International Journal of Computer Applications (0975 – 8887)
Volume 187 – No.28, August 2025

The Contributions of AI Usage Over AUVs' Path-Following and Path-Planning

Abeer Ali Sirelkhatim
School of Mechanical & Electrical Engineering, UESTC
Chengdu, China

Mustafa Osman Ali
Department of Electrical & Electronics Engineering
Nile Valley University, Sudan

Bei Peng
School of Mechanical & Electrical Engineering, UESTC
Chengdu, China

ABSTRACT
The increasing of interesting in exploration of the underwater environment unknown and complexity, accentuated the need of the underwater vehicles with trusted robust control systems. Over a few pasts dedicates, much of researchers and scholars shown a huge racing for designing and implementing navigation systems supporting autonomy for underwater vehicles.

This paper will explore the basic concepts of Autonomous Underwater Vehicle (AUV) control systems and terms. Two different criteria of algorithms for AUVs' path trajectory will discussed and explained (Path Following and Path Planning).

Also, some path trajectory algorithms which have been designed with aiding of AI techniques are discussed through this research work; Where the study shows the similarities and differences between different types, and then assesses the benefits gained from the use of AI technology.

General Terms
Path following – Path planning, Reinforcement Learning – Deep Reinforcement Learning – degrees of freedom (DoF).

Keywords
Autonomous, unmanned, trajectory, rewards, agent, obstacle.

1. INTRODUCTION
With the development of large-scale and complicated automation systems, networked control has attracted significant attention in the past two decades. In view of various applications, coordinated motion control serving as an important branch of networked control has been widely studied [1]. Roughly speaking, in recent years, the role of autonomous unmanned vehicle (AUV) has become more and more important. Equipped with a series of various technical sensors, AUV can conduct continuous operation without human supervision in specific environment. In addition, it can work independently adjusting to the changes of the environment to complete its goal task. Due to the remarkable maneuvering characteristics and high cost-effective advantages, AUV can replace personnel in high-risk environments [2]. In general, the less investment, good maneuverability and flexible control, give AUV wide applied in many fields, such as scientific observation, resource investigation, oil and gas engineering, military applications etc., [3]. With the continuous development and maturity of AUV technology, developing algorithms to address AUV tasks has become a research hotspot in the latest years. Generally, AUV is advantageous over human due to its deployment simplicity, low cost and low risk, and can even complete dangerous tasks in unknown complex environments that might be inaccessible to human beings [4].

2. AUTONOMOUS VEHICLES
The term autonomous agent generally refers to an entity that makes its own choices about how to act in a specific environment without any influence from a leader or a global plan. According to Craig W. Reynolds on his article [5]: {The term "autonomous agent" is used in many contexts, so the following is an attempt to locate the terminology of this paper in relation to other fields of study. An autonomous agent can exist in isolation, or it can be situated in a world shared by other entities. A "data mining" agent is an example of the former, and a controller for a power grid is an example of the latter. A situated agent can be reactive (instinctive, driven by stimulus) or it can be deliberative ("intellectual" in the classic AI sense). An autonomous agent can deal exclusively with abstract information ("softbot", "knowbot", or "information agent") or it can be embodied in a physical manifestation (a typical industrial robot or an autonomous vehicle). Combinations of situated, reactive, and embodied define several distinct classes of autonomous agents}.

2.1 Definitions and Terms
There is, as of yet, a lack of consistency in the nomenclature and taxonomy of unmanned vehicles (UVs). As per the précis authors of [6] and the Expert commentators of their work; they have adopted the following terms. Here is an exploration of common acronyms, synonyms and key terms as per Brendan Gogarty et. al. collection in their article [6]:

2.1.1 UVs
Any vehicle which operates without a human direct physical contact.

2.1.2 UV variants
The four acronyms used to describe UVs operating in different environments are:

2.1.2.1 UAVs (unmanned aerial vehicles),

2.1.2.2 UGVs (unmanned ground vehicles),

2.1.2.3 USVs (unmanned [water] surface vehicles), and

2.1.2.4 UUVs (unmanned underwater vehicles).

2.1.3 UCV variants
Refers to weaponized UVs. UVs designed specifically for this purpose usually include the term 'combat' within the acronym; hence a UCAV is an unmanned combat aerial vehicle.

2.1.4 Drones
The term 'drone' is arguably the most common and widespread synonym for UVs. In particular it is used to refer to unmanned aerial vehicles (UAVs).

2.1.5 Remote vehicles
These generally refer to vehicles over which a human has direct, albeit remote, control. For instance, a human operator receives visual images from cameras or sensors on-board a UV and steers it by cable (tethered control) or wireless signal (remote control). This form of human/machine interface is referred to as 'teleoperated' control.

2.1.6 Robotics
The more autonomous forms of UVs are often referred to as robots or robotic systems. The Oxford English Dictionary (OED) describes a robot as "a machine … designed to function in place of a living agent, esp. one which carries out a variety of tasks automatically or with a minimum of external impulse".

2.2 Vehicle Autonomy
Unmanned Vehicles (UVs) vary in their form and complexity, but perhaps the most important distinguishing feature is how much a UV can operate wisely without human control and direction.

Modern UVs are all 'controlled' to one degree or another; however modern technology platforms and artificial intelligence (AI) give UVs the capacity to function without direct human intervention. UAVs in current use can, for instance, be set general patrol coordinates and then left to pilot themselves; while surveillance UGVs can independently patrol long stretches of border, only alerting a human controller when suspicious activity is detected [6,7].

According to the Society of Automotive Engineers (SAE) and the National Highway Traffic Safety Administration (NHTSA) [8,9,10], vehicle autonomy is classified into six levels, ranging from Level 0 (no automation) up to Level 5 (full automation). At Level 0, the vehicle is entirely controlled by a human driver. As reader progresses through the levels, the degree of autonomy increases, with Level 5 vehicles being capable of fully autonomous operation in their specific environments and conditions without any human intervention. These six levels of automation are as following:

2.2.1 Level 0 (No Autonomy)
The UV is entirely tele-operated by a human.

2.2.2 Level 1 (Robot Assistance)
The UV provides some automated functionality, for example staying at a set depth (set by the operator) or prohibiting the operator to maneuver into obstacles. The operator is still in full control of the UV.

2.2.3 Level 2 (Task Autonomy)
The UV is able to execute motions under the guidance of the operator. For example, way-points could be set to which the UV will travel with no further input from the operator.

2.2.4 Level 3: (Conditional Autonomy)
The UV generates task strategies, but requires a human to select which one to undertake. For example, when exploring an environment, the UV may identify several different routes to take, with the human selecting the most appropriate one.

