International Journal