advantages and performance is a must in facing the uncertainty of the environment. Therefore, a new type of marine robot USV design is proposed in this study.

Comparisons to conventional monohull design are used concerning catamaran USV to understand better the performance characteristics of the proposed design [49][50]:

- Better ability in seakeeping. One of the most significant performance abilities is using catamaran USV design using rocking motion in cutting wind and waves while sailing. The complex design of the catamaran boat allows for better performance when navigating in the waves with a minor rocking motion than the monohull design.

- Better hydrostatic resistance performance. Compared to monohull, catamaram have better hydrostatic resistance due to the displacement of the catamaram, mainly on the main body deeper from the water surface. This contributes in reducing the waterline and wave resistance on the water surface.

- Stable stability and rotation performance. Due to the longer design of the two propellers inside the USV enables good steering when taking turns in avoiding obstacles compared to monohull design. Furthermore, the large surface help in minimizing the USV from overturning

- Large surface area. Since the USV carry out multiple sensors and missions loads to perform certain task, large surface area to carry all the equipment need to be in demand and catamaram model design satisfied the needs of it.

Based on several consideration advantages offered, the catamaran design is used in this study. Two propellers are installed within the two hulls of the catamaran ship: the right propeller and the left propeller. Bridges are added between two hulls to load components and sensors, along with a brushless motor for each propeller. In order to monitor the surrounding environment, web-camera is placed on top of the USV. In Fig. 4 The catamaran USV has the advantage of navigating through waves in the water surface, with the efficient operation and low cost offered. The easiness of the catamaran boat can be used for monitoring, search and rescue, and navigation.

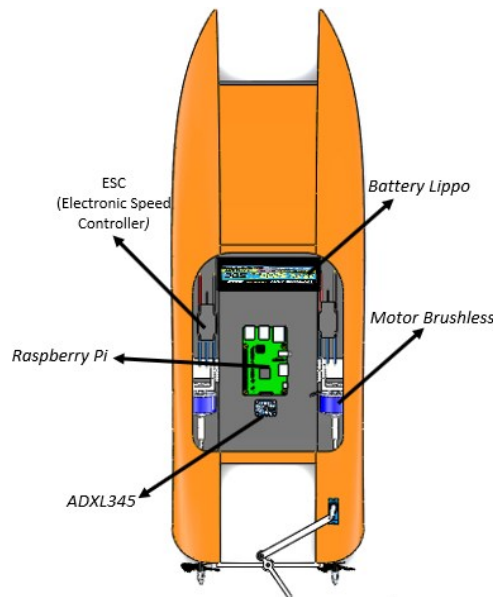**Fig. 4.** Upper view of the catamaran USV

2.6. Maneuver

The actuators of the USV in this study are brushless motors and servos. The brushless motor is the rate driver of the USV while the servo is used to adjust the angle of the turning direction of the USV. Fig. 5 is an illustration of the direction of the USV.

$$x\ center\ =\ x + (w/2)/2 \tag{2}$$

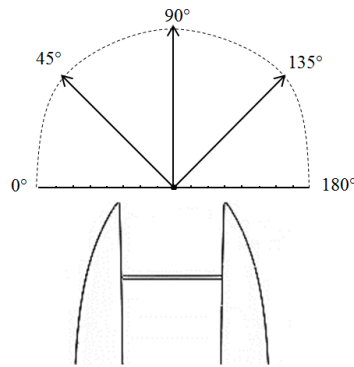$$y\ center\ =\ y + (h/2)/2 \tag{3}$$

Fig. 5. Robot boat navigation servo degree

The actuator is driven by Arduino while Jetson Nano is used as a microprocessor in the image detection process using a webcam camera. Fig. 6 shows all the components used in this research.
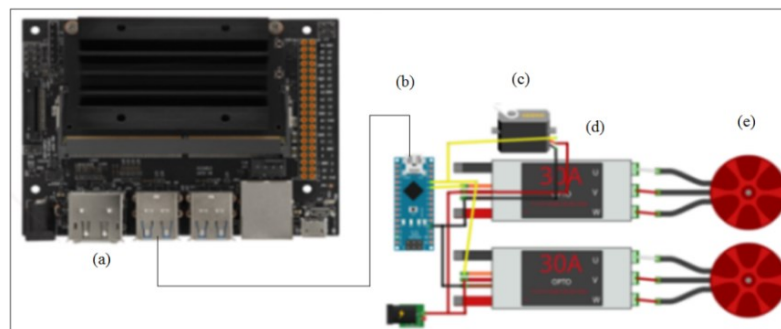
Fig. 6. (a) Jetson Nano (b) Arduino NANO (c) Servo Motor (d) ESC Motor (e) Brushless Motor

