An AI AUV Enabling Vision-based Diver-following and Obstacle Avoidance with 3D-modeling Dataset

Yu-Cheng Chou
Institute of Undersea Technology
National Sun Yat-sen University
Kaohsiung, Taiwan
ycchou@mail.nsysu.edu.tw

Hsin-Hung Chen
Institute of Undersea Technology
National Sun Yat-sen University
Kaohsiung, Taiwan
hhchen@faculty.nsysu.edu.tw

Chau-Chang Wang
Institute of Undersea Technology
National Sun Yat-sen University
Kaohsiung, Taiwan
chauwang@mail.nsysu.edu.tw

Hui-Min Chou
Institute of Undersea Technology
National Sun Yat-sen University
Kaohsiung, Taiwan
az70021@gmail.com

Chua-Chin Wang
Department of Electrical Engineering
National Sun Yat-sen University
Kaohsiung, Taiwan
ccwang@ee.nsysu.edu.tw

2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS) | 978-1-6654-1913-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/AICAS51828.2021.9458431

***Abstract*—This paper presents an AUV with AI (artificial intelligence), which is able to perform real-time optical vision-based diver-following and forward looking altimeter-based obstacle avoidance. The AI AUV is equipped with thrusters and a standard navigation-related sensor suite. A diver detection convolutional neural network, a suite of motion controllers, and a diver detection payload device are developed to enable diver-following functionality of the AUV. An obstacle avoidance algorithm based on forward looking altimeters is developed to enhance the waypoint navigation security of the AUV with obstacle avoidance functionality. The diver-following and the obstacle avoidance capabilities of the Taiwan Moonshot AUV under different scenarios are evaluated through hardware-in-the-loop simulations. In addition, the designated single diver following capability of the Taiwan Moonshot AUV is also verified through closed water experiments conducted in a towing tank.**

Keywords—Artificial Intelligence; Autonomous Underwater Vehicle; Diver-following; Obstacle Avoidance.

I. INTRODUCTION

A team of human divers and AUVs usually cooperate to carry out certain underwater missions, such as the inspection of ship hulls and submarine pipelines, the study of marine species migration, search and rescue, and surveillance. The divers typically lead the tasks and interact with the AUVs that follow the divers at certain stages of a mission [1]. In these applications, it is important that an AUV is able to follow and interact with a human diver. The diver-in-the-loop guidance reduces operational overhead by eliminating the need of teleoperation or complex mission planning of an AUV [2].

Based on the qualitative comparisons among some representative diver-following AUVs over the past decade [2], AUVs that have the best diver-following performance are classified into two types, including the one with an optical camera [1] and the other with an active sonar [3]. In order to have an eco-friendly underwater operation that is not intrusive and disruptive to the ecosystem, passive sensors without emitting energy, such as optical cameras, are preferred over active sensors. On the other hand, active sonars are useful for diver-following applications in unfavorable visual conditions [3]. Compared with sonar devices, the main limitation of optical cameras is their short underwater visibility range. However, optical cameras have the advantages of high resolution, high frame rate, low cost, and high application popularity.

This research is focused on the development of an AI (artificial intelligence) AUV, the Taiwan Moonshot AUV, which enables optical vision-based diver-following and forward looking altimeter-based obstacle avoidance capabilities. The Taiwan Moonshot AUV, developed by the Institute of Undersea Technology (IUT) at the National Sun Yat-sen University (NSYSU), is equipped with horizontal and vertical thrusters and a standard sensor suite including an altimeter with a depth sensor, a Doppler velocity logger, a gyrocompass, and an underwater acoustic modem.

A diver-following control system consisting of three main components, including a diver detection convolutional neural network (CNN), a detection strategy planner, and a suite of motion controllers, is developed. A diver detection unit consisting of a single board computer and a webcam module is developed and installed on the Taiwan Moonshot AUV as a payload device. The diver detection CNN and the detection strategy planner are executed on the single board computer of the diver detection unit. In addition, an obstacle avoidance algorithm based on two forward looking altimeters is developed and integrated with the waypoint navigation algorithm previously developed for the Taiwan Moonshot AUV. Two altimeters will be installed on the front side of the AUV, facing forward, as payload devices to obtain the distance readings between the AUV and obstacles.

