Object Recognition in Pond Environments Using Deep Learning

Suhaila Sari[1], Ng Wei Jie[2], Nik Shahidah Afifi Md Taujuddin[3], Hazli Roslan[4], Nabilah Ibrahim[5], Mohd Helmy Abd Wahab[6]
Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia[1, 2, 3, 5, 6]
Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia[4]

***Abstract*—Complicated underwater environment, such as visibility limitations and illumination conditions pose significant challenges for underwater imaging and its object recognition performance. These issues are especially critical for applications involving autonomous underwater vehicles (AUVs) or robotic systems involved in object recognition tasks during search-and-retrieval operations. Moreover, high-turbidity underwater image datasets, especially for pond environments, remain scarce. Therefore, this study focuses on establishing a pond underwater images dataset and evaluating the deep learning-based object recognition architecture, You Only Look Once Version 5 (YOLOv5), in recognizing multiple objects in respective underwater pond images. The dataset contains self-captured 1116 underwater pond images, which are annotated with LabelImg for object recognition and dataset generation. Under varying depths, camera distances, and object angles, the YOLOv5 reaches a mean accuracy mAP 50-95 of 87.96%, demonstrating its effectiveness for recognizing multiple objects in pond underwater environments.**

***Keywords*—Dataset; deep learning; object recognition; pond; underwater image; YOLO**

I. Introduction

Complicated underwater environments, such as visibility limitations and illumination conditions, pose significant challenges for underwater imaging and its object recognition performance. These aspects may degrade the underwater image quality and lower the performance of object recognition systems. These issues are especially critical for applications involving autonomous underwater vehicles (AUVs) or robotic systems involved in object recognition tasks during search-and-retrieval operations.

Underwater images can be divided into three categories: lakes, seas, and ponds. The main distinction between ponds and lakes is their size, which is described by surface area and depth. Ponds are usually defined by their smaller surface areas and shallower depths compared to lakes, which usually have greater width and deeper basins [1]. Due to limited water input and reduced depth, ponds often have minimal water circulation. It causes stagnant conditions for the water body. These characteristics may decrease light penetration. Consequently, it causes low-visibility environments that produce poorly contrasted and blurred underwater views. The presence of suspended particles, which are commonly brownish or greenish in appearance, is typically associated with nutrient loading. These particles promote algal blooms [2]. The presence of algae not only discolor the water (green, brown, or red) but also degrades water quality through oxygen reduction during the decomposition phase [3]. These environmental conditions pose significant challenges for underwater imaging and object recognition. The visibility limitations may affect the accurate acquisition of object features such as shape and color. To increase the performance of object recognition techniques, the development of a sufficiently large and representative dataset is crucial. However, high-turbidity underwater image datasets, especially for pond environments, remain scarce [4]. The development of such datasets may limit the development and evaluation of robust object recognition models in pond ecosystems. This gap is particularly important for applications involving autonomous underwater vehicles (AUVs) or robotic systems involved in object recognition tasks during search-and-retrieval tasks, where accurate object detection is essential. This study addresses the challenge by developing and assessing a multiple object recognition system tailored for underwater pond images by utilizing the You Only Look Once Version 5 (YOLOv5) deep learning architecture. The specific objectives of this study are as follows:

1) To develop a dedicated dataset for multiple object recognition under pond-specific underwater conditions.

2) To evaluate the existing approach for multiple object detection in underwater pond images.

Given the limitations of low visibility and the existence of turbidity in underwater pond environments. The research question is: Can a deep learning-based approach recognizes multiple objects using a self-collected and annotated dataset?

This study focuses on establishing a pond underwater image dataset and evaluating the deep learning technique, YOLOv5, in object recognition accuracy for multiple objects in respective underwater pond images. The successfully created dataset for this study can be accessed via the following link [a].

This study is organized as follows: Section II discusses existing datasets and related works; Section III explains the methodology, including dataset development and object recognition model; Section IV provides the results and analysis; Section V discusses the performance of YOLOv5, challenges, limitation and future direction of this work; and Section VI concludes key findings and future directions of the research.

[a] Created dataset: http://tinyurl.com/32z25b

II. Related Works

In recent years, advancements in underwater exploration and imaging technologies have led to an increasing number of underwater visual data. Despite these developments, visual images acquired in underwater environments often exhibit significant visual degradation. Common problems include limited visibility, low contrast, non-uniform lighting, blurred textures, diminished color fidelity, and image noise [5]. These issues are primarily caused by the interaction of light with existing particles in water. They absorb and scatter incoming rays, thereby reducing image clarity. The phenomenon of turbidity related to a high concentration of underwater particles contributes to the haziness frequently observed in underwater images [4]. The Marine Underwater Environment Database (MUED) addresses this challenge by providing 8,600 underwater images across 430 distinct object groups. It consists of complex backgrounds, multiple salient entities, and variations in position, illumination, spatial orientation, and turbidity levels [6]. While raw underwater images often display high-level color distortion and contrast loss, the accompanying reference images are corrected to reflect more accurate color representation, enhanced brightness, and visibility. These comprehensive datasets are critical in overcoming the lack of annotated benchmark images. It offers valuable resources for training, validation, and performance evaluation in underwater image enhancement and object recognition research. Fig. 1 shows examples of the MUED underwater images.