2.2.5 Level 4: (High Autonomy)
The UV can plan and execute missions based on a set of boundary conditions specified by the operator. The operator does not require to select which one the UV should do; however, they are there to oversee the task execution.

2.2.6 Level 5 (Full Autonomy)
The UV requires no human input at all. It is deployed into the environment and left with no operator oversight.

Due to this increasing level of independence, UVs are often referred to as "autonomous vehicles". However, it is clear that, at present, no agent in active military or commercial use is actually "autonomous", in the sense that they are completely independent or self-governing. According to, the UV's control is varied between "semi-autonomous agent" and "fully autonomous agent" [8,11]. The semi-autonomous agents are given broad operating instructions by operators, but are left to carry out routine functions within those parameters, such as navigation or monitoring operations. Critical decisions, such as whether to fire weapons or follow a suspect target off routine patrol paths are currently left to a human operator to veto or directly control. While fully autonomous agents would not require such a human veto. Rather, they would be given general instructions and then left to fulfil their directives according to their programming and artificial intelligence.

As per [12], The wide range of operational contexts implies that truly autonomous vehicles must be able to follow spatial trajectories (path following), avoid collisions along these trajectories (collision avoidance), and maintain a desired velocity profile (velocity control). In addition, autonomous vehicles are often underactuated by the fact that they operate with three generalized actuators (propeller, elevation, and rudder fins) in six degrees of freedom (6-DoF).

3. AUTONOMOUS VEHICLE'S ENVIRONMENT AND APPLICATIONS
Autonomous vehicles are being used in a virous wide scope of applications such as military, industrial and commercial applications. In the military, autonomous vehicles offer navigation, secure communication and reconnaissance. Furthermore, they are being used in mobile edge computing, cellular communication, package delivery, smart healthcare, intelligent transportation systems, video surveillance missions, precision agriculture, power-line inspection, remote sensing, search and rescue, and performing relief operations in disaster environments [11]. Autonomous vehicles have the capability to access remote or dangerous areas, facilitate environmental monitoring and capture high resolution imagery. These wonderful vehicles are helpful in monitoring as they bridge the constraints in limited-access, dynamic, harsh and complex environments.

In article [7] a detailed Comprehensive review of AVs impacts on air, land, water, noise & light pollution. In this article the authors made a review of the existing literature on the environmental impacts of AVs was carried out using the search engines Scopus and Web of Science during August 2020. A first search included the most commonly used keywords to refer to AVs, i.e., "Automated" or "Autonomous" or "Self-driving" or "Driverless" and "Vehicle(s)" or "Car(s)", yielding thousands of references. The search was subsequently extended to terms related to environmental impacts on different physical environments, including noise and light pollution as shown in Table 1.

Table 1. Specific keywords used in the literature review

Keyword	Usage
Air	"Emissions" OR "Pollution" OR "Global Warming" OR "Greenhouse" OR "Carbon" OR "Air Quality"
Land	"Built Environment" OR "Land Use" OR "Urban Form" OR "Territorial Impact"

Water	“Water Pollution” OR “Water Contamination” OR “Aquatic Toxicity” OR “Water Consumption”
Others	“Noise Pollution”, “Light Pollution”

Extracting from Table 1 taken from article [7] which concerns about “Environmental impacts of autonomous vehicles”, rea der can notice clearly the AVs can be used in one of three ideal environments: air, land, and water; where each environment has its special considerations that should be handle with.

Of Course, no one should forget an out-earth (space) environment which has its special conditions that must be taken into consideration. As per Michael D. Watson et. al. [13] “Human exploration outside of the Earth planetary system (beyond Earth orbit) requires autonomous operation of the vehicle to deal with communication latencies, crew size limits, and vehicle complexity. A fully autonomous vehicle of this complexity will require multiple autonomous algorithms working cooperatively within a set of mission objectives and system constraints. The understanding of the physics of the systems, system interactions, and environmental interactions is essential to the system engineering of this complex system”. Here for an example, a description of a space unmanned vehicle that developed by European Space Agency (ESA) is explored:

“The Automated Transfer Vehicle (ATV) being developed by ESA is an unmanned vehicle that can be configured to provide the International Space Station (ISS) with up to 5500 kg of dry supplies (e.g., hardware, food and clothes) and liquid and gas supplies (up to 840 kg of water; up to 100 kg of gases (air, nitrogen, oxygen); up to 860 kg of refueling propellant). ATV can provide propulsion support to the ISS by using up to 4700 kg of propellant. The total net payload is estimated to be at least 7500 kg. Finally, ATV can also remove up to 6500 kg of waste from the Station” [14]. This detailed description informs the reader how much huge and importance of this ATV, also he can notice the special demands of the environment where ATV serves in.

This article focuses on Autonomous underwater vehicles (AUVs); Where autonomous term is widely used than unmanned term although they are likely same. AUV is any vehicle that is able to operate underwater without a human occupant. The first device, which can be classified as AUV, was developed in 1957 in the USA, Applied Physics Laboratory, University of Washington and named SPURV (Special Purpose Underwater Research Vehicle), designed to research in the Arctic waters. SPURV hull was made of aluminum and its shape is somehow like a torpedo. SPURV had classical hydrodynamic shape and it was driven by a screw. Control of this AUV was carried out by means of acoustic communications. SPURV had been successfully used in oceanographic research until 1979, and declare itself as a reliable and practical tool to explore the ocean [15].

4. AUVS CONTROL TECHNIQUES

AUVs, need a robust navigation system that capable to face underwater challenges because of three main difficulties. First of all, AUVs are highly nonlinear multi-input multioutput systems with strong coupling and time-varying hydrodynamic coefficients of dynamics. Secondly, AUVs and environment models are often poorly known. Thirdly, most AUVs are designed as underactuated, that is, their degrees of freedom (DoF)s are greater than the number of independent actuators. All of these make it necessary to further study trajectory tracking problem of AUVs. The main issues of the AUV’s navigation system are: localization, positioning, path tracking, guidance, and control during a long period of duty cycle. The complexity that arises when combining the control objectives, a complicated hydrodynamic environment and disturbances, and the physical design with three generalized actuators spurs an intriguing control challenge [12,16]. Therefore, in order to develop an accurate and robust navigation and control system for an AUV, it is necessary to derive an adaptive algorithm for estimation of AUV dynamics.

Two famous path controlling techniques for AUVs are widely used: Path Following Technique and Path Planning Technique. each of these two techniques have its own passionate researchers and developers.