Motor navigation on the USV uses the division of several areas on the frame to determine the direction of the servo as a turning angle. The result of reading the coordinates of the frame through the object detection system determines the turning angle and motion of the motor when it is far and near from the detected object. Fig. 7 illustrates the area division of the frame for motor navigation. There are 6 areas divided in one frame, namely: (1) Far Left, (2) Far Center, (3) Far Right, (4) Near Left (5) Near Center, (6) Near Right. The camera pixel size used is 640 x 480 pixels with the division of each area worth one-third of the pixel length and two-thirds of the width for long distance classification.

Fig. 7. Area of division in camera frame

3.   Results and Discussion

3.1. Training Results

Shown in Fig. 8 are the result of the CNN model training after 6000 iteration steps. The mAP after 1800 iterations is 92%. After 6000 iterations, the mAP value hits 99%. On the other hand, the value of Loss is decreasing as the number of iterations increases. This example's occurrence, mAP and Loss value, is typical of learning. To put it another way, the mAP trend runs counter to the loss value. The best trained model is selected as the final trained model to be used as the USV detection models after performance testing of the trained models. The learning model has achieved its best mAP at 1800 iterations, but the present iteration's loss value is still large, and the learning model's greatest results were obtained at 6000 iterations which is shown in Table 4.
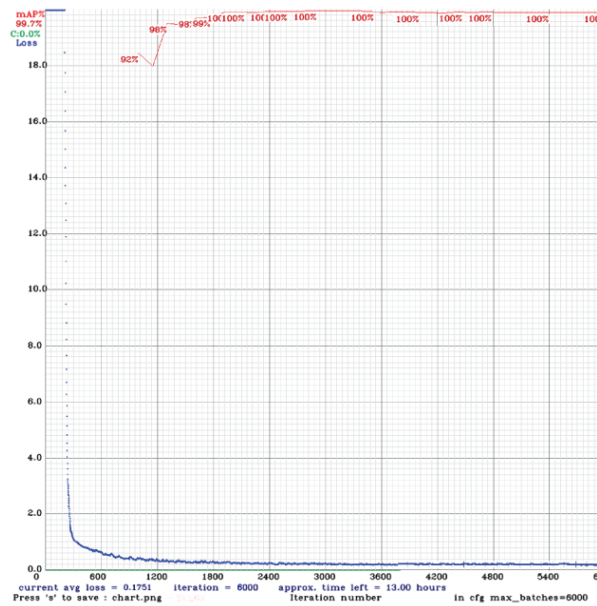

Fig. 8. Training graph

Table 4. Loss value of each epoch

Epoch	Loss
1000	0.3190
2000	0.2478
3000	0.2768
4000	0.2157
5000	0.1544
6000	0.1542

3.2. Distance and Detection Accuracy

Navigation testing is done by considering two variable which are light and distance variables. The light level is measured with a luxmeter at three different times: morning (07.00 - 10.00), afternoon (10.00 - 13.00), and evening (13.00 - 16.00). Distance measurements are measured between the object and the USV by 10cm to 120cm. Testing is done with three different objects while observing the USV navigation response. The results of object detection are the average of five experimental testing with various distance variations. Table 5, Table 6, and Table 7 are the results of navigation experiments based on the light level of each object.

At distances greater than 10 cm and up to 90 cm, Boat 1 produces average accuracy that is greater than 80%. In Table 5, the detection has a poor average of accuracy at a distance of 10 cm. At a distance of 100 cm to 120 cm, detection on the camera became less accurate, with an accuracy below 70%. While detection with an average above 50% is at a distance of 40cm to 100cm with the best detection

accuracy located at a distance of 90cm, Boat 2 has an accuracy below 50% at a distance of 10cm to 30 cm. Accuracy drops to less than 50% at distances greater than 100 cm. While testing from 30 cm to 110 cm has a strong detection accuracy level that exceeds 80%, obstacle objects in the test results at a distance of 10 cm to 20 cm the camera is too close to the item such that numerous bounding boxes appear at once.