Furthermore, the Enhancing Underwater Visual Perception (EUVP) dataset was proposed as a representative database that includes a paired and an unpaired collection of 20,000 underwater images (of poor and good quality) that can be used for training [7]. This dataset was specifically created to accurately reflect the current characteristics and challenges encountered in underwater settings. The dataset is a large assortment of video clips taken in different underwater settings. The videos have varying degrees of water turbidity, lighting, and image degradation. This collection is a wide set to allow researchers to develop and evaluate algorithms aimed at applications like image enhancement, restoration, and object recognition with a particular focus on underwater settings. Using the EUVP dataset, researchers can tackle the special visual problems presented by underwater imaging and try to enhance the quality and analysis of underwater videos and images. Fig. 2 presents example images derived from the EUVP dataset.

Recognition of objects in underwater surveillance, exploration, and assessment is essential in analyzing zones of interest. Not only is object identification important but also information must be extracted from the objects, thereby necessitating object recognition to be an imminent research domain under low-light and underwater environments [8]. In recent years, object recognition has been utilizing deep learning techniques as well. Object recognition methods based on deep learning have been extensively used in various fields [9][10][11], with specific applications in underwater image scenes [12][13][14]. One of the techniques is based on You Only Look Once architecture. Building upon YOLOv2, YOLOv3 introduced several key advancements. It utilized a feature extraction backbone called Darknet-53, consisting of 53 convolutional layers, which improved the model's ability to extract meaningful features [15]. The YOLOv5 also offers support for the latest computer vision algorithms such as instance segmentation, enabling multiple object recognition in an image or videos [16].

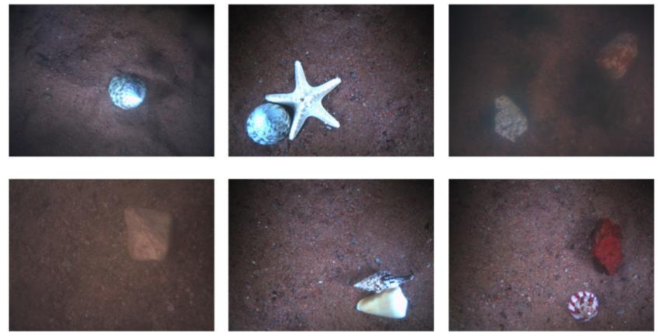


Fig. 1. MUED underwater images [6].

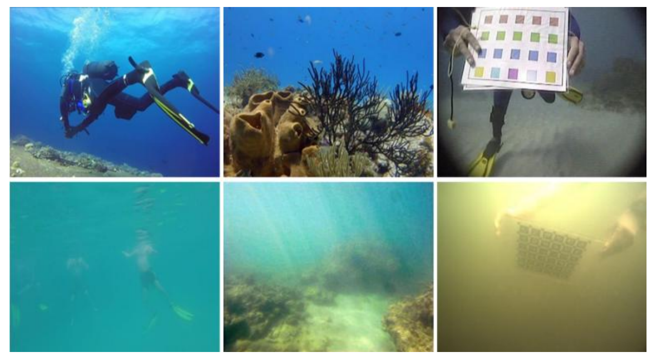


Fig. 2. Example of EUVP dataset images [7].

III. Methodology

A. Data Preparation

The dataset consists of five different classes representing real-world objects: Male, Female, Airplanes, Car, and Helicopter. The images are categorized into five classes based on the number of objects in the image: 1 object, 2 objects, 3 objects, 4 objects, and 5 objects. The images were captured under different conditions and positions to ensure a total of 1116 images. The images were captured at a pond in Taman Muhibbah, Saleng, 81400, Senai, Johor. The images were captured based on three different object heights from the surface (20 cm, 40 cm, and 60 cm), three different object-to-camera distances (10 cm, 20 cm, and 30 cm), and four different object surface directions (0°/360°, 90°, 180°, 270°). These factors ensure a diverse and comprehensive dataset of 1116 images. Fig. 3 shows the toys representing real-world objects.

