

Article type: Review Article

HYDRIA - An integrated underwater ROV system for archaeological research

Georgia Peleka*; Georgios Zampokas; Giorgos Giouvanis; Konstantinos Paraskevas; Charalampos Polychroniadis; Thanasis Antamis; Michalis Skoumperdis; Ioannis Mariolis; Dimitrios Giakoumis; Dimitrios Tzovaras Centre for Research and Technology Hellas / Information

Technologies Institute, Thessaloniki, Greece.

*Corresponding Author: Georgia Peleka

Centre for Research and Technology Hellas/Information Technologies Institute, Thessaloniki, Greece. Email: gepe@iti.gr Received: Mar 25, 2025 Accepted: May 21, 2025 Published Online: May 28, 2025 Website: www.joaiar.org

License: © **Peleka G** (2025). *This Article is distributed under the terms of Creative Commons Attribution 4.0 International License*

Volume 2 [2025] Issue 1

Abstract

This paper introduces HYDRIA, an advanced underwater Remotely Operated Vehicle (ROV) for archaeological research, featuring modular hardware and sophisticated software for semiautonomous operations. The ROV includes a robust underwater unit with integrated navigation, a versatile robotic arm with haptic feedback, RGB and stereo cameras for high-resolution imaging and depth image acquisition, as well as powerful LED lighting. Processing and communication utilize laptops and embedded GPU units on the surface control station connected in a custom local network. The software framework includes modules for stereo reconstruction, 3D fusion, object detection, and archaeological information visualization. The system enables efficient exploration of submerged archaeological information on images for verification. HYDRIA was tested in real- world conditions in several scenarios at an underwater archaeological site, including 3D reconstruction of archaeological areas, manipulation of tools essential for excavation (e.g. airlift) using the robotic arm, and autonomous detection of archaeological findings.

Keywords: Underwater ROV; Archaeology; Computer vision; Robotic manipulator.

Introduction

Underwater archaeology has undergone a remarkable transformation, driven by rapid technological advancements. This field, historically reliant on the physical capabilities of human divers, now extensively utilizes Remotely Operated Vehicles (ROVs), marking a significant shift towards high-tech exploration methods. ROVs, originally developed for oceanographic research and offshore operations, have become pivotal in uncovering and preserving the mysteries lying beneath the ocean's surface. Nowadays, Remotely Operated Vehicles (ROVs) have wide applications in underwater applications such as submarine warfare [1], oil and gas exploration [2], marine science [3] and underwater archaeology [4]. Much research work has focused on the handling of remotely operated underwater vehicles in terms of the degree of human intervention in the handling of the vehicle [5] and the accuracy of the position of moving robotic arms attached to the underwater vehicles [6]. In order to solve these issues, the assistance of computer vision is essential in order to adjust the corresponding sensors. Initially, by locating objects of interest the working area of the robotic arm is defined [7], whereas suitable mathematical models calculate the appropriate torque/force and the slipperiness of the object in order to allow efficient manipulation [8].

The integration of ROVs in underwater archaeology merges history with modern engineering, equipping them with advanced imaging and robotic arms to explore previously inaccessible sites. Cutting- edge imaging, akin to deep-sea surveys, provides unprecedented insights into ancient shipwrecks and **Citation:** Peleka G, Zampokas G, Giouvanis G, Paraskevas K, Polychroniadis C, et al. HYDRIA - An integrated underwater ROV system for archaeological research. J Artif Intell Robot. 2025; 2(1): 1019.

submerged settlements. The intersection of robotics and Al enhances efficiency and scope, enabling detailed and systematic surveys, significantly advancing our understanding of historical events. We introduce "HYDRIA," a customized Underwater ROV for archaeology, integrating essential hardware and software, including perception, navigation, control, communication, and haptic feedback. HYDRIA is considered a mediumsmall ROV compared to established solutions in underwater surveing, such as "Leopard" from Seaeye [10], and "Triton XL" from Searov [11]. Furthermore, smaller ROVs such as "vLBV & vLBC" from Teledyne Marine [12], and "Mission Specialist Defender" from VideoRay [13], either don't utilize a robotic arm, or the robotic arm lacks haptic sensitivity. Experimental results from the pilot deployment in Island Modi, Eastern Trizinia, Greece, verify its capabilities in pottery shard verification.