4.1 Path Following Technique

Due to the AUVs’ impressive maneuverability and versatility, they have been extensively used in a variety of underwater tasks, including pipeline tracking, seafloor mapping, underwater structure inspections, and military operations. To accomplish these tasks effectively, accurate path-following control needs to be guaranteed [17].

However, today's marine applications put forward higher and higher requirements for the autonomy of AUV. The AUVs that usually do not have good autonomy and are generally limited to pre-planning or pre-programming tasks. They work well in known and structured environments, but not in uncertain and dynamic ones. Therefore, to realize the autonomy of AUV, it is necessary for it to have strong abilities of environmental perception and understanding, adjustment of control policies, and task planning. The path planning and following of AUV, which determines the application prospect of AUV in the marine _eld, can only be realized with accurate control technology, in consideration of its energy consumption, motion characteristics, speed constraints, etc. [3]. Therefore, autonomous control that can adapt to the changes of marine environment is the core technology to realize the autonomy of AUV. Simply, path following technique is to provide the AUV with a robust mechanism capable to configure the detailed characteristics of the path that must be followed, hence the AUV should trace the segments of the path periodically to correct its direction till reach its target end. Finally, path following control aims at forcing a vehicle to converge to and follow a desired spatial path, without any temporal specifications [18].

4.2 Path Planning Technique

At present, the most of AUVs used for deep-water exploration are underactuated AUVs. It’s only including stern thruster generally, and steering and pitching are realized through vector propulsion or rudder. Path planning is one of the core problems in the underactuated AUV fields. Its purpose is to find an optimal path from the beginning to the end. The path planning environment is either static or dynamic. In a static environment, the global environmental information such as terrain, obstacles and disturbances are known and a path can be planned ahead of the detection. However, for the dynamic environments, the global environmental information is unknown and the path needs to be planned in real-time [19]. Relatively speaking, the real time path planning in the dynamic environments has more practical significance and great difficulty.

Path planning for AUVs is broadly categorized into two sub-areas: global path planning and local path planning. A global path planner employs known information about the operational environment to return a path from the start point to the goal while avoiding any obstacles. In contrast, a local path planner recalculates the path returned by the global path planner as needed to avoid unexpected moving obstacles such as ships, boats, swimmers, other AUVs, etc. [20,21].

Always, efficient path planning has always been a challenging task for underwater vehicles, and aimed at finding an optimal path from the initial to target points with full consideration of path length, safety and smoothness. In the latest years, many researchers paid more attention to optimal path planning and environment adaptation. And so does this article, its main research work which this overview made for, is mainly focuses on path planning and environment adaptation for AUV. In spite of many successful advances in path following and path planning, there are still some limitations for the practical engineering [21]: For path following and tracking: There were many points that should be taken in the consideration to work on and to be improved. First, some way points-based path following methods ignored the derivation of desired position, and the Line-of-Site (LoS) law showed unsmooth characteristic during switching stage. Second, some controllers based on the Lyapunov function existed the singularity of yaw angle in surge law. In practical underwater engineering, there will usually be a yaw angle error of 90◦, resulting in damage to the actuated mechanism with these methods.

In the other hand, path planning: Most traditional path planning methods might only consider the issues in static environment, which required exact map information before planning, and usually could not achieve real-time planning. However, most real underwater environments are dynamic with movable obstacles, bringing great challenges to the planning algorithms. Consequently, it is significant to explore a utility algorithm to achieve the real-time path planning in complex dynamic environments.

Either path following or path planning the problem of imbuing agents with a decision-making mechanism can be hard to tackle as the designer himself does not have a clear idea of what the agent should do, or does not have the time to provide all the possible acting options. Hence, the need for machine learning techniques arises [20-23].

5. MACHINE LEARNING TECHNOLOGIES

Due to the challenges and complexity of the underwater environment, and the large number of the path factors tests and their rapidly varies, therefore a large number of control methods to overcome these difficulties have been proposed by researchers over the last decades. Different motion controllers for trajectory tracking, way-point tracking, path following, and path-planning in the literature are existed. Such as examples for autonomous based on the motion controllers for trajectory tracking system are the work carried out by Juan Li et. al. [22] and L. Lapierre et. al. [18]. Both of these works built on using Lyapunov theory. In these works, in order to solve the multi-UUV formation tracking problem. Firstly, a single UUV kinematics path-tracking controller and a dynamic path tracking controller are designed by using backstepping techniques and Lyapunov theory. Juan Li et. al. reported that: “the formation method based on the leader and the virtual structure can handle the flexibility of the formation very well. Each UUV can track the expected speed of the formation very well. Making up for the inadequacy of the virtual structure formation method and leader-follower formation method. The algorithm implements three UUVs to track the desired path in a certain curve formation. The formation control mission of UUV was well achieved”.

According to Wenjie Shi et. al. [16], in studies about trajectory tracking problem of AUVs, the dynamic models are often decoupled or linearized to enable potential applications of various classic controllers. But these researches are shown limitation in solving some of path tracking or its planning issues, neither overcome all the issues.

The use of machine learning (ML) techniques in overcoming some of the challenges associated with AUV path planning problems such as safety and obstacle avoidance, energy consumption, and optimal time and distance travelled remains an active research area. The (ML) algorithms are classified as: under supervised, unsupervised, and reinforcement learning [24].

5.1 Reinforcement Learning Technology

Later since a few years, agent learning based on Reinforcement Learning (RL) has been introduced into the AUV design and research to improve its autonomy. Reinforcement Learning is a method for an agent controller to learn optimal control policy through interaction with the environment [3]. The policy defines which action the controller should take when the agent is in a certain environmental state. Under the current policy, after the controller tries to select and execute an action in a certain state, it will receive a reward signal provided by the reward function defined in advance by the designer in the environment. This reward signal reflects the quality of the actions performed by the controller and is used to update the control policy. The controller main task is to learn a policy that maximizes the total cumulative reward.

As a branch of machine learning, (RL) is usually modeled as a Markov decision process (MDP) [3], that mainly consists of five elements: agent, environment, state, action and reward. In (RL), an agent interacts with the environment by acquiring the environment state, performing actions and obtaining rewards. The scenario of the (RL) mechanism exactly is as shown in Figure 1. Where (s_t) represent the environmental state at time (t), and (a_t) is the agent performs an action after obtaining the state (s_t), and the environmental state is transformed from (s_t) to (s_{t+1}) at time ($t+1$). Then after the environment generates feedback reward (r_{t+1}) to the agent in the new state (s_{t+1}). The agent will update the learned policy with the reward signal and perform a new action (a_{t+1}) in the new state. The agent will optimize the policy by continually interacting with the environment until an optimal policy is learned. The agent's goal is to maximize the long-term cumulative rewards.