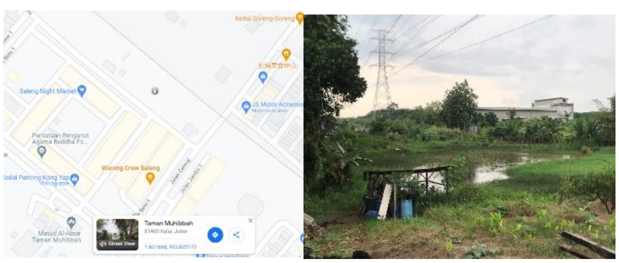
The images are captured with a waterproof action camera GoPro Hero 5. The dataset is processed using LabelImg software to label the images. The object recognition and image enhancement techniques are developed using Google Colaboratory, executed on a personal desktop with the following specifications:

- Windows 10 Home
- Intel(R) Core (TM) i7-6700
- 12GB RAM, 240GB SSD, and 1TB HDD

The GoPro Hero 5, known for its versatility and durability, is well-suited to meet the demands of project requirements. The selected resolution of 1080p ensures a balance between video clarity and file size, optimizing storage efficiency while preserving visual details. The camera's compact design and robust build make it suitable for various environments, aligning seamlessly with the dynamic nature of the project. Fig. 4 shows the parameter setting for the GoPro Hero 5. Fig. 5 shows the location of the experiment in the pond at Taman Muhibbah.

Fig. 3. Toys used in the project.: (a) Airplane, (b) Helicopter, (c) Car, (d) Male, (e) Female.

Fig. 4. Parameter setting for GoPro Hero5 camera.



(a) (b)

Fig. 5. Location 2 for experiment; (a) Taman Muhibbah pond from google map, (b) Taman Muhibbah pond.

B. Data Distribution

In this study, two software programs, namely Google Colaboratory and LabelImg, are utilized. The LabelImg software is employed for labeling objects in the gathered images to create a dataset. Another software tool employed in this research is Google Colaboratory. It was utilized to perform object recognition tasks via the YOLOv5 deep learning framework. During the dataset preparation process, thorough consideration was given to ensure a balanced and representative distribution of underwater images across all data subsets. The dataset was divided into three main folders: 80% of the images were allocated for training, 10% for testing, and the remaining other 10% for validation purposes [17]. This division was implemented to support effective model training and facilitate rigorous performance evaluation. It also enables fine-tuning of hyperparameters in the validation phase. The specific dataset distribution across these subsets is presented in Table I.

TABLE I DISTRIBUTION OF UNDERWATER IMAGES DATASET

Object quantity	Train folder		Test folder		Valid folder		Total Image No.
	%	*Image No.*	%	*Image No.*	%	*Image No.*	
1	80	144	10	18	10	18	180
2	80	288	10	36	10	36	360
3	80	288	10	36	10	36	360
4	80	144	10	18	10	18	180
5	80	29	10	4	10	3	36
Combine images (all)	80	893	10	112	10	111	1116

C. You Only Look Once Version 5 (YOLOv5)

Object recognition in underwater images encountered a number of issues such as poor visibility, dynamic illumination, and particle pollution. The YOLOv5 is a complex object recognition technique that can be adapted to address such issues. This section discusses the key mechanisms and steps involved in utilizing YOLOv5 for object recognition for underwater images. Fig. 6 demonstrates the flowchart of YOLOv5. It utilizes a single neural network that can directly predict bounding boxes and class probabilities from entire images in a single evaluation step. There are three elements that build up the architecture, which are the Backbone, Neck, and Head.

The process of training YOLOv5 for underwater object recognition includes some important steps as follows:

1) Data preparation: The underwater images are annotated using bounding boxes and object-class-specific labels.

2) Data augmentation: To improve the robustness and generalization of the model, a variety of data augmentation techniques can be employed. The examples are random cropping, scaling, flipping, color correction, and underwater-specific augmentations like simulating turbidity and altering light conditions.

3) Loss function: The YOLOv5 employs a combination of several loss functions to train the model, which are the Localization Loss, Confidence Loss and Classification Loss.

4) Optimization: Model training is carried out using stochastic gradient descent (SGD) or the Adam optimizer. Hyperparameters such as learning rate, batch size, and momentum are carefully tuned to achieve the best performance results. In many cases, transfer learning is used, starting with a pre-trained model trained on a large dataset and then fine-tuning it on the target underwater dataset.

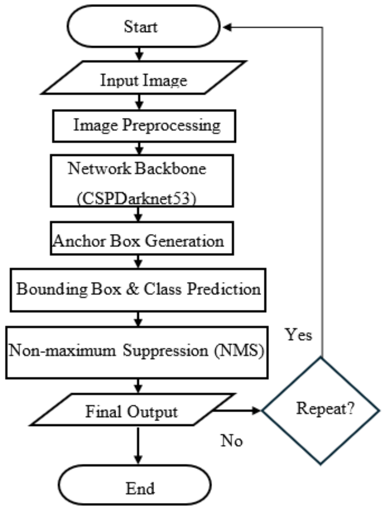


Fig. 6. YOLOv5 approach for object recognition.