The diver-following capability as well as the waypoint navigation with obstacle avoidance capability of the Taiwan Moonshot AUV under different scenarios are evaluated through hardware-in-the-loop (HIL) simulations. In addition, the designated single diver following capability of the Taiwan Moonshot AUV is also verified through closed water experiments conducted in a towing tank.

II. OPTICAL VISION-BASED DIVER-FOLLOWING

This section presents the diver detection CNN, the motion controllers, and the diver detection unit, which are the main components enabling the optical vision-based diver-following capability of the Taiwan Moonshot AUV.

A. Diver Detection CNN

This research adopts a light weight CNN called Tiny-YOLOv3 [4] as the diver detection algorithm. Tiny-YOLOv3 has a total of 22 layers, including convolutional layers, max pooling layers, up-sampling layers, and concatenation layers, as shown in Fig. 1. Tiny-YOLOv3 can produce two feature maps with two different scales to facilitate the detection of large and small objects. Tiny-YOLOv3 was trained on a

This research is supported by the Ministry of Science and Technology, Taiwan.

workstation with Intel CPU and Nvidia GPU processors through Darknet deep learning framework under Ubuntu operating system, as shown in TABLE I. The training dataset contains a total of 5386 diver images.

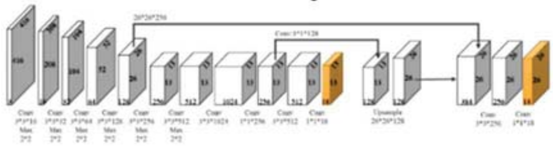

Fig. 1. Network architecture of Tiny-YOLOv3.

TABLE I. TRAINING ENVIORNMENT OF TINY-YOLOV3

Processor	CPU: Intel Xeon Gold 5122 GPU: GeForce RTX 2080 Ti
Operating system / Deep learning framework	Ubuntu / Darknet
Programming language	C language

B. Montion Controllers

A set of three PID motion controllers are developed to enable the diver-following capability of the Taiwan Moonshot AUV. The AUV main computer, which is a Raspberry Pi single board computer, executes the three motion controllers. The data needed by the motion controllers are sent from the diver detection unit, including the center coordinates (x, y) and the dimensions (w, h) of detected bounding boxes, as shown in Fig. 2. The desired and feedback values for the three controllers are shown in Fig. 3. For the yaw controller, the target value is x_0, the x-coordinate of an image center, and the feedback value is x, the x-coordinate of a bounding box center. For the heave controller, the target value is y_0, the y-coordinate of an image center, and the feedback value is y, the y-coordinate of a bounding box center. For the surge controller, the target value is 0.2 and the feedback value is $\left(\frac{h_1 \times w_1}{640 \times 480}\right)$, the area ratio of a bounding box to the entire image.

The three PID controllers generate motion commands to bring the bounding box of the diver to the center of the camera view and to keep the size of the bounding box constant. The Ziegler-Nichols method is used to obtain the PID gains of the heave, yaw, and surge controllers, as shown in TABLE II.

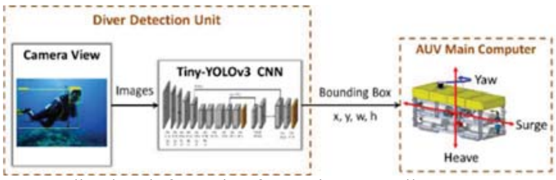

Fig. 2. Bounding box information for motion controllers.