It makes sense that accuracy would decline as distance grew. However, Table 5 result demonstrates that the accuracy varies. The circumstance demonstrates that the relationship between distance and accuracy cannot be seen as linear. The accuracy of the detection may be impacted by additional variables like the angle of the boat relative to the sea and wind. (–) is the bounding box unable to locate in x distance.

Table 5. Object detection morning time

Lux Meter	Distance (cm)	Accuracy (5 trials)		
		Boat1	**Boat2**	**Obstacle Object**
1353	10	63.8	43.5	-
	20	80.4	34.8	-
	30	92.3	44.2	93.6
	40	97	52.8	94.4
	50	95.4	60.4	96.4
	60	90.8	75.2	80.5
	70	91.8	68	91.6
	80	80.2	71.6	91.6
	90	83.5	83.75	86.25
	100	72.4	76	89.4
	110	63.2	39.2	83.4
	120	50.4	48.4	65.4

In Table 6 Boat 1 has a sufficient level of accuracy at a distance of 10cm to 20 cm. Good detection accuracy at the 80% level is produced at a distance of 30 cm to 110 cm. At a distance of 120 cm the camera detection decreased with an average value of 75%. In the second object test Boat 2 camera detection on the object experienced a sufficient level of accuracy with an average of 50% at a distance of 30 cm to 120 cm while the distance of 10 cm and 20 cm accuracy decreased below 30% due to the size of the frame on the camera for the Boat 2 object is too small so that the ship object is not fully 100% captured by the camera. Testing the third object, namely the Obstacle object, has an average accuracy level above 80% from a distance of 30 cm to 110 cm at a distance of 10-20 cm the algorithm experiences several repeated detections because the object is too close so that many bounding boxes appear. (–) is the bounding box unable to locate in x distance.

Table 6. Object detection afternoon time

Lux Meter	Distance (cm)	Accuracy (5 trials)		
		Boat1	**Boat2**	**Obstacle Object**
3795	10	61.2	27	-
	20	57.8	26.6	-
	30	86	56	94.2
	40	91.4	59.2	93.8
	50	91.2	48.8	94.6
	60	95.6	58.8	94
	70	92	54.2	90.4
	80	98.8	54.6	89
	90	90.4	52.6	88
	100	81.2	51	90.6
	110	66.8	60	84.6
	120	75.8	59.2	75.2

Table 7 shows the result of the object detection testing with each object in the evening time conditions. Data collection was carried out at 13:00 - 16:00 within 5 time trials at each distance

testing. After testing the detection accuracy at 13:00 - 16:00, Boat 1 result shows low average accuracy at a distance of 10 cm. Detection result from distance of 20 cm – 100 cm result are average above 80%. Camera detection decreased at a distance of 110 cm – 120 cm. On Boat 2 the detection accuracy of the camera decreases with an average below 30% at a distance btween 10 cm to 20 cm. At a distance of 30 cm – 120 cm the detection accuracy results are average of 50%. The obstacle object test results at a distance of 10 cm – 20 cm are unreadable due to multiple bounding box overlapping the intended object detection in a frame. At a testing distance of 30 cm – 120 cm the camera accuracy with an average of 80%. The light intensity measurements was taken on 13:00 - 16:00 were 1758 lux. (–) is the bounding box unable to locate in X distance.

Table 7. Object Detection Evening Time

Lux Meter	Distance (cm)	Accuracy (5 trials)		
		Boat1	**Boat2**	**Obstacle Object**
	10	54.6	24	-
	20	85.8	24.6	-
	30	91.8	45.6	89.8
	40	90.2	40.8	91.2
	50	87.6	51	89.4
	60	90.6	48.2	90.6
1353	70	86.8	44	92.2
	80	90.4	59	91.4
	90	87.6	60.8	92.8
	100	92	57.8	92.6
	110	79.6	59.4	95
	120	79.6	54.6	81.4