At inference time, YOLOv5 feeds the input image to the trained network to get the bounding boxes and class probabilities. The subsequent steps are:

1) Image preprocessing: The original image is resized to a specific size and normalized. Additional preprocessing steps can include color correction and noise reduction to deal with the specific nature of underwater images.

2) Forward pass: The preprocessed image is passed into the network, resulting in the generation of feature maps via the backbone, neck, and head modules.

3) Post-processing: Non-Maximum Suppression (NMS) is applied to eliminate duplicate bounding boxes and retain the most confident predictions. Additional post-processing may include object size and confidence thresholds specific to underwater object filtering.

YOLOv5 is a significant improvement in real-time object recognition with improved accuracy and efficiency for object recognition underwater. Its architecture, training process, and inference steps are meticulously designed to handle the unique challenges of underwater environments.

IV. Result and Analysis

This section examines the results and analysis of the underwater image dataset and the object recognition system using Google Colaboratory. The results and analysis are discussed in the following order: i) Dataset development for pond underwater images and ii) YOLOv5 performance analysis for all combined pond underwater images.

A. Dataset Development for Pond Underwater Images

The created dataset can be accessed via the following link [a]. Fig. 7 shows the underwater image acquisition arranged based on the Car, Male, Female, Helicopter, and Airplane categories. It displays the images of each object according to different underwater conditions. The settings show examples of different quantities of objects with depths from the surface varying from 20 cm, 40 cm, and 60 cm inside the pond. The distance of the object to the camera and the object's angle are fixed, with a distance of 30 cm from the object to the camera and the angle of objects all set to 0°. Despite the brownish tone of the pond, which contrasts with the greenish tint of the lake, the objects remain apparent in the photographs, with constant settings for object depths and camera distances. However, with these settings for object distances from the surface and camera, the objects can still be seen in the images.

Fig. 8 shows the example of underwater image object-to-camera distance setting for image acquisition for the pond, set to 10 cm, 20 cm, and 30 cm. The distance from the surface and angle of the object are fixed, with the distance from the surface set to 20 cm and the angle of the object remaining at 0°. When the distance of objects is nearer to the camera, the objects can still be seen, although the distance from the surface differs. It can be observed that for nearer distances, there are possibilities of objects at the edge being partially displayed in the image due to the camera's capturing width limitations. In some cases, the objects may overlap with neighboring objects for higher quantity objects in an image. It is critical to recognize that these visual impacts may differ in the pond due to its brownish color compared to the lake's greenish tint.

Fig. 9 shows the examples of underwater image angle settings for the images in the pond captured at 0°, 90°, 180°, and 270°. The depth from the surface and object-to-camera distance are fixed, with the depth set to 20 cm from the surface and the distance of the object from the camera set to 30 cm. When the object distances from the surface and camera are longer, the image turbidity tends to be higher, and object visibility becomes less clear. The visual effects witnessed in the pond, marked by its brownish hue, may diverge from those in the lake, distinguished by a greenish tint. These environmental distinctions have the potential to influence the overall clarity and appearance of objects captured within the images.

B. YOLOv5 Performance Analysis for All Combined Underwater Images

In this analysis, the focus is on a meticulous evaluation of YOLOv5's performance, leveraging a dataset explicitly tailored for underwater images in a pond environment. The assessment encompasses critical metrics such as confusion matrices, result graphs, result values, and illustrative images, providing a detailed perspective on the model's prowess in object recognition. The aim is to deliver a comprehensive understanding of YOLOv5's efficacy in detecting and identifying objects, particularly amid the distinctive challenges posed by underwater conditions in pond environments.

C. YOLOv5 Training Performance: Confusion Matrix for Pond Images

Fig. 10 shows the confusion matrix for pond images for all combined images. The confusion matrix shows four crucial values: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The diagonal elements represent correctly predicted samples. In this specific case, out of a total of 892 samples, 892 samples were accurately predicted, leading to an overall accuracy of 100%. From these 892 samples, 67 instances of airplane classes were detected, 56 instances of car classes were detected, 50 instances of female classes were detected, 45 instances of male classes were detected, and 61 instances of helicopter classes were detected.

Fig. 7. Different surface distance settings for pond underwater images (distance of object to camera is 30 cm, and angle of object is 0 °).

Fig. 8. Different camera distances from the object setting for pond underwater images (depth of water surface is 20cm and angle of object is 0 °).