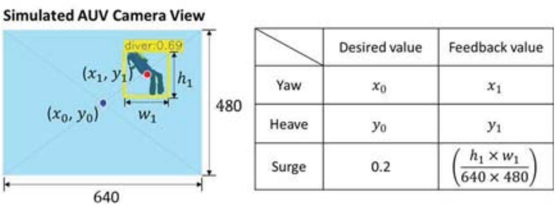

Fig. 3. Desired and feedback values of three motion controllers.

TABLE II. PID GAINS OF MOTION CONTROLLERS

	K_P	K_P	K_D
Yaw	0.03	0.0001	0.07
Heave	0.12	0.01077	0.3433
Surge	0.54	0.082	0.891

C. Diver Detection Unit

The main components of the diver detection unit include a LattePanda single-board computer, a webcam module, and two 24V-to-12V DC-DC converters, as shown in Fig. 4. The diver detection unit has an input voltage of 24 VDC, two communication interfaces, a data rate of 4 Hz, and a maximum power consumption of 36 W. The resolution of the webcam image is 640 × 480 pixels. The underwater field of views are 63° horizontally and 48° vertically. The RS-232 interface is used for underwater data transfer between the diver detection unit and the AUV main computer. The Ethernet interface is used for data transfer after the AUV is retrieved back to the surface vessel, surface platform, or the ground.

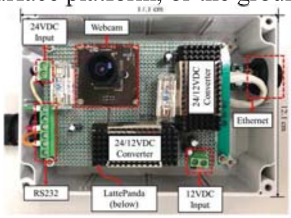

Fig. 4. Diver detection module and its specifications..

III. FORWARD ALTIMETER-BASED OSTACLE AVOIDANCE

Two altimeters will be installed on the front side of the Taiwan Moonshot AUV, facing forward, as payload devices to obtain the distance readings between the AUV and obstacles, as shown in Fig. 5. The primary obstacle avoidance procedure based on two forward looking altimeters is shown in Fig. 6. When the AUV begins to execute the waypoint navigation missions, the waypoint navigation part of the algorithm will navigate the AUV towards the desired waypoints one after another by changing the AUV's heading and speed. Meanwhile, based on the distance readings from the two forward looking altimeters, the obstacle avoidance part of the algorithm will determine the AUV's speed, turning direction, and moving distance to avoid obstacles. When the distance readings indicate that no obstacles are within a pre-defined distance threshold in front of the AUV, the AUV will switch back to the waypoint navigation mode.

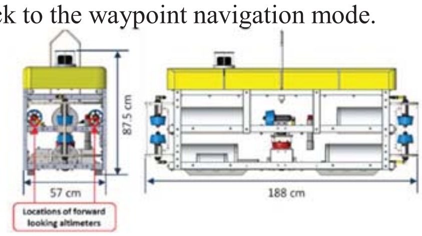

Fig. 5. Locations tentative to install two forward looking altimeters.

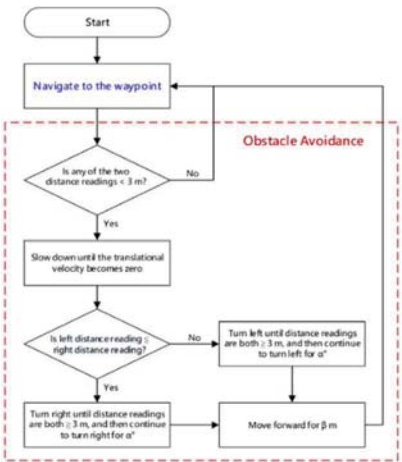

Fig. 6. Primary obstacle avoidance rules based on forward looking altimeters.

IV. 3D-Modeling Underwater Dataset Augmentation

Ever since AI algorithms using DNN or CNN were highly promoted in early 2010's, e.g., [6], YOLO has been known one of the best options for real-time objection recognition tools. Theoretically, at least 10,000 pictures with various viewpoints, sizes, and positions for the same object is needed to achieve 90% accuracy by commercially available tools, e.g., [7]. It is hard to massive amount of samples for rare object underwater, e.g., special plankton, sea spider, etc. To resolve this problem, a dataset augmentation method using 3D modeling is proposed in this investigation.