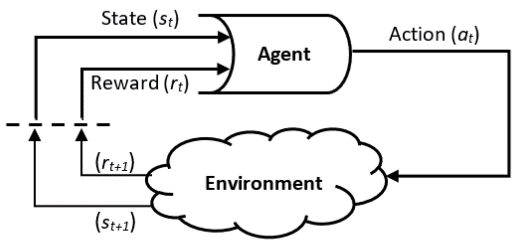

Fig 1: The main framework of (RL) algorithm

In (RL), a general policy (Π) maps states to actions and has Markov property. The probability of taking action (a) in the current state (s) is only related to the current state, and has nothing to do with other factors [3]. The policy in general can be formulated as:

$$\Pi\,(a \mid s) = p\,(a_t = a \mid s_t = s). \qquad (1)$$

5.2 Deep Reinforcement Learning Technology

Deep learning (DL) is a branch of machine learning based on a set of algorithms that attempt to model high level abstractions in data by using multiple processing layers, with complex structures or otherwise, composed of multiple non-linear transformations. (DL) also known as deep structured learning, hierarchical learning or deep machine learning [24].

Recently, Runsheng Yu et al. [25] applied the latest Deep Reinforcement Learning (DRL) to AUV path following task. DRL combines the advantages of Deep Learning (DL) and Reinforcement Learning, and can realize the end-to-end autonomous learning and control with the raw high-dimensional environment perception information input to the behavior action. Runsheng Yu et al. claimed that AUV with DRL can achieve better control effect than a PID controller in simulation experiments. However, because it is difficult to define an effective reward function and it usually provides very sparse reward signals, the agent needs a lot of time and samples to explore and test before learning an effective control policy. Therefore, the authors of [25] reported that: the traditional (RL) and (DRL) methods are still difficult to apply directly to the actual AUV system.

In a similar manner, Zhenzhong Chu et. al. [19] have proposed a DRL path planning method based on double deep Q Network (DDQN) In order to improve the AUV's path planning capability in the unknown environments. Their work is created from an improved convolutional neural network, which has two input layers to adapt to the processing of high-dimensional environments. Zhenzhong Chu's team has implemented their proposed algorithm on an actual AUV. They declared their algorithm can achieve path planning, and also can achieve better planning effectively in the known environments with assistance of NURBS. In addition, it is more suitable for real-time path planning.

In order to speed up the agent learning, researchers of [26] propose interactive reinforcement learning (IRL) based on reward shaping in traditional reinforcement learning. Interactive reinforcement learning allows designers and even non-technical personnel to train agents by evaluating their behavior. In this way, human experience and knowledge can be embedded into autonomous learning of agent to speed up its learning.

6. AI-TRAJECTORY RECOGNITION

A great an opportunity for powerful control system is raised up based on artificial intelligence and neural networks. This intelligent system can be adaptive control systems or control systems based on environmental measurement makes AUV more autonomous and give it the opportunity to choose the further purpose of following and build a route around obstacles. In this part of article different previous research work regarding controlling path following and path planning are discussed.

6.1 AI Aid for Path-Following

Yexin Fan et. al. based on the RL framework they introduce an improved TD3 algorithm to address the AUV path-following control problem (Note: TD3 stands for Twin-Delayed Deep Deterministic policy gradient algorithm. It's a model-free, online, and off-policy reinforcement learning algorithm based on the DDPG). The developed algorithm's main structure and dataflow are lies on a basic TD3 module that contains six neural networks, as well as two essential components that have been enhanced to improve the algorithm's performance: an improved experience replay and the policy smoothness regulation. Specifically, the improved experience replay employs both TD errors and episodic return to evaluate the significance of each experience and concurrently sampled experiences based on both sets of significances. The policy smoothness regulation introduces a dynamic parameter, which adaptively adjusts the smoothness constraint based on the AUV's current state [17].

Authors Yexin Fan et. al. declared that their simulation results showed that the proposed algorithm achieved faster convergence speed and better tracking results than other RL agents, including DDPG, vanilla TD3, and TD3-PER. In addition, their algorithm showcased better generalization capabilities across varying path configurations and exhibited superior robustness in handling uncertainties and disturbances, which outperformed the commonly used controller for AUVs. The simulations and real-world experiments have demonstrated the superiority of the proposed approach. Trajectories of AUV path following under different shaped paths using the control policy trained via TD3-IMP and the PID controller are shown in Figure 2.

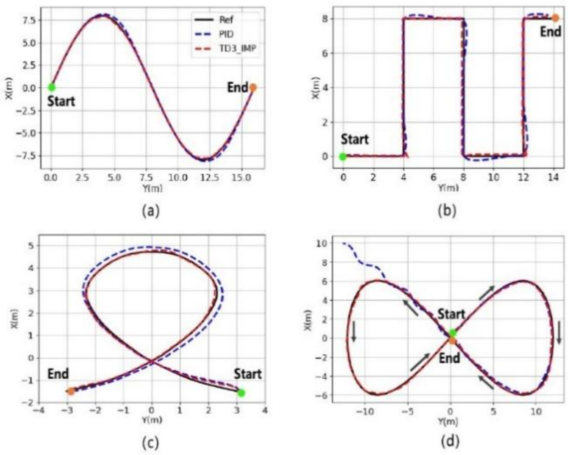

Fig 2: Trajectories of UUV path following trained via TD3-IMP and the PID controller as per [17]. (a) Sine-wave path. (b) Comb scanning path. (c) Closed curve path. (d) Eight-shaped path.

By Figure. 2, authors announced that both the PID controller and the TD3-IPM controller can generalize well to simple path-following tasks in (a) and (b). However, for more complicated paths (c) and (d), the performance of the PID controller shows an obvious degradation, where overshoots are exhibited in (c) and being unable to finish the task of (d).

Qilei Zhang et. al in [3], proposed a deep interactive reinforcement learning method for path following of AUV by combining the advantages of deep reinforcement learning and interactive RL. In addition, since the human trainer cannot provide human rewards for AUV when it is running in the ocean and AUV needs to adapt to a changing environment, they further propose a (DRL) method that learns from both human rewards and environmental rewards at the same time. They test their methods in two path following tasks: straight line and sinusoids curve following of AUV by simulating in the Gazebo platform. They reported that their experimental results show that with their proposed deep interactive RL method, AUV can converge faster than a DQN learner from only environmental reward. Moreover, they believe AUV learning with their deep RL from both human and environmental rewards can also achieve a similar or even better performance than that with deep interactive RL and can adapt to the actual environment by further learning from environmental rewards. The main results of this work are shown in Figure 3.