Fig. 9. Different angle of object setting for dataset for pond underwater images (depth from water surface is 20 cm, distance of object to camera is 30 cm).

D. YOLOv5 Training Performance: Result Graph for Pond Images

Fig. 11 shows the training result graph for pond images for all combined images. The YOLOv5 model, trained over 150 epochs, exhibited notable improvements in alignment and performance between epochs 50 and 100, as indicated by the decreasing training losses. The evaluation metrics, including precision, recall, mAP 50, and mAP 50-95, provided a comprehensive overview of the model's ability to accurately detect and classify objects. During the analyzed epochs, precision consistently increased, indicating better object classification, while recall remained high, demonstrating effective object recognition. The mAP 50 and mAP 50-95 metrics, which assess the mean average precision at different IoU thresholds, reflected the model's proficiency in localization and classification tasks.

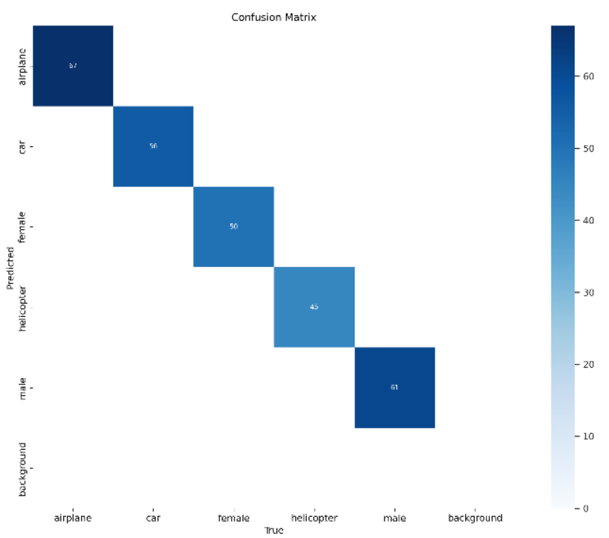


Fig. 10. Training confusion matrix for pond images (combine all images) YOLOv5.

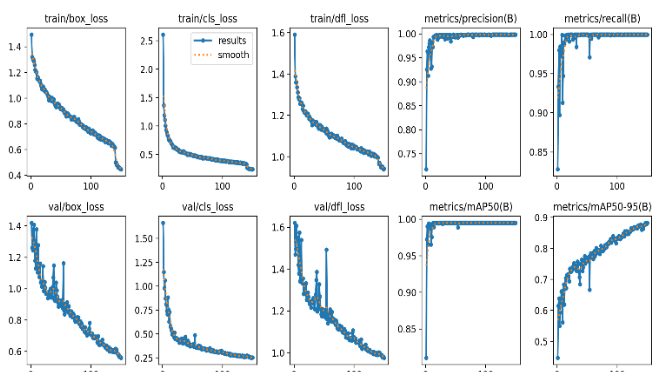


Fig. 11. Training result graph for pond images (combine all images) YOLOv5.

E. YOLOv5 Validation Performance: Confusion Matrix for Pond Images

Fig. 12 shows the confusion matrix for pond images for combined all images. The confusion matrix shows four crucial values: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The diagonal elements represent correctly predicted samples. In this specific case, out of a total of 112 samples, 112 samples were accurately predicted, leading to an overall accuracy of 100%. From these 112 samples, 67 instances of airplane classes were detected, 56 instances of car classes were detected, 50 instances of female classes were detected, 45 instances of male classes were detected, and 61 instances of helicopter classes were detected.

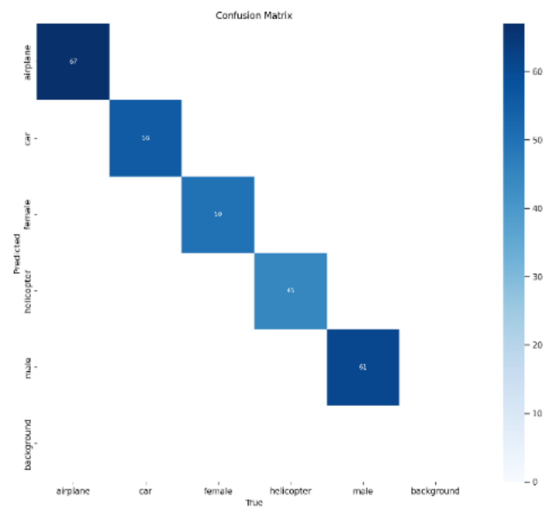


Fig. 12. Validation confusion matrix for pond images (combine all images) YOLOv5.