3.3. USV Performance

In the realistic environment, changes of light as well as water droplets from the water would influence the target detection performance. Each detection results from live feed are reviewed manually. Detection results from three different time period are illustrated in Fig. 9.
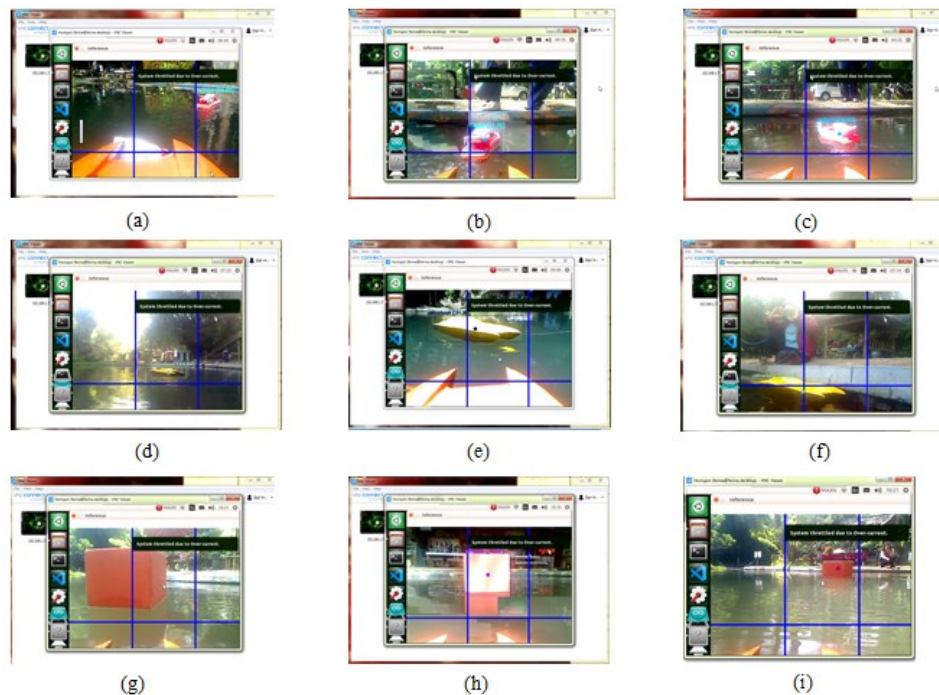

Fig. 9. Object detection result under different environment (a-c) Boat 1 Detection from morning, afternoon and evening time (d-f) Boat 2 Detection from morning, afternoon and evening time (g-i) Obstacle object detection from morning, afternoon and evening time

The upper, middle and lowe row of Fig. 9 listed above are the detection results achieved under different environmental condition using YOLOv4-tiny algorithm. During the experiment reflective light occured on the afternoon time Fig. 9 (b), (e), (h). The YOLOv4-tiny has potential to properly classified the designated object with high accuracy despite the environmental condition. However, the experimental results demonstrate that the proposed YOLOv4-tiny is unstable agaits changes of the environmental condition despite its tolerance to the USV speed.

3.4. Distance and Manuever Accuracy

To evaluate the servo degree, detection accuracy and navigation response of The USV, distance and manuever accuracy test is conducted. The detection targets from the live-feed camera are put into different location on the frame. Table 8 shows the result of the of maneuver testing at a distance between 10 – 120 cm. The result shown at 10 cm distance the USV correctly take action to stop at 10 cm with 100% accuracy with 5 time trials. Expected result of the USV between 20 – 120 cm results are the USV need to approach designated object which detected by the camera. The result shown the optimal distance for the object to be detected are between 70 – 90 cm which resulted a 100% accuracy. Compare to Table 9 at a distance 10 cm the USV also have the same expected response which to stop at 10 cm with 100% accuracy. Optimal detection distance along with the navigation response of the USV with Boat 2 object are achieved with a distance at 90 – 100 cm, resulting in 100% accuracy. Lastly in Table 10 at a distance 10cm the USV response is to stop in front of the object which from the result of the overlapping bounding box. This effect the detection classification since it is unrecognizable by the algorithm and the USV response is to stop. If the USV detect the ostacle object above 10 cm the USV would recognize the obstacle object and take action to turn and avoid. Based on the experiment the optimal distance achieved for the USV to take turn for the obstacle object is at a distance of 40 – 110 cm. A is the The USV stop, B is the The USV approach an object and C is the The USV approach and avoid an object.