Related work

In [5], authors propose a novel joystick with closed-loop control functions for ROVs, aiming to simplify and enhance the efficiency of manual piloting tasks like inspection or complex manipulation. The joystick inputs generate desired velocities or positions, controlled by a closed-loop positioning system with feedback, which automatically compensates for ROV dynamics and environmental loads. Full-scale experiments demonstrate the performance of the closed-loop control joystick and high-lighting its potential in ROV operations.

The work in [9] presents the implementation and results of the StereoFusion algorithm for real- time 3D dense reconstruction and camera tracking in underwater environments. Unlike KinectFusion, which it is based on, StereoFusion utilizes a stereo camera to generate a depth map, which is then used to incrementally build a volumetric 3D model of the environment. This model aids in camera tracking, enhancing the capabilities of underwater ROVs. The ROV recognizes the object and its exact position in order to be more autonomous from the human agent. The algorithm has been successfully tested in both lake and ocean settings with different ROVs, and future work is focused on adapting it for acoustic sensors and developing a vision-based monocular system with similar functionalities.

The work in [14] details a stereo-vision system designed for underwater 3D object detection, along with a novel method for camera calibration that avoids complex underwater procedures. It is focusing on underwater cooperative intervention tasks. Utilizing state-of-the-art technologies, computer vision methods and ROS middleware, the system was tested on underwater stereo images of cylindrical pipes to calculate the energy consumption, heat dissipation and the efficiency of the camera calibration network.

The study presented in [8] focuses on implementing force control algorithms for an electrically driven manipulator mounted on a ROV, specifically for underwater operations in aquaculture. The manipulator is unique in its use of electrical systems over the more common hydraulic systems. The paper describes laboratory experiments with a two-degree-of-freedom manipulator, implementing impedance and admittance control, and using CAD modeling for calculations. The manipulator arm utilizes a mathematical model based on the Newton-Euler equations to replace the underwater manipulator of the system to create a semi-autonomous system. The work in [6] presents an innovative three-finger underwater gripper equipped with an integrated force/torque sensor. The gripper is characterized by its advanced actuation system, based on electric servo motors and cable transmission, and is designed to be part of a complex robotic system comprising an AUV and a dexterous 7-DOF arm. The gripper's design aims to facilitate complex manipulation and cooperation tasks in underwater environments, enhancing the capabilities of autonomous underwater vehicles in tasks like object recovery and transportation.

Regarding [7], the paper details the development of a novel optoelectronic force/torque sensor for robotic applications in order to achieve more complex and more subtle manipulations of the arm through the sensors, the number of arms (fingers) of the grasp and the degrees of freedom. Utilizing optoelectronic components as sensible elements, the sensor's design offers a simple yet reliable implementation, making it easy to integrate into various robotic systems like robotic hand fingers or industrial grippers. The paper outlines the sensor's basic working principle and design, provides experimental data to demonstrate its theoretical model and main features, and discusses the implementation of a sensor prototype, including its validation as an intrinsic tactile sensor.

In [4], research focuses on developing a machine learning model, specifically the YOLOv8 model, trained on a custom dataset of underwater videos for identifying ancient pottery fragments near a submerged shipwreck. The goal is to integrate this object detection system into an ROV for automated pottery shard recognition, which could significantly improve the efficiency and accuracy of underwater archaeological excavations and analysis. The paper presents the model's development methodology and comprehensive experimental results.

HYDRIA ROV design and implementation

To design and develop HYDRIA, our Remotely Operated Vehicle (ROV), a rigorous scientific methodology was employed starting with the extraction of user requirements from marine archaeologists. Based on the user requirements the system architecture was formulated, whereas special consideration was given to the selection of every component, assessing their technical specifications, compatibility, and efficacy within the system, which according to the end-user should be lightweight and easily deployable. For this reason, we have selected as a base of our vehicle the BlueROV2 underwater RoV, which is both lightweight (about 12 Kg) and easily expandable with additional hardware.