F. YOLOv5 Validation Performance: Visual Result for Pond Images

Fig. 13 shows examples of images randomly chosen from the validation set of pond images for all combined images. The images chosen include 1 object, 2 objects, 3 objects, 4 objects, and 5 objects. The system detected the objects with correct classes, which are airplane, car, helicopter, male, and female, with the average accuracy detected value above 90%.

G. YOLOv5 Validation Performance: Overall Accuracy Mean Average Precision 50-95 (mAP 50-95) for Pond Images

Table II shows the overall accuracy mean average precision 50-95 (mAP 50-95) while executing the validation. The mAP 50-95 for all classes is 88.3%, with airplane at 90.1%, car at 89.1%, female at 86.4%, helicopter at 87.4%, and male at 88.7%. Increasing the number of validation images may contribute to a more comprehensive evaluation, enabling the model to generalize better diverse scenarios and, consequently, enhance overall accuracy.

Fig. 13. Validation visual result for pond images (combine all images) YOLOv5.

H. YOLOv5 Test Performance: Confusion Matrix for Pond Images

Fig. 14 shows the testing confusion matrix for pond images for all combined images. The confusion matrix shows four crucial values: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The diagonal elements represent correctly predicted samples. In this specific case, out of a total of 112 samples, were accurately predicted, leading to an overall accuracy of 100%. From these 112 samples, 55 instances of airplane classes were detected, 53 instances of car classes were detected, 61 instances of female classes were detected, 69 instances of male classes were detected, and 56 instances of helicopter classes were detected.

TABLE II VALIDATION OVERALL ACCURACY MEAN AVERAGE PRECISION 50-95 (MAP 50-95) FOR POND IMAGES (COMBINE ALL IMAGES) YOLOV5

Class	Images	mAP50	mAP50-95
All	110	0.995	0.883
Airplane	110	0.995	0.901
Car	110	0.995	0.891
Female	110	0.995	0.864
Helicopter	110	0.995	0.874
Male	110	0.995	0.887

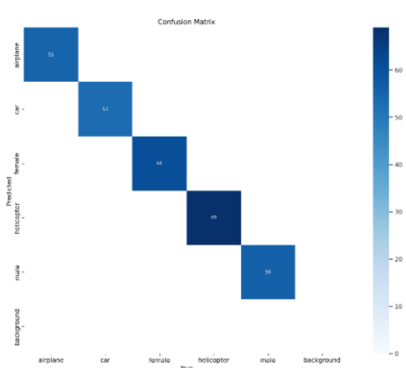


Fig. 14. Test confusion matrix for pond images (combine all images) YOLOv5.

I. YOLOv5 Test Performance: Overall Accuracy Mean Average Precision 50-95 for Pond Images

Table III shows the overall accuracy mean average precision 50-95 (mAP 50-95) while executing the testing. The mAP 50-95 for all classes is 87.4%, with airplane at 89.9%, car at 86.6%, female at 85.6%, helicopter at 87.6%, and male at 87.2%. Increasing the number of testing images may contribute to a more comprehensive evaluation, enabling the model to generalize better to diverse scenarios and, consequently, enhance overall accuracy.

J. YOLOv5 Test Performance: Visual Result for Pond Images

Fig. 15 shows examples of images randomly chosen from the test set of pond images for all combined images. The images chosen include 1 object, 2 objects, 3 objects, 4 objects, and 5 objects. The system detected the objects with correct classes, which are airplane, car, helicopter, male, and female, with the average accuracy detected value above 90%.

K. Overall Evaluation of YOLOv5 Performance

This section focuses on comparing the mAP 50-95 (Mean Average Precision where the IoU threshold is 0.5-0.95) results obtained by YOLOv5 for underwater images in the pond environment. A detailed examination of the mAP scores in each location aims to reveal potential variations in the model's performance under different environmental conditions. Table IV shows the evaluation of YOLOv5 performance.

TABLE III TEST OVERALL ACCURACY MEAN AVERAGE PRECISION 50-95 (MAP 50-95) FOR POND IMAGES (COMBINED ALL IMAGES) YOLOV5

Class	Images	mAP50	mAP50-95
All	112	0.995	0.874
Airplane	112	0.995	0.899
Car	112	0.995	0.866
Female	112	0.995	0.856
Helicopter	112	0.995	0.876
Male	110	0.995	0.872

TABLE IV EVALUATION OF THE YOLOV5 PERFORMANCE

Overall Images of object category and mode	mAP 50-95% for YOLOv5
Train	88.20%
Validation	88.30%
Test	87.40%
Average accuracy	**87.97%**

Under the pond classification category, YOLOv5 records an average accuracy of 87.97%. Despite excellent performance being recorded by YOLOv5, there are slight accuracy variations across various stages of the dataset, i.e., training, validation, and testing.