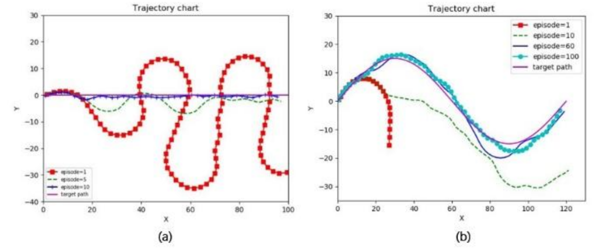

Fig 3: Path following of AUV's simulating in the Gazebo platform as per [3]: (a) straight line and (b) sinusoids curve

An analysis and modeling of AUV 3D path tracking problem were accomplished in a case study of a cylindrical helix path for the work done by Yushan Sun et. al. in the article [27]. In this work, the training environment for 3D path tracking was designed by applying deep reinforcement learning DDPG algorithm. A method of selecting actions based on positive distribution was adopted to maintain the exploratory action selection. The rudders angles and their rates of change was added to be the new term in reward function, and a boundary reward was also designed to form a part of the reward function. The new reward function was shown to be effective to lower the frequency of steering. The LOS method with the integral term added was adopted to provide an indication of the target course angle and target flight path angle. Furthermore, to enable the controller to observe the current disturbance and adjust outputs, a currents disturbance observer was proposed. The observer was found to perform very well in terms of anti-disturbance.

In the end of their work, Yushan Sun et. al. reported that: "Training and simulation experiments about the cylindrical helix path tracking were carried out. The controller proposed in this article was proven to be successful in high-precision path tracking, and the anti-disturbance ability and convergent speed were improved.". Figure 4 shows the path tracking performances of simulations that done through [27] which applying three kinds of different controllers.

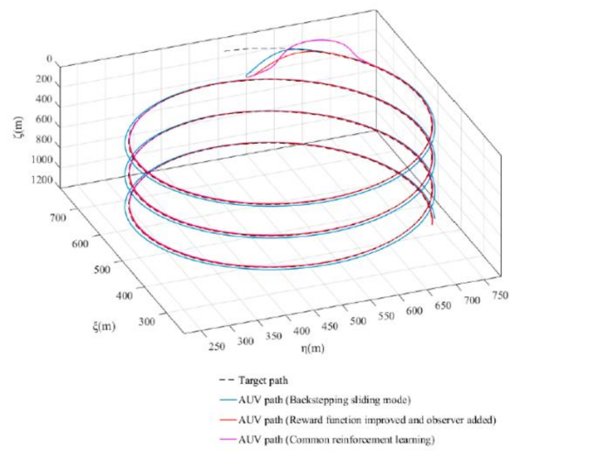

Fig 4: The path tracking performances results of simulation as per [27]

Some scholars employed the PPO and PID assisted PPO algorithms to control the under-actuated cross-rudder AUV in the trajectory tracking task, and further used the PPO algorithm to complete underwater spatial trajectory tracking and obstacle avoidance tasks simultaneously. Here is an example where DRL agents were trained using state-of-the-art RL algorithm PPO and deployed to tackle the hybrid objective of 3D path following and collision avoidance by an AUV presented by Simen T. Havenstrøm and his colleagues in [12]. Where a curriculum learning approach was utilized to train the agent with increasing levels of complexities starting with path following, followed by the introduction of complexities in the obstacle layouts and ultimately the introduction of ocean currents. The AUV was operated by commanding three actuator signals in the form of propeller thrust, rudder, and elevator fin angles. A PI-controller maintained a desired cruise speed, while the DRL agent operated the control fins. The agent made decisions based on the observation of the state variables of the dynamical model, control errors, the disturbances, and sensory inputs from a Forward-Looking Sonar (FLS). As a conclusion the authors reported that: "From the current studies, it is clear that DLR using curriculum learning can be an effective approach to taming an underactuated AUV with 6-DOF to achieve the combined objective of path following and collision avoidance in 3D.". Figure 5 shows the pure path-following test, as expected, higher λr are better at path following.

But in spite of their valuable amazing results; they reported also: "However, it is also important to stress that despite the demonstrated potential of the DRL approach holds, it will have very limited acceptability in safety-critical applications because the whole learning process happens in a black-box way, thereby lacking it explain-ability and analyzability".

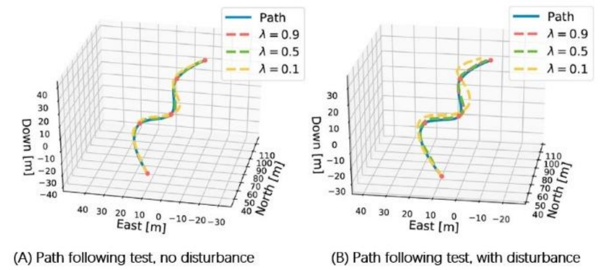

Fig 5: The pure path-following test as per [12]

Yuan Fang et. al. in their paper [2], a DDPG algorithm in DRL method is adopt to train the agent for posture control of the ECA_A9 type AUV based on simulation platform Gazebo and open-source project UUV Simulator. This article trains the agent for posture control of AUV based on the DDPG algorithm. During the training process, the reward function of a single episode starts to converge around the 15th episode, begins to converge from the 40th episode, and fully converges at the 60th episode. Now that the AUV agent has learned to adjust and maintain its posture with θ in the range of (-15◦, +15◦), ψ in the range of (-10◦, +10◦), while keeping φ at 0◦. In this paper, the parameters of stably converged agent model are chosen to complete the subsequent research work. Compared with previous work, it is obvious that the amount of AUV training task in this paper is much smaller, and the DRL agent can be deployed for missions after 80 episodes of training, which takes approximately 1.75 h because of the computing power limitation, rather than 0.75 h in theory. The hardware configuration in this research is Intel i7 7800X CPU and 64G memory.

On this basis as per [2], the position-tracking task of AUV for targets in different orientations in three-dimensional space is completed, achieving a six-degree-of-freedom control of AUV. Additionally, by decomposing the trajectory control task of

AUV in three-dimensional space into multiple position-tracking missions, the trajectory control of AUV in the underwater horizontal plane and underwater three-dimensional space is realized, demonstrating the significant task generalization ability of the control methods proposed; refer to Figure 6.