Table 8. Boat 1 detection navigation accuracy based on servos degree

Distance (cm)	0°	90°	180°	Accuracy (%)
10	A (1)	A (1)	A (1)	100%
20	A (1)	A (0.2)	A (0.2)	100%
30	B (0.6)	B (0.6)	B (0.8)	67%
40	B (0.6)	B (0.6)	B (0.8)	67%
50	B (0.6)	B (1)	C (1)	86%
60	B (0.8)	B (1)	C (1)	93%
70	B (1)	B (1)	C (1)	100%
80	B (1)	B (1)	C (1)	100%
90	B (1)	B (1)	C (1)	100%
100	B (1)	B (0.8)	C (1)	93%
110	B (0.8)	B (0.4)	C (1)	73%
120	B (0.4)	B (0)	C (0)	13%

Table 9. Boat 2 detection navigation accuracy based on servos degree

Distance (cm)	0°	90°	180°	Accuracy (%)
10	A (1)	A (1)	A (1)	100
20	A (0)	A (0.4)	A (0.2)	26.7
30	B (0.2)	B (0.4)	B (0.4)	33.3
40	B (0.2)	B (0.6)	B (0.4)	33.3
50	B (0.4)	B (0.6)	B (0.8)	60
60	B (0.4)	B (0.6)	B (0.8)	60
70	B (0.6)	B (0.8)	B (1)	80
80	B (0.8)	B (1)	B (1)	93
90	B (1)	B (1)	B (1)	100
100	B (1)	B (1)	B (1)	100
110	B (0.6)	B (0.6)	B (0.6)	60
120	B (0.2)	B (0.4)	B (0)	20

Table 10. Obstacle object detection navigation accuracy based on servos degree

Distance (cm)	0°	90°	180°	Avg Accuracy (%)
10	A (1)	A (1)	A (1)	100%
20	A (0)	A (0)	A (0)	100%
30	A (0)	A (0)	C (0.8)	26%
40	C (1)	C (1)	C (1)	100%
50	C (1)	C (1)	C (1)	100%
60	C (1)	C (1)	C (1)	100%
70	C (1)	C (1)	C (1)	100%
80	C (1)	C (1)	C (1)	100%
90	C (1)	C (1)	C (1)	100%
100	C (1)	C (1)	C (1)	100%
110	C (1)	C (1)	C (1)	100%
120	C (1)	C (0.8)	C (0.8)	86%

4. Conclusion

In conclusion, this study introduced a novel YOLOv4-tiny detection method for the real-time identification of water surface targets in diverse environmental conditions. The results clearly demonstrate the superiority of our approach over conventional color detection methods, with a remarkable mean Average Precision (mAP) of 99.51% and an Intersection of Union (IoU) of 87.80%. By employing a 2MP Logitech webcam-type camera as a visual tool in conjunction with YOLOv4-tiny, we achieved an impressively low error value of 0.1542. This innovative system equips Unmanned Surface Vehicles (USVs) with the ability to effectively operate amidst unpredictable environmental changes, significantly enhancing their detection adaptability. Looking forward, our research roadmap includes: (1) The integration of GPS and the development of an advanced navigation algorithm to empower the USV's path planning capabilities during detection missions. (2) We are committed to optimizing our detection model to expedite training and elevate the overall performance of the USV. The implications of this study extend far beyond the realm of autonomous navigation. Our work paves the way for safer maritime operations, precise environmental monitoring, and enhanced scientific research capabilities. The unpredictable factor such as varying light, winds, and water flows are added as future work. As we step into the future, we are excited about the transformative potential of this research in revolutionizing the field of unmanned surface vehicle technology.

References

[1] R. K. Chandana and A. C. Ramachandra, "Real time object detection system with YOLO and CNN models: A review," *arXiv Prepr. arXiv2208.00773*, 2022, https://doi.org/10.48550/arXiv.2208.00773.

[2] Y.-L. Chang, A. Anagaw, L. Chang, Y. Wang, C.-Y. Hsiao, and W.-H. Lee, "Ship Detection Based on YOLOv2 for SAR Imagery," *Remote Sens.*, vol. 11, no. 7, p. 786, Apr. 2019, https://doi.org/10.3390/rs11070786.

[3] W. Zhang, X. Gao, C. Yang, F. Jiang, and Z. Chen, "A object detection and tracking method for security in intelligence of unmanned surface vehicles," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, no. 3, pp. 1279–1291, 2022, https://doi.org/10.1007/s12652-020-02573-z.

[4] L. Cheng *et al.*, "Water Target Recognition Method and Application for Unmanned Surface Vessels," *IEEE Access*, vol. 10, pp. 421–434, 2022, https://doi.org/10.1109/ACCESS.2021.3138983.