The mAP 50-95 metric is an overall assessment of the performance of the model, considering precision and recall at different Intersection over Union (IoU) thresholds. High mAP values observed on training, validation, and test datasets. It demonstrates that YOLOv5 has high proficiency in object recognition and classification in underwater images.

Training Performance: The mAP 50-95 score of 88.20% in training exhibits that the model has learned to recognize and classify objects present in the training dataset successfully. The high score indicates the model's ability to generalize well to the training data.

Validation Performance: The validation mAP 50-95 score of 88.30% shows that the model maintains its performance when applied to unseen data during training. This stability between train and validation scores proves that the model is not overfitting and can generalize well to new data.

Testing Performance: The testing mAP 50-95 of 87.40% is less than the train and validation. However, it still represents good performance. This small reduction in accuracy is to be expected because the testing dataset is made of completely different underwater image dataset. The high-test score proves the strength and reliability of the model for real-world scenarios.

The consistency of high mAP scores throughout all datasets shows the model's effectiveness and reliability.

V. DISCUSSION

The results obtained from this study demonstrate the strong performance of YOLOv5 in multiple object recognition within pond environments. The average mAP 50–95 of 87.97% shows the robustness of YOLOv5 in object recognition under presence of turbidity and low visibility in pond environments.

One of the challenges is the lack of pond underwater datasets. Most datasets focus on marine and lake setting, with many with lower level of turbidity. These differences affect the results of observations across different water body environments.

The YOLOv5, while effective, still have room for improvement in object recognition, especially in the case of high-turbidity conditions. For future work, the use of image enhancement techniques as image processing, may further improve model accuracy in object recognition.

Despite these challenges, this study fills a critical gap and offers a valuable contribution for future research in underwater object recognition, especially for pond environments.

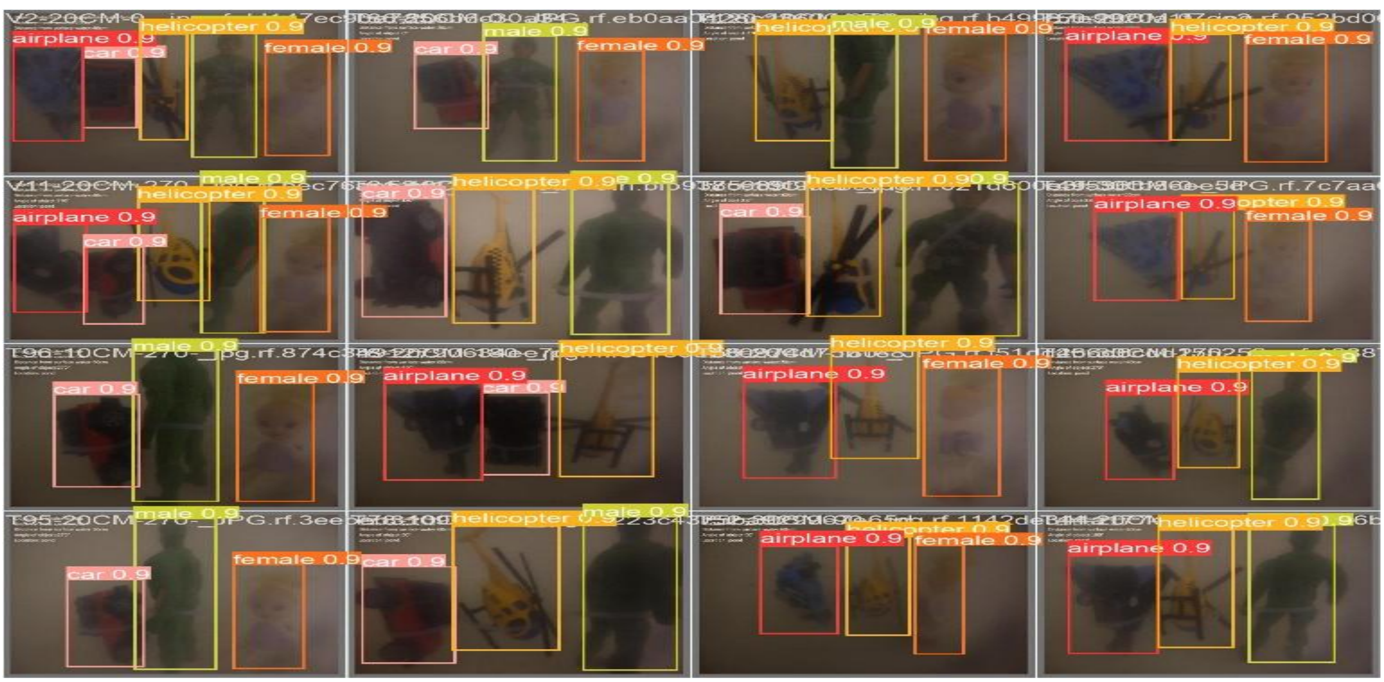


Fig. 15. Test: visual result for pond images (combine all images) YOLOv5.