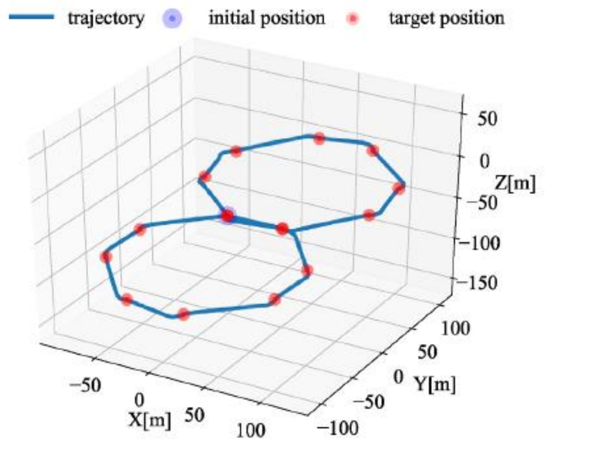

Fig 6: The results of the underwater three-dimensional spatial trajectory control as per [4].

6.2 AI Aid for Path-Planning

AUVs' path planning system generates guidance information such as feasible trajectories after considering the starting and goal points and the following constraints: the shortest possible distance to the target, path smoothness, a safe distance from the obstacles, and environmental disturbances. The control system provides the required forces and moments to satisfy these purposes [28]. AUVs have highly nonlinear uncertain dynamics. Therefore, adaptive control methods are extensively used in the AUV control problems. These vehicles traverse in unstructured and unexplored environments that have been affected by environmental factors such as ocean waves and currents, etc. To provide safe guidance in such an environment, AUVs must calculate or modify their feasible and collision-free paths based on the new information collected from their surroundings [19,28,29]. Hereby, this article will explore three different published papers focusing mainly about path planning approaches.

The purpose of [29] is to establish a safe, real-time, and robust method of collision avoidance that improves the autonomy of AUVs. The proposed method based on active sonar, which utilizes the state-of-the-art deep reinforcement learning algorithm to learn the processed sonar information to navigate the AUV in an uncertain environment. This paper has mainly studied a DRL method that realizes AUV reactive collision avoidance behavior by learning a reward function to determine the mapping between perception information and actions.

Although the algorithm proposed in [29] achieved good results, there are still many problems that have not been solved as authors reported. For example, the sample utilization rate is low, the reward function is too simple, and it is difficult to balance the exploration and exploitation so that the algorithm does not become trapped around a local minimum and instability. In addition, the heading adjustment is too frequent, and the adjustment angle is larger than the deep learning algorithm. The authors declared that they will solve these mentioned problems in their future work; deciding that they will attempt to improve the deep deterministic strategy gradient (DDPG) algorithm.

Article [28] proposes an adaptive motion planning and obstacle avoidance technique based on deep reinforcement learning for an AUV. The research employs a twin-delayed deep deterministic policy algorithm (TD3), which is suitable for Markov processes with continuous actions. Environmental observations are the vehicle's sensor navigation information. Motion planning is carried out without having any knowledge of the environment. A comprehensive reward function has been developed for control purposes.

Behnaz Hadi et. al. authors of [28] used a rectangular area as a training environment. Their training environment includes AUV, obstacles, and target locations. The 1.4-meter AUV is enclosed in a circle with a radius of 0.7 m. The AUV has a nominal speed of 1.5 m/s and a maximum speed of 2 m/s. Each obstacle has a radius of 1.5 m, and the target area has a radius of 3 m. Figure 7 depicts an avoidance zone of 0.5 m around each obstacle. The target point is randomly considered at the border of the rectangle. At the start of each episode, the AUV's initial position is chosen randomly in a 5*5 square zone in the center of the training area. The AUV's heading angle is set to zero. The coordinates of the obstacles and the target are generated at random. A collision occurs when the AUV's circumferential circle and the avoidance circle (red-dash) converge as shown in Figure 7-(b) as per [28].

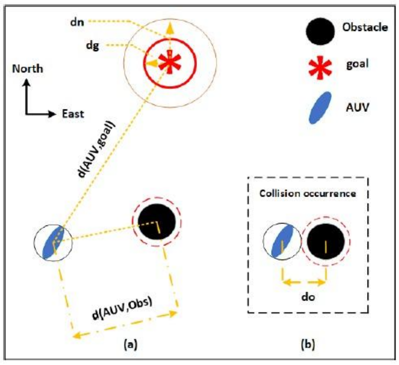

Fig 7: Environment Items as per [28]: (a) Representation of AUV, Obstacle and Target. (b) A collision occurrence

The authors of [28] reported their final result saying that: "Designing the reward function is an essential part of implementing reinforcement learning. It is carried out so that the AUV can produce a short, safe, and directional path towards the target while considering practical constraints such as energy consumption, actuator saturation, and a reduction of the control signal's sudden fluctuations."

A distinguish work for path planning is done by Zhenzhong Chu et. al. [19]. In this work, an improved double deep Q network (DDQN) path planning algorithm is proposed for an underactuated AUV under ocean current disturbance in an unknown environment, and a Non-Uniform Rational B-spline Path Smoother (NURBS) algorithm is adopted to make the route smooth. So, in this work in order to verify the effectiveness of the proposed path planning algorithm, a simulation environment shown in Figure 8 with casually unstructured obstacles is configured. Especially, the ocean current is considered in the simulation environment. The blue-purple areas represent the obstacles. The gray zones indicate that navigation is prohibited within a certain range from the obstacles. This distance is generally set in the range of 3 m to 5 m. The constant ocean current is represented by cyan areas. And then the proposed DDQN algorithm is applied for AUV path planning in the simulated environments. It is compared with the APF algorithm in terms of real-time performance and a variety of state-of-the-art path planning algorithms, such as PRM, RRT/RRT*, GA, PSO and QPSO about the length of the path, travel time and the smoothness rate to validate the efficiency and robustness of the proposed model.

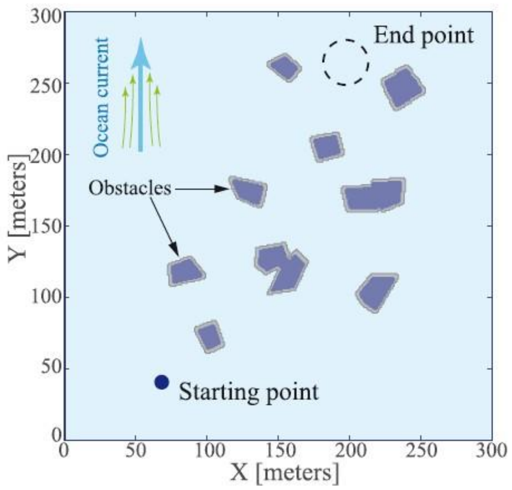

Fig 8: Simulation environments with arbitrarily shaped obstacles considering ocean current as per [19]

The authors of [19] conclusion their results as: “an improved DDQN algorithm of framework was proposed for the path planning of the underactuated AUVs with ocean current disturbances and unknown global environmental information. The improved CNN with two input layers ensures the fusion of different dimensional environment state variables. The dynamic and composite reward function effectively carried out the local path planning with obstacle avoidance under the ocean current disturbance. Simulation and comparison results validated that the improved DDQN algorithm can achieve path planning, and also can achieve better planning effectively in the known environments with assistance of NURBS.”