3D modeling attains two features: multiple viewpoints and back-grounds. Referring to Fig. 7, assume the target object is placed at origin, (0,0,0). At least 26 viewpoints can be generated, which are from (x,y,z), where $x,y,z \in \{-1, 0, +1\}$ and x,y,z =0 at the same time. The amount of information that a viewer can collect from an object is proportional to the projection area from his/her viewpoint. Therefore, those viewpoints with maximum information is (x,y,z), $|x|=|y|=|z|=1$. Namely, the 8 corners of the cube in Fig. 5 are the best selections. As for the added background, basic selections include: clear water, dark water, coral reef, rocks, seagrass. With reference to Fig. 8, the examples of clown fish and the diver images can be easily generated.

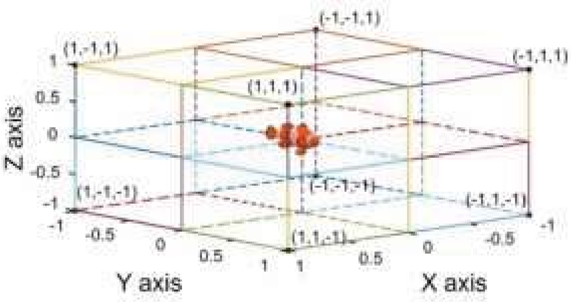

Fig. 7. Multiple 3D modeling viewpoints

Fig. 8. Examples of 3D modeling : (a) clown fish; (b) diver

V. Simulations and Experiments

This section presents the results from hardware-in-the-loop simulations to verify the diver-following and obstacle avoidance capabilities of the Taiwan Moonshot AUV. The HIL simulation platform used in this research was developed in our previous work [5]. This section also presents the results from towing tank experiments to verify the single diver-following capability of the Taiwan Moonshot AUV in a closed water environment.

A. Diver-following HIL Simulations

Two HIL simulations are presented to demonstrate the diver-following performance of the Taiwan Moonshot AUV under two different motions of the diver. In the first HIL simulation, the diver moves along a straight path with a constant speed of 0.2 m/s and maintains at a constant altitude. The simulation results show that the AUV can follow the diver and the deviation of the AUV trajectory from a straight line is roughly within 15 cm. As for the diver detection performance, the precision index is 98.2% and the recall index is 79.2%, as shown in the left of 9.

In the second HIL simulation, the diver moves along a square path with a constant speed of 0.2 m/s and maintains at a constant altitude. The simulation results show that the AUV can follow the diver; bounding boxes of the diver are roughly around the center of the camera view. As for the diver detection performance, the precision index is 99.3% and the recall index is 95.9%, as shown in the right side of Fig. 9.

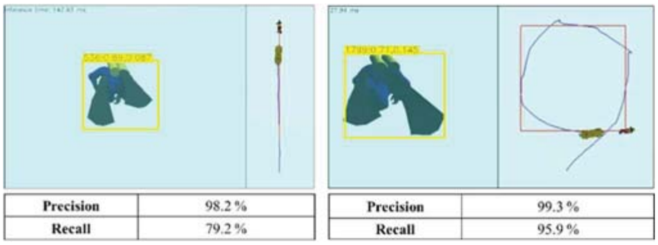

Fig. 9. Diver-following HIL simulation results for which the diver moves along a straight path (left) and a square path (right).

B. Obstacle Avoidance HIL Simulations

Two HIL simulations are presented to demonstrate the obstacle avoidance performance of the Taiwan Moonshot AUV in two different-sized water tanks with two different-shaped obstacles. In the first HIL simulation, the dimension of the water tank is specified as 50 m (length) × 16 m (width) × 3.5 m (water depth). There are two obstacles placed in the water tank. The first obstacle is a cylindrical object whose dimension is specified as 1.0 m (diameter) × 3 m (height). The second obstacle is a rectangular object whose dimension is specified as 0.2 m (length) × 1.0 m (width) × 1.5 m (height). As shown in Fig. 10, the simulation results indicate that the AUV is able to avoid the two obstacles and arrive at the four desired waypoints to complete the waypoint navigation tasks in a larger water tank environment.