[5] X. Zhao, P. Sun, Z. Xu, H. Min, and H. Yu, "Fusion of 3D LIDAR and Camera Data for Object Detection

in Autonomous Vehicle Applications," *IEEE Sens. J.*, vol. 20, no. 9, pp. 4901–4913, 2020, https://doi.org/10.1109/JSEN.2020.2966034.

[6] T. Zhang, X. Liu, Y. Lib, and M. Zhang, "Sea-surface object detection scheme for USV under foggy environment," *Indian J. Geo-Marine Sci.*, vol. 50, no. 11, pp. 960–968, 2022, https://doi.org/10.56042/ijms.v50i11.66765.

[7] C. Barrera, I. Padron, F. S. Luis, and O. Llinas, "Trends and challenges in unmanned surface vehicles (Usv): From survey to shipping," *TransNav Int. J. Mar. Navig. Saf. Sea Transp.*, vol. 15, 2021, https://doi.org/10.12716/1001.15.01.13.

[8] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues," *Array*, vol. 10, p. 100057, 2021, https://doi.org/10.1016/j.array.2021.100057.

[9] G. Li, Z. Song, and Q. Fu, "Small boat detection for radar image datasets with yolo V3 network," in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*, pp. 1–5, 2019, https://doi.org/10.1109/ICSIDP47821.2019.9173163.

[10] Y. Chen, X. Chen, J. Zhu, F. Lin, and B. M. Chen, "Development of an Autonomous Unmanned Surface Vehicle with Object Detection Using Deep Learning," in *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, pp. 5636–5641, 2018, https://doi.org/10.1109/IECON.2018.8591129.

[11] J. Taipalmaa, N. Passalis, and J. Raitoharju, "Different Color Spaces In Deep Learning-Based Water Segmentation For Autonomous Marine Operations," in *2020 IEEE International Conference on Image Processing (ICIP)*, pp. 3169–3173, 2020, https://doi.org/10.1109/ICIP40778.2020.9190960.

[12] H. S. Cunha *et al.*, "Water tank and swimming pool detection based on remote sensing and deep learning: Relationship with socioeconomic level and applications in dengue control," *PLoS One*, vol. 16, no. 12, p. e0258681, 2021, https://doi.org/10.1371/journal.pone.0258681.

[13] D. Hema and D. S. Kannan, "Interactive color image segmentation using HSV color space," *Sci. Technol. J*, vol. 7, no. 1, pp. 37–41, 2019, https://doi.org/10.22232/stj.2019.07.01.05.

[14] X. Wang, J. Yang, P. Ruan, and P. Wang, "An Improved Unsupervised Color Correction Algorithm for Underwater Image," in *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1215–1220, 2021, https://doi.org/10.1109/ICIEA51954.2021.9516076.

[15] W. Jung, J. Woo, and N. Kim, "Recognition of the light buoy for scan the code mission in 2016 Maritime RobotX Challenge," in *2017 IEEE Underwater Technology (UT)*, pp. 1–5, 2017, https://doi.org/10.1109/UT.2017.7890288.

[16] A. Al-Kaff, F. M. Moreno, A. de la Escalera, and J. M. Armingol, "Intelligent vehicle for search, rescue and transportation purposes," in *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*, pp. 110–115, 2017, https://doi.org/10.1109/SSRR.2017.8088148.

[17] S. Yakin, T. Hasanuddin, and N. Kurniati, "Application of content based image retrieval in digital image search system," *Bull. Electr. Eng. Informatics*, vol. 10, no. 2, pp. 1122–1128, 2021, https://doi.org/10.11591/eei.v10i2.2713.

[18] M. A. Hossain, M. I. Hossain, M. D. Hossain, N. T. Thu, and E.-N. Huh, "Fast-D: When non-smoothing color feature meets moving object detection in real-time," *IEEE Access*, vol. 8, pp. 186756–186772, 2020, https://doi.org/10.1109/ACCESS.2020.3030108.

[19] R. Prados, R. García, N. Gracias, L. Neumann, and H. Vågstøl, "Real-time fish detection in trawl nets," in *OCEANS 2017 - Aberdeen*, pp. 1–5, 2017, https://doi.org/10.1109/OCEANSE.2017.8084760.

[20] Y. Liu, Z. Bellay, P. Bradsky, G. Chandler, and B. Craig, "Edge-to-fog computing for color-assisted moving object detection," in *Big Data: Learning, Analytics, and Applications*, vol. 10989, pp. 9–17, 2019, https://doi.org/10.1117/12.2516023.