Through rigorous integration processes and fine-tuning, we've orchestrated a system that utilizes diverse hardware and software elements, resulting in an ROV that exhibits functionality and performance in the assistance of underwater archaeological research. HYDRIA was designed in a way to meet the predefined functional requirements and the required technical characteristics of each individual subsystem and how they are interconnected, to allow seamless communication and performance. An image of the final RoV is presented at Figure 1 and a summary of each subsystem and its functions is given below.

The RoV frame is constructed by High-Density Polyethylene (HDPE) material, known for its high strength-to-density ratio, in order to be both lightweight and strong. This frame consists

the skeleton of the system upon which the individual subsystems are mounted and through which they are powered and interconnected. In the core of the system is the RoV controller which communicates with the operator on the surface via the Q-GroundControl platform, which runs on a laptop computer. The RoV's main system also includes an HD navigation camera from which the operator receives a video stream so that she can guide the RoV underwater. The navigation control module ensures precise maneuverability in challenging underwater conditions.

The power supply and communication subsystem are the subsystem on the surface which converts the power supply of the power grid into the format required by RoV and encodes the communication network so that it can be connected to the RoV and its other subsystems. The transfer of this energy and the communication signals is done with a special cable (tether) that is connected to the RoV and the power supply subsystem.

The USBL subsystem provides the position of the RoV in relation to the operator on the boat, or any other point of deployment. It consists of two modules, one on the RoV (transponder) and one on the surface (receiver), which communicate with each other via sound waves and provide the position of the transponder in relation to the one of the receivers. The RTK-GPS subsystem includes an antenna and an RTK corrections module providing accurate GPS position on a computer at the deployment site. Therefore, by combining RTK-GPS with USBL measurements the precise geolocation of the RoV can be acquired in terms of longitude and latitude. The DVL subsystem is located at the bottom of the RoV. By transmitting and receiving sound frequencies and analyzing their changes due to the Doppler effect the DVL can provide precise measurements of the RoV's movement. Thus, we can then control with greater accuracy the movement of the RoV and we also have the ability to hold the position of the RoV unchanged relative to the seabed.

The subsystem of the robotic manipulator consists of the robotic arm, the controller, ip camera, and the force/torque sensor. The manipulator is located on the RoV, whereas the control interface runs on a computer on the surface. From there the operator can manually teleoperate the arm by having a view from the cameras and access to data such as the positions of the axes, the force/torque values of the haptic sensor and more.



Figure 1: The complete RoV system with the main hardware modules annotated.

The stereoscopic camera subsystem consists of two RGB cameras positioned with a baseline distance of 33.5 cm, a configuration carefully tailored to optimize the objectives of stereoscopic reconstruction. This baseline distance is particularly well suited to our target operating range of 0.5 to 3 m, achieving a balance between depth accuracy and spatial coverage. In addition, this setting is adapted to the dimensions of the remotely operated vehicle, ensuring practicality and stability.

All subsystems are integrated into a unified system with centralized control and exchange of the required data between the subsystems.

ROV perception

This chapter presents the core modules which comprise the perception ability of the HYDRIA ROV for underwater archaeological research. The modules are tailored for execution on a NVIDIA Jetson AGX Orin embedded GPU device, to allow flexibility and energy efficiency. Their software implementations are integrated under the Robot Operating System (ROS) [15] to facilitate communication and data transfer between them.