In the second HIL simulation, the dimension of the water tank is specified as 50 m (length) × 8 m (width) × 3.5 m (water depth), which is narrower than the one used in the first simulation. Additionally, the two obstacles used in the second simulation are the same as those used in the first simulation. As shown in Fig. 11, the simulation results indicate that the AUV is able to avoid the two obstacles and arrive at the four desired waypoints to complete the waypoint navigation tasks in a smaller water tank environment.

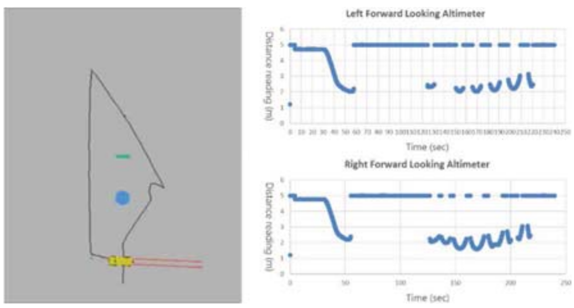

Fig. 10. Obstacle avoidance HIL simulation results for which the AUV needs to avoid two obstacles and arrive at four waypoints in a larger tank.

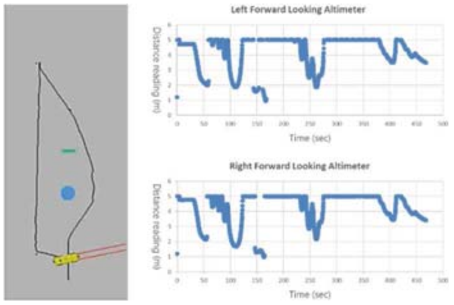

Fig. 11. Obstacle avoidance HIL simulation results for which the AUV needs to avoid two obstacles and arrive at four waypoints in a smaller tank.

C. *Closed Water Experiments*

Diver-following experiments for the surge-yaw and heave-yaw hybrid following modes were conducted with a single diver moving in front of the Taiwan Moonshot AUV, as show in Fig. 12. The diver detection results for the surge-yaw and heave-yaw following experiments indicate that the trained Tiny-YOLOv3 obtains high detection accuracy in these two experiments, as shown in Fig. 13. Overall, the experimental results show that the Taiwan AUV is able to follow a diver who slowly performs a linear or planar motion in a closed water environment.

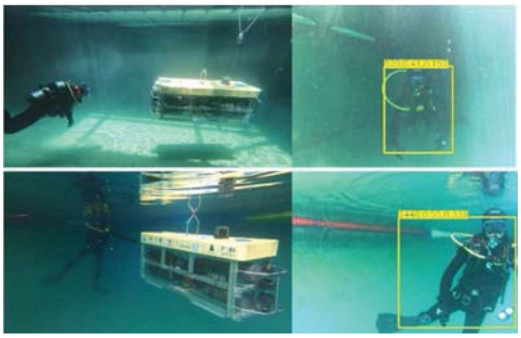

Fig. 12. Diver-following experiments conducted using the Taiwan Moonshot AUV in a towing tank.

Surge-Yaw Following

Tiny-YOLOv3	
Ground truth	2128
Detected	1966
TP	1928
FP	38
FN	200
Precision	98.07 %
Recall	90.60 %
mAP	90.39 %

Heave-Yaw Following

Tiny-YOLOv3	
Ground truth	1198
Detected	1007
TP	961
FP	46
FN	237
Precision	95.43 %
Recall	80.22 %
mAP	78.29 %

Fig. 13. Diver detection results for surge-yaw and heave-yaw following experiments.