[21] S.-H. Tsai and Y.-H. Tseng, "A novel color detection method based on HSL color space for robotic soccer competition," *Comput. Math. with Appl.*, vol. 64, no. 5, pp. 1291–1300, 2012, https://doi.org/10.1016/j.camwa.2012.03.073.

[22] J. K. Basak, B. G. K. Madhavi, B. Paudel, N. E. Kim, and H. T. Kim, "Prediction of total soluble solids

and pH of strawberry fruits using RGB, HSV and HSL colour spaces and machine learning models," *Foods*, vol. 11, no. 14, p. 2086, 2022, https://doi.org/10.3390/foods11142086.

[23] Y. Ardiyanto, I. T. Sujoko, W. A. Wibowo, V. D. H. Nugraha, and F. E. Saputra, "Prototype design of unmanned surface ship to detect illegal fishing using solar power generation technology," *J. Electr. Technol. UMY*, vol. 3, no. 1, pp. 14–18, 2019, https://doi.org/10.18196/jet.3149.

[24] D. Chowdhury, S. Mandal, D. Das, S. Banerjee, S. Shome, and D. Choudhary, "An Adaptive Technique for Computer Vision Based Vehicles License Plate Detection System," *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, pp. 1-6, 2019, https://doi.org/10.1109/OPTRONIX.2019.8862406.

[25] O. Surinta and S. Khruahong, "Tracking people and objects with an autonomous unmanned aerial vehicle using face and color detection," in *2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON)*, pp. 206–210, 2019, https://doi.org/10.1109/ECTI-NCON.2019.8692269.

[26] C. Powers, R. Hanlon, and D. Schmale, "Tracking of a Fluorescent Dye in a Freshwater Lake with an Unmanned Surface Vehicle and an Unmanned Aircraft System," *Remote Sens.*, vol. 10, no. 2, p. 81, Jan. 2018, https://doi.org/10.3390/rs10010081.

[27] A. Aqthobilrobbany, A. Handayani, D. Lestari, M. Muladi, R. Andrie, and O. Fukuda, "HSV Based Robot Boat Navigation System," 2020, pp. 269–273, https://doi.org/10.1109/CENIM51130.2020.9297915.

[28] C. Liu *et al.*, "Development of USV Autonomy: Architecture, Implementation and Sea Trials," *Brodogradnja*, vol. 73, no. 1, pp. 89–107, 2022, https://doi.org/10.21278/brod73105.

[29] P. Yang, C. Song, L. Chen, and W. Cui, "Image Based River Navigation System of Catamaran USV with Image Semantic Segmentation," *2022 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*, pp. 147-151, 2022, https://doi.org/10.1109/WRCSARA57040.2022.9903932.

[30] B. Y. Suprapto, M. A. Kurniawan, M. K. Ardela, H. Hikmarika, Z. Husin, and S. Dwijayanti, "Identification of Garbage in the River Based on The YOLO Algorithm," *Int. J. Electron. Telecommun.*, vol. 67, no. 4, Jul. 2023, https://doi.org/10.24425/ijet.2021.137869.

[31] D. Freire, J. Silva, A. Dias, J. M. Almeida, and A. Martins, "Radar-based target tracking for Obstacle Avoidance for an Autonomous Surface Vehicle (ASV)," *OCEANS 2019 - Marseille*, pp. 1-6, 2019, https://doi.org/10.1109/OCEANSE.2019.8867477.

[32] Z. Gan, J. Zheng, Z. Jiang, and R. Lu, "Reward Shaping-based Double Deep Q-networks for Unmanned Surface Vessel Navigation and Obstacle Avoidance," *IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1-6, 2022, https://doi.org/10.1109/IECON49645.2022.9968970.

[33] N. Hassan, K. W. Ming, and C. K. Wah, "A Comparative Study on HSV-based and Deep Learning-based Object Detection Algorithms for Pedestrian Traffic Light Signal Recognition," in *2020 3rd International Conference on Intelligent Autonomous Systems (ICoIAS)*, pp. 71–76, 2020, https://doi.org/10.1109/ICoIAS49312.2020.9081854.

[34] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263–7271, 2017, https://doi.org/10.1109/CVPR.2017.690.

[35] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020, https://doi.org/10.48550/arXiv.2004.10934.