Stereo reconstruction

Considering the limitations and constraints outlined in the introduction, the method presented in [16] stands out as the most suitable for the task of real-time 3D reconstruction in the underwater domain. This approach utilizes quasi-dense matching, a stereo matching algorithm optimized for speed and capable of achieving real-time performance on modern GPUs. The method involves two key steps. First, in a sparse feature matching stage, a set of sparse features is extracted from the stereo images using a feature detector, such as FAST or Harris Corners. The features are then matched across the images using a robust matching technique, such as Normalized Cross-Correlation (NCC). Then, in the Dense matching stage, the depth information from the sparse feature matches is propagated to a semidense representation of the scene. This is done by propagating depth to neighboring pixels until a predefined similarity threshold is reached. The similarity measure is based on the Zero Mean Normalized Cross Correlation (ZNCC) metric. Therefore, a quasi-dense disparity map of the depicted scene is generated, which can be converted to depth using camera parameters. The method's effectiveness is demonstrated through visual evaluation on underwater data, showcasing its ability to produce accurate and real-time 3D reconstructions in this challenging environment, while remaining robust to noise, occlusion, and specular highlights.

Stereo fusion

To generate a full 3D reconstruction of a scene of interest, a fusion method must be applied to combine information from multiple frames into a single 3D model (Figure 3). This study utilizes the seminal algorithm derived from KinectFusion [17], designed for real-time 3D mapping and tracking specifically tailored for Microsoft Kinect and analogous RGBD sensors. In this work, we adapt the method to work with our current setup for underwater scenarios. RGB information is provided by the left channel of the stereo pair, while depth is acquired by the Stereo Reconstruction method presented above.

Similarly to [9], we employ the InfiniTAM [18] framework, which is based on Kinect stereo fusion, a technique employing two depth cameras to construct a 3D model of a scene. These cameras, mounted on a rigid body, capture depth images from

different viewpoints, generating a 3D point cloud that is then quantized into a voxel grid. Visibility of each voxel is determined, and rays are cast through visible voxels to update the 3D model. The software implementation, based on the InfiniTAM framework adapted for ROS and a stereo camera, leverages parallel execution on a GPU for optimal performance. InfiniTAM facilitates real-time 3D reconstruction and tracking from a single RGB-D sensor, utilizing volumetric integration of depth images to build comprehensive 3D models of scenes, particularly beneficial for large or complex reconstructions. However, unlike in [9], in our case the depth images are provided by the real- time stereo reconstruction method presented in the Stereo Reconstruction section above.

Object detection

In underwater archaeological excavations, there is increased motivation for the automated detection of objects of interest. However, the challenging underwater environment necessitates innovative solutions for efficient and accurate object detection. This paper addresses this need by employing state-of-the-art deep learning techniques, with a focus on the You Only Look Once (YOLO) framework, which demonstrates impressive performance [19] in various tasks, based on the methodology presented by [4]. First, a specific dataset for pottery shard fragments and stone anchors is constructed, after careful annotation of pottery shard and anchors instances. Then, the YOLOv8-small model is trained on the aforementioned dataset, following the [4] instructions, until convergence. The model is deployed on the NVIDIA Jetson AGX device, requiring minimal latency.

Archaeological information overlay

Archaeological photography relies on scale indication for precise documentation. Whether using a physical or virtual measuring rod, incorporating scale in photographs establishes a standard for assessing object dimensions and spatial relationships. This ensures accurate analysis, aiding researchers in interpreting size, proportions, and spatial distribution of archaeological features. Without scale indicators, photos lack context, hindering reliable research and understanding of excavated environments. In addition, the use of a compass in underwater photography is essential to provide orientation and spatial context to the acquired images, since underwater environments can be disorienting, with limited visibility and lack of familiar landmarks. The compass image serves as a navigational aid, allowing archaeologists to create accurate maps of underwater sites and aiding in the reconstruction of underwater landscapes, ensuring that the documentation of archaeological discoveries below the surface is geospatially informative.

To calculate the scale of the pixels in an image given the distance of the foreground in the scene (*Depth*), information about the camera specifications, such as the camera's Field of View (*FOV*) and the pixel dimensions of the sensor (*W*) is used. The formula for calculating the pixel scale is as follows:

$$pixel \ scale = \frac{size}{w} \tag{1}$$

where *Size* corresponds to the actual length that an object at distance *Depth* from the camera should have in order to exactly fit withing the width of the image sensor. In that case, *Size* can be calculated by (2)

$$size = 2 \tan\left(\frac{FOV}{2}\right) Depth$$
 (2)

Thus, the *pixel scale* in meters/pixel for a specific distance *Depth* from the camera (estimated from the mean values in meters in the central area of the acquired depth image) can be inferred. Then, the virtual scale (ruler with alternating black and white areas) is constructed according to archaeological specifications, defining white and black intervals of predefined actual dimension. In the context of the u- ArchaeoRoV work, a 10 cm interval is chosen for the virtual ruler, capable of providing sufficient resolution and distinct intervals.