VI. Conclusion

This study presents a contribution to underwater object recognition by developing a self-captured and annotated dataset for pond environments. The created dataset can be accessed via the following link [a]. Through investigation utilizing the datasets, the performance evaluation of YOLOv5 with the average mAP 50-95 of 87.97% metric exhibits that the model has a high capability in recognizing and classifying multiple objects in underwater images. The high accuracy shown on training, validation, and test datasets demonstrates that YOLOv5 has good generalization capability to unseen data and maintains stable performance in real-world underwater environments. However, this study has limitations, including the need for image pre-processing and more image datasets. Future studies can consider further optimization of the YOLOv5 model and its use in numerous underwater environments to improve its robustness and accuracy. Overall, this study is hoped to offer new insights and resources for object recognition not only in pond environments, but also in the real-world underwater environments.

Acknowledgment

This research is supported and funded by Universiti Tun Hussein Onn Malaysia (UTHM) through Tier 1(Q377) Grant.

References

[1] J. Britton, "Formation and features of ponds and lakes," Study.com, 2023 [Online]. Available: https://study.com/learn/lesson/pond-vs-lake-formation-features.html.

[2] MSU Extension Natural Resources, "What is that color in my pond?" 2022. [Online]. Available: https://www.canr.msu.edu/news/what-is-that-color-in-my-pond.

[3] K. F. Smith, J. Stuart, and L. L. Rhodes, “Molecular approaches and challenges for monitoring marine harmful algal blooms in a changing world,” Frontiers in Protistology, vol. 1:1305634, 2024.

[4] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," IEEE Transactions on Image Processing, vol. 29, pp. 4376-4389, Nov. 2019.

[5] E. Lucas et al., " Underwater image enhancement and object detection: Are poor object detection results on enhanced images due to missing human labels?" in 2025 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW) pp. 1435-1440, Feb. 2025.

[6] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using Generative Adversarial Networks," in IEEE International Conference on Robotics and Automation (ICRA), pp. 7159-7165, Sept. 2018.

[7] Y. Ma et al., "A new database for evaluating underwater image processing methods," in IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1-6, Nov. 2018.

[8] N. A. I. N. Azlin et al., "Object recognition system with image enhancement for lake underwater images," Evolution in Electrical and Electronic Engineering, vol. 4(1), pp. 559-565, May 2023.

[9] G. Rani, S. Khandelwal, N. Kundu and V. S. Dhaka, "Automatic segmentation and classification of skin cancer cells using thresholding and Deep Learning based techniques," in IEEE International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), pp. 1-6, Aug. 2023.

[10] H. H. Kang, A. Ab Malik, M. A. Osman, Y. Yusop and I. H. Hamzah, " Healthy and unhealthy oil palm tree detection using Deep Learning method," International Journal of Advanced Computer Science and Applications(IJACSA), vol. 16(4), pp. 748-757, 2025.

[11] N. Das and S. Das, "Epoch and accuracy based empirical study for cardiac MRI segmentation using Deep Learning technique," PeerJ, Mar 2022.

[12] L. Chen et al., "Underwater object detection in the era of Artificial Intelligence: Current, challenge, and future," arXiv preprint arXiv: 2410.05577, Oct. 2024.

[13] L. Chen et al., "Underwater object detection in the era of Artificial Intelligence: Current, challenge, and future," arXiv preprint arXiv:2410.05577, Oct. 2024.

[14] Y. S. Tsai, C. T. Tsai, and J. H. Huang, "Multi-scale detection of underwater objects using attention mechanisms and normalized Wasserstein distance loss," The Journal of Supercomputing, vol. 81, pp. 1-33, Mar. 2025.

[15] Y. Yang et al., "Image defogging based on amended Dark Channel Prior and 4-Directional L1 Regularisation," IET Image Processing, vol. 15, pp. 2454-2477, Apr. 2021.

[16] J. H. Kim, N. Kim, Y. W. Park, and C. S. Won, "Object detection and classification based on YOLO-V5 with improved maritime dataset," Journal of Marine Science and Engineering, vol. 10, no. 3, pp. 377, Mar. 2022.

[17] A. Katuria and S. Mukherjee, "How to Train YOLO v5 on a custom dataset," DigitalOcean Community Tutorials, Mar. 4, 2025. [Online]. Available: https://www.digitalocean.com/community/tutorials/train-yolov5-custom-data.