[36] L. I. Bo, X. I. E. Xiaoyang, W. E. I. Xingxing, and T. A. N. G. Wenting, "Ship detection and classification from optical remote sensing images: A survey," *Chinese Journal of Aeronautics*, vol. 34, no. 3, pp. 145-163, 2021, https://doi.org/10.1016/j.cja.2020.09.022.

[37] A. Aqthobilrobbany, A. N. Handayani, D. Lestari, Muladi, R. A. Asmara, and O. Fukuda, "HSV Based Robot Boat Navigation System," *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pp. 269-273, 2020, https://doi.org/10.1109/CENIM51130.2020.9297915.

[38] A. N. Handayani, F. A. Pusparani, D. Lestari, I. M. Wirawan, O. Fukuda, and A. Aqthobirrobbany, "Robot Boat Prototype System Based on Image Processing for Maritime Patrol Area," in *2022 2nd International*

Seminar on Machine Learning, Optimization, and Data Science (ISMODE), pp. 310–315, 2022, https://doi.org/10.1109/ISMODE56940.2022.10181007.

[39] M. Sportelli *et al.*, "Evaluation of YOLO Object Detectors for Weed Detection in Different Turfgrass Scenarios," *Appl. Sci.*, vol. 13, no. 14, 2023, https://doi.org/10.3390/app13148502.

[40] M. Pan, Y. Liu, J. Cao, Y. Li, C. Li and C. -H. Chen, "Visual Recognition Based on Deep Learning for Navigation Mark Classification," in *IEEE Access*, vol. 8, pp. 32767-32775, 2020, https://doi.org/10.1109/ACCESS.2020.2973856.

[41] M. Singh and S. Indu, "BGR to HSV based Text Extraction from Manuscripts Using Slidebars," in *2021 Asian Conference on Innovation in Technology (ASIANCON)*, pp. 1–6, 2021, https://doi.org/10.1109/ASIANCON51346.2021.9544674.

[42] V. R. Joseph, "Optimal ratio for data splitting," *Stat. Anal. Data Min.*, vol. 15, no. 4, pp. 531–538, 2022, https://doi.org/10.1002/sam.11583.

[43] S.-A. N. Alexandropoulos, C. K. Aridas, S. B. Kotsiantis, and M. N. Vrahatis, "Multi-Objective Evolutionary Optimization Algorithms for Machine Learning: A Recent Survey," *Approximation and Optimization: Algorithms, Complexity and Applications*, pp. 35–55, 2019, https://doi.org/10.1007/978-3-030-12767-1_4.

[44] S.-A. N. Alexandropoulos, C. K. Aridas, S. B. Kotsiantis, and M. N. Vrahatis, "Multi-Objective Evolutionary Optimization Algorithms for Machine Learning: A Recent Survey BT - Approximation and Optimization : Algorithms, Complexity and Applications," *Approximation and Optimization: Algorithms, Complexity and Applications*, pp. 35–55 2019, https://doi.org/10.1007/978-3-030-12767-1_4.

[45] Y. Zhang, Y. Yuan, Y. Feng, and X. Lu, "Hierarchical and Robust Convolutional Neural Network for Very High-Resolution Remote Sensing Object Detection," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5535-5548, 2019, https://doi.org/10.1109/TGRS.2019.2900302.

[46] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 779–788, 2016, https://doi.org/10.1109/CVPR.2016.91.

[47] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "Scaled-yolov4: Scaling cross stage partial network," in *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*, pp. 13029-13038, 2021, https://doi.org/10.1109/CVPR46437.2021.01283.

[48] G. Shao, Y. Ma, R. Malekian, X. Yan, and Z. Li, "A Novel Cooperative Platform Design for Coupled USV–UAV Systems," *IEEE Trans. Ind. Informatics*, vol. 15, no. 9, pp. 4913–4922, 2019, https://doi.org/10.1109/TII.2019.2912024.

[49] E. Giraldo-Pérez, E. Betancur, and G. Osorio-Gómez, "Experimental and statistical analysis of the hydrodynamic performance of planing boats: A Comparative study," *Ocean Engineering*, vol. 262, p. 112227, 2022, https://doi.org/10.1016/j.oceaneng.2022.112227.

[50] P. Asgharian and Z. H. Azizul, "Proposed Efficient Design for Unmanned Surface Vehicles," *arXiv Prepr. arXiv2009.01284*, 2020, https://doi.org/10.48550/arXiv.2009.01284.