As for the orientation information, it is generated by taking the orientation of the ROV from its compass, which gives the angle from the North and a triangular icon that in its original form the triangle's tip is at 90 degrees (North). The compass angle is used to rotate the triangular icon to correspond to the actual direction of the North according to the ROV's compass.

To visualize the orientation and scale information, we overlay them on top of the captured images. The compass pointing North is overlaid in the upper right corner of the image, and the scale used as a measurement standard isoverlaid in the right part of the image below the compass, as seen in Figure 2.

Deployment feedback from underwater archaeological site

In the context of the research project u-archaeoRov https:// uarchaeorov.eu/, which is co-financed by the European Union and by national funds through the E.P. Competitiveness, Entrepreneurship & Innovation (EPANEK 2014-2020) we have implemented pilot tests of the underwater ROV presented in the previous sections that were successfully completed on November 3, 2023 in the area southeast of Poros.

The pilot application was implemented in the context of the marine archaeological research conducted by the Hellenic Institute of Marine Archaeology (H.I.M.A. - I.EN.A.E) in the Argosaronic Gulf and more specifically on the rocky island of Modi. Since 2009, a Mycenaean shipwreck of the Late Helladic III B-C period (13th/12th century BC) has been excavated at the northern point of the island. The cargo of the merchant ship that was wrecked in one of the most critical periods for Aegean prehistory, consists mainly of ceramic vessels for transporting products.

During the experiments, HYDRIA ROV performed successfully the assistive tasks related with the aforementioned functionalities, including also a collaborative excavation task where the robotic manipulator was holding the airlift tube, while the diver/archaeologist directed with his hand the seabed deposits towards the airlift nozzle (Figure 4).



Figure 2: Simulating the verification of underwater archaeological sample by overlaying orientation and scale information on top of an image of the excavation site at Modi.



Figure 3: The 3D mesh output of the presented stereo fusion process (bottom) from video frames exploring the underwater scene (top).



Figure 4: Underwater photo of the ROV holding the airlift tube while the diver/archaeologist using his hand directs the sea bed deposits towards the airlift nozzle.

Moreover, HYDRIA ROV contributed to the communication of the diver and the support vessel through a mobile phone available underwater (see top of ROV in Figure 4), allowing the diver to communicate with the surface or even access the internet underwater. The above contributes significantly to safety during scientific diving and increase the effectiveness and efficiency of the work of scuba divers. The RoV was also used by the archaeologists for conducting teleoperated survey of the area around the excavation site resulting to the detection of a stone anchor belonging to the Mycenaean shipwreck.

Conclusion

This paper introduces an integrated Underwater Remotely Operated Vehicle (ROV) optimized for archaeological research exploration in underwater environments, assisting archaeologists/divers in their task within the challenging underwater environment. It presents the hardware and software components which were integrated to provide a full system with perception, navigation, control, communication and grasping capabilities, as well as the initial feedback from field tests in an actual underwater archaeological site. This system can also be used in various other use cases, such as marine equipment inspection, or surveillance in underwater infrastructure. Its main advantages over existing methods, is its haptic feedback capabilities of the robotic arm, and the overall size and weight of the ROV compared to another robotic arm equipped ROVs. Moving forward, the system's adaptability and precision pave the way for continued advancements and further experiments.

Declarations

The authors would like to thank the Hellenic Institute of Marine Archaeology (H.I.M.A.) for providing access to their underwater excavations at Modi, Greece, allowing the testing of the developed ROV at the pilot site under their supervision and collaboration.

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project name u-Archaeo-RoV, code: T2EDK- 03656).