7. SUMMARY AND DISCUSSION

As per tens of reviewed articles focus in path following or path planning, the squeezed gain valuable information tells that there still need for more research, simulations, and experiments in this area. This exploration of the various different results and recommendations extracted from these articles, shown the future work in trajectory of AUVs is wide open. Specially, with the varies and rapid of technology which encourages everyone interested to join the teams for designing and implementing algorithms for AUVs tracking or planning trajectory.

The reviewed previous work discussed through this paper, were extracted in the Table 2. The aim of this extraction is to highlight the phases of comparisons between different articles models, algorithms, tools, and paths trajectories' environmental planes. The Table 2 is organized in a manner that simplified main details of each article; however, comparing process through all discussed articles become easy and noticeable.

This research review indicates that, the majority of researches have done in path flowing. While the minority focused on path planning. This pointer raised out due to the complexity and unknown nature of underwater environment. The path planning needs to deal with multi changeable parameters with robust algorithm capable to make decision instantly and simultaneously with forward direction correction. Also, other small unknown details may occur while navigates of vehicle needs to be considered, processed, and compared by the controller system with gained rewards for every slot of time. These continuously reprocessed issues, are considered as the main backbone for each AUV trajectory algorithm, and need to be updated always.

8. CONCLUSION

AUV is any vehicle able to operate underwater without a human occupant. Vehicles in underwater environments struggling of three main difficulties: AUVs are highly nonlinear multi-input multioutput systems with strong coupling and time-varying hydrodynamic coefficients of dynamics, AUVs and environment models are often poorly known, and most AUVs are designed as underactuated, that is, their (DoF)s are greater than the number of independent actuators. Therefore, the AUV should has a navigation system capable for: localization, positioning, path tracking, guidance, and control during a long period of duty cycle.

In order to develop an accurate robust navigation and control system for an AUV, it is necessary to derive an adaptive algorithm for estimation of AUV dynamics. Two famous path controlling techniques for AUVs are widely used: Path Following Technique and Path Planning Technique.

According to the authors' deep research over previous work in the area of AUVs trajectory, the majority of researches have done in path flowing. While the minority focused on path planning. Due to the complexity and unknown nature of underwater environment, however, path planning demands consideration of much ponded details at any slot of times while vehicle navigates to reach its target.

As per this review's conclusion, The AI-Techniques have shown high valuable contribution through their various tools on developing and enhancing the techniques of AUV's trajectory models and applications.

9. ACKNOWLEDGMENTS

We would like to express our deep gratitude to ***Dr. Li Min*** for his valuable contributions and keen guidance on our research work topic and simulation.

Table 2: Extraction and highlight some main issues from reviewed articles trajectories

Article	Trajectory	AI-Aid	Results Summary as Authors Reported	Environmental Path Model
[17]	Path Following	TD3-Reinforcement Learning Algorithm	Faster convergence speed and better tracking comparing to DDPG, Vanilla TD3, and TD3-PER.	-Sinewave Path -Comb Scanning Path -Closed Curve Path -Eight-shaped Path
[3]	Path Following	Deep Interactive Reinforcement Learning Algorithm	AUV learning from both human and environmental rewards can also achieve a similar or even better performance than that with deep	-Straight Line -Sinusoid Curve

			interactive RL and can adapt to the actual environment by further learning from environmental rewards.	
[27]	Path Following	Deep Reinforcement Learning DDPG Algorithm	The controller proposed in this article was proven to be successful in high-precision path tracking, and the anti-disturbance ability and convergent speed were improved.	Cylindrical helix Path
[12]	Path Following	Reinforcement Learning Algorithm - PPO	DLR using curriculum learning can be an effective approach to taming an underactuated AUV with 6-DOF to achieve the combined objective of path following and collision avoidance in 3D	Tackling the hybrid objective of 3D path following and collision avoidance by an AUV
[2]	Path Following	DDPG Algorithm	The position-tracking task of AUV for targets in different orientations in three-dimensional space is completed, achieving a six-degree-of-freedom control of AUV. This work done in three-dimensional space into multiple position-tracking missions.	Underwater horizontal plane three-dimensional space structured as eight's shape consist of 14-positions.
[29]	Path Planning	State-of- the-art object detection Deep Reinforcement Learning Algorithm Combined with YOLOv8 model	Although the algorithm is working well and achieved good results, but while utilizing low rate of samples and the reward function is too simple. Therefore, the authors declared some future work that they plan to proceed on.	The proposed method based on active sonar, which utilizes the state-of-the-art deep reinforcement learning algorithm to learn the processed sonar information to navigate the AUV in an uncertain environment.
[28]	Path Planning	Twin-Delayed deep Deterministic Policy Algorithm (TD3)	The AUV can produce a short, safe, and directional path towards the target while considering practical constraints such as energy consumption, actuator saturation, and a reduction of the control signal's sudden fluctuations.	A rectangular area as a training environment includes AUV, obstacles, and target locations.
[19]	Path Planning	Double Deep Q Network (DDQN) algorithm in addition to (NURBS) algorithm is adopted to make the route smooth	Simulation and comparison results validated that the improved DDQN algorithm can achieve path planning, and also can achieve better planning effectively in the known environments with assistance of NURBS.	The improved CNN with two input layers ensures the fusion of different dimensional environment state variables: (ocean current disturbances and unknown global environmental information).

10. REFERENCES

[1] Zhengzhong Chu, Fulun Wang, and Tingjun Lei, "Path Planning Based on Deep Reinforcement Learning for Autonomous Underwater Vehicles Under Ocean Current Disturbance," IEEE Transactions on Intelligent Vehicles, vol. 8, NO. 1, pp 108-120, Jan. 2023.

[2] Wenjie Shi, Shji Song, Cheng Wu, and Philip Chen, "Multi Pseudo Q-Learning-Based Deterministic Policy Gradient for Tracking Control of Autonomous Underwater Vehicles," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, No. 12, pp 3534-3546, Dec 2019.

[3] Yexin Fan, Hongyang Dong, Xiaowei Zhao, and Peter Denissenko, "Path*Following Control of Unmanned Underwater Vehicle Based on an Improved TD3 Deep Reinforcement Learning," in Process IEEE Transactions On Control Systems Technology, vol 32, pp 1904-1919, Mar 2024.