D. *3D-modeling Compared with Benchmark Datasets*

Many widely recognized marine or ocean related datasets are spread over the internet, e.g., ImageNet [8], Fish4Knowledge [9]. We have conducted the same learning and testing procedures to the 2 datasets and ours to make a fair comparison using the following key performance indices (KPI). Fig. 14 is graphical demonstration of the these KPIs.

- UIQM (underwater image quality measure)
- UCIQE (underwater color image quality evaluation)
- UICM (underwater image colorfulness measure)
- UISM (underwater image sharpness measure)
- UIconM (underwater image contrast measure)

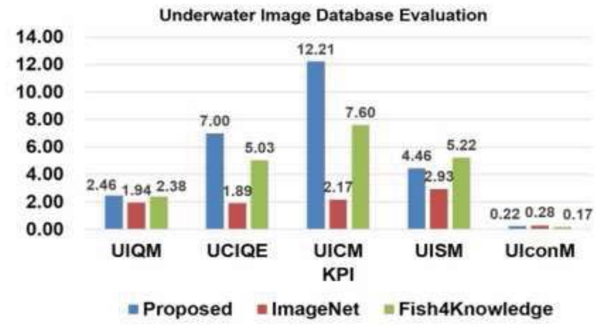

Fig. 14. Performance comparison with benchmark datasets

VI. Conclusions

This paper presents an artificial intelligence AUV, the Taiwan Moonshot AUV, which is able to perform optical vision-based diver-following and forward looking altimeter-based obstacle avoidance. The diver detection CNN, the motion controllers, and the diver detection unit, which enable the optical vision-based diver-following capability of the AUV have been presented. The primary obstacle avoidance procedure based on two forward looking altimeters has been illustrated as well.

HIL simulations have been presented to demonstrate that the Taiwan Moonshot AUV can follow a single diver who moves with different trajectories. In addition, HIL simulations have also been presented to demonstrate that the Taiwan Moonshot AUV can avoid obstacles and arrive at desired waypoints in closed water environments of different sizes. Moreover, experimental results obtained from a towing tank have been presented to show that the Taiwan Moonshot AUV is able to detect a single diver and follow the detected diver, who slowly performs a linear or planar motion in a closed water environment.

Acknowledgment

This research is supported by the Ministry of Science and Technology of Taiwan, under grant numbers 109-2218-E-110-008, 109-2218-E-110-007, 109-2218-E-110-009, 108-2218-E-110-003, 108-2218-E-110-002, 108-2218-E-110-004, 107-2218-E-110-005, 107-2218-E-110-004, and 107-2218-E-110-006.

References

[1] M. J. Islam, M. Fulton, and J. Sattar, "Toward a Generic Diver-Following Algorithm: Balancing Robustness and Efficiency in Deep Visual Detection," J IEEE Robotics Automation Letters, vol. 4, no. 1, pp. 113-120, 2018.

[2] M. J. Islam, J. Hong, and J. Sattar, "Person following by autonomous robots: A categorical overview," International Journal of Robotics Research, 2019.

[3] F. Mandić, I. Rendulić, N. Mišković, and Đ. Nađ, "Underwater object tracking using sonar and USBL measurements," Journal of Sensors, vol. 2016.

[4] J. Redmon and A. Farhadi. "YOLO: Real-time object detection." https://pjreddie.com/darknet/yolo/ (accessed 2021).

[5] Y.-C. Chou, H.-H. Chen, C.-C. Wang, C.-C. Wang, and W.-H. Chen, "A Hardware-in-the-Loop Simulation Platform for Development of AUV Control Systems," in Proc. 2019 IEEE Underwater Technology (UT), 2019: IEEE, pp. 1-6.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, You Only Look Once: unified, real-time, object detection, in Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779788, Dec. 2016.

[7] https://www.itread01.com/content/1548042482.html

[8] http://www.image-net.org/

[9] http://groups.inf.ed.ac.uk/f4k/