The data supporting the results reported in this paper are located in a private repository and can be accessed after contacting the corresponding authors.

References

- 1. JR Benedict. The unraveling and revitalization of US Navy antisubmarine warfare. Naval War College Review. 2005; 58: 93-120.
- T Salgado-Jimenez, J Gonzalez-Lopez, L Martinez-Soto, E Olguin-Lopez, P Resendiz-Gonzalez, M Bandala-Sanchez. Deep water ROV design for the Mexican oil industry, in OCEANS'10 IEEE SYD-NEY. IEEE. 2010: 1-6.
- D McLean, P Macreadie, D White, P Thomson, A Fowler, A Gates, et al. "Offshore Technology Conference Asia," in Offshore Technology Conference Asia. OnePetro. 2018.
- K Paraskevas, I Mariolis, G Giouvanis, G Peleka, G Zampokas, D Tzovaras. "Underwater Detection of Ancient Pottery Sherds Using Deep Learning", in International Journal on Cybernetics & Informatics (IJCI). 2023; 12: 6.
- F Dukan, AJ Sørensen. Joystick in Closed-loop Control of ROVs with Experimental Results. IFAC Proceedings. 2012; 45: 293-300.
- G Palli, L Moriello, U Scarcia, C Melchiorri. An Underwater Robotic Gripper with Embedded Force/Torque Wrist Sensor, IFAC-PapersOnLine. 2017; 50: 11209-11214.
- G Palli, L Moriello, U Scarcia, C Melchiorri. Development of an optoelectronic 6-axis force/torque sensor for robotic applications, Sensors and Actuators A: Physical. 2014; 220: 333-346.
- P Leite Machado. Robotic Arm with Base Force/Torque Sensing Towards Underwater Manipulation. Instituto Superior Técnico, Lisboa, Portugal. 2019.

- M Rossi, P Trslić, S Sivčev, J Riordan, D Toal, G Dooly. "Real-Time Underwater StereoFusion" Sensors. 2018; 18: 3936.
- 10. https://www.saabseaeye.com/uploads/leopard_datasheet_ rev1_%281%29.pdf
- 11. https://global-uploads.webflow.com/61bc9c83993276d6130d6 aa1/620b77d5c292cf6daf9f927b_59%20_%20Triton%20XL%20 Datasheet.pdf
- 12. https://www.teledynemarine.com/en-us/products/SiteAssets/ SeaBotix/SeaBotix%20vLBV300%20Data%20Sheet-2021.pdf
- 13. https://videoray.com/wp-content/uploads/2023/04/2022-Spec-Sheet-MSS-Defender-4- 20.pdf#page=1.00
- 14. F Oleari, F Kallasi, DL Rizzini, J Aleotti, S Caselli. "An underwater stereo vision system: From design to deployment and dataset acquisition," OCEANS 2015 - Genova, Genova, Italy. 2015: 1-6.
- 15. Stanford Artificial Intelligence Laboratory, et al. Robotic Operating System. 2018. Available at: https://www.ros.org.

- G Zampokas, K. Tsiolis, G. Peleka, I. Mariolis, S. Malasiotis and D. Tzovaras, "Real-Time 3D Reconstruction in Minimally Invasive Surgery with Quasi-Dense Matching," 2018 IEEE International Conference on Imaging Systems and Techniques (IST), Krakow, Poland. 2018: 1-6.
- RA Newcombe, et al. KinectFusion: Real-time dense surface mapping and tracking, 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel. Switzerland. 2011: 127-136.
- Kähler O, Prisacariu VA, Ren CY, Sun X, Torr P, Murray D. Very High Frame Rate Volumetric Integration of Depth Images on Mobile Devices. IEEE Trans Vis Comput Graph. 2015; 21: 1241–1250.
- 19. Terven J, Cordova-Esparza D. "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond." 2023. arXiv preprint arXiv:2304.00501.
- 20. https://www.dropbox.com/s/0j0puotbfs7dj9s/br_bluerov2_ datasheet_rev2022-R4ROV.pdf?dl=0