[4] Yuan Fang, Zhenwei Huang, Jinyun Pu, and Jinsong Zhang, "AUV position tracking and trajectory control based on fast-deployed deep reinforcement learning method," Ocean Engineering, vol 245, Feb 2022.

[5] Qilei Zhang, Jinying Lin, Qinxin Sha, Bo He, and Guangliang Li, "Deep Interactive Reinforcement Learning for Path Following of Autonomous Underwater Vehicle," IEEE Access, vol 8, pp 24258-24268, Jan 2020.

[6] Mohammed Imran Chowdhury, and Daniel G. Schwartz, "UUV On-Board Path Replanning using PRM-A," IEEE Global Oceans 2020: Singapore – U.S. Gulf Coast, Apr 2021.

[7] Jian Wang, Zhengxing Wu, Shuaizheng, Min Tan, and Junzhi Yu, "Real-Time Path Planning and Following of a Gliding Robotic Dolphin Within a Hierarchical Framework," IEEE Transactions on Vehicular Technology, vol. 70, No. 4, pp 3243-3255, Apr 2021.

[8] Juan Li, Duorui Xue, and Jainxin Zhang, "Multi-UUV Formation Coordination Control Based on Combination of Virtual Structure and Leader," IEEE International Conference on Mechatronics and Automation (ICMA), pp 1574-1579, Oct 2018.

[9] L. Lapierre, D. Soetanto, and A. Pascoal, "Nonlinear Path Following with Applications to the Control of Autonomous Underwater Vehicles," 42nd IEEE International Conference on Decision and Control,

pp1258-1261, Mar 2004.

[10] Jiwei Hu, Bo Jin, Huiping Li, Wesheng Yan, Mingyong Liu, and Rongxin Cui, "A DMPC-Based Approach to Circular Cooperative Path-Following Control of Unmanned Underwater Vehicles," IEEE 28th International Symposium on Industrial Electronics, pp 1207-1212, Jun 2019.

[11] Zongyu Zuo, Jiawei Song, and Qing-Long Han, "Coordinated Planner Path-Following Control for Multiple Nonholonomic Wheeled Mobile Robots," IEEE Transactions on Cybernetics, vol. 52, No. 9, pp 9404-9413, Sep 2022.

[12] Jintao Su, Jianping Lou, and Xiaolu Jiang, "Overview of intelligent vehicle core technology and development," IOP Conf. Series: Earth and Environmental Science 769, 2021.

[13] Carig W. Reynolds, "Steering Behaviors for Autonomous Characters," website [CteSeer]. Available: https://www.researchgate.net/publication/2495826.

[14] Goncalo Neto, "From Single-Agent to Multi-Agent Reinforcement Learning: Foundational Concepts and Methods," Neto G. Learning theory course, 2005. Available: https://users.cs.utah.edu/~tch/CS6380/resources/Neto-2005-RL-MAS-Tutorial.pdf.

[15] Brendan Gogarty, and Isabel Robinson, "Unmanned Vehicles: A (Rebooted) History, Background and Current State of the Art," Journal of Law and Information Science, vol 21, pp 1-34, 2011.

[16] Kiruthika D, and Jeevishaa S, "Unmanned Vehicles: Navigating the Environmental Implications," JSS Journal for Legal Syudies and Research, pp 178-193, 2024.

[17] Oscar Silva, Ruben Cordera, Esther Gonzalez, and Soledad Nogues, "Environmental impacts of autonomous vehicles: A review of the scientific literature," Elsevier-Science of The Total Environment, vol 830, pp 1-11, Jul 2022.

[18] Richard Nunno, "Autonomous Vehicles: State of the Technology and Potential Role as a Climate Solution," Jun 2021, Available: www.eesi.org/papers.

[19] Syed Agha Hassnain Mohsan, Nawaf Qasem Hamood Othman, Yanlong Li, Mohammed H. Alsharirif, and Muhammad Asghar Khan, "Unmanned aerial vehicles (UAVs): practical aspects, applications, open challenges, security issues, and future trends," Journal of Intelligent Service Robotics, vol 16, pp 109-137, Mar 2023.

[20] Michael D. Waston, Stephen B. Johnson, and Luis Trevino, "System Engineering of Autonomous Space Vehicles," IEEE-International Conference on Prognostics and Health Management, Feb 2015.

[21] P. Amadieu, G. Beckwith, B. Dore, J. P. Bouchery, and V. Pery, "Automated Transfer Vehicle (ATV) Structural and Thermal Model Testing at ESTEC," esa bulletin, vol 111, pp 95-104, Aug 2002.

[22] Salimzhan A. Gafurov, and Evgeniy V. Klochkov, "Autonomous Unmanned Underwater Vehicles Development Tendencies," Elsevier-Procedia Engineering, vol 106, pp 141-148, 2015.

[23] Simon Watson, Daniel A. Duecker, and Keir Groves, "Localisation of Unmanned Underwater Vehicles (UUVs) in Complex and Confined Environments: A Review," *Sensors* 2020, *20*(21), 6203; https://doi.org/10.3390/s20216203, Oct 2020.

[24] Yexin Fan, Hongyang Dong, Xiaowei Zhao, and Petr Denissenko, "Path-Following Control of Unmanned Underwater Vehicle Based on an Improved TD3 Deep Reinforcement Learning," IEEE Transactions on Control Systems Technology, vol 32, pp 1904-1919, Sep 2024.

[25] Behnaz Hadi, Alireza Khosravi, and Pouria Sarhadi, "Deep reinforcement learning for adaptive path planning and control of Autonomous Underwater Vehicle," Elsevier-Applied Ocean Research, vol 129, Dec 2022.

[26] Runsheng Yu, Zhenyu Shi, Chaoxing Huang, Tenglong Li, and Qiongxiong ma, "Deep Reinforcement Learning Based Optimal Trajectory Tracking Control of Autonomous Underwater Vehicle," 36th Chinese Control Conference (CCC), Sep 2017.

[27] Jens Kober, J. Andrew Bagnell, and Jan Peter, "Reinforcement Learning in Robotics: A Survey," International Journal of Robotics Research, vol 32, issue 11, pp 1238-1274, Sep 2013.

[28] Jurgen Schmidhuber, "Deep learning in neural networks: An overview," Elsevier-Neural Networks, vol 61, pp 85-117, Jan 2015.

[29] Faraz Ahmed, Tom Creutz, Christian Ernst Siegfried Koch, and Bilal Wehbe, "A Deep-Learning Approach for Visual Detection of an AUV Docking Station," IEEE: OCEANS 2024 - Halifax, Sep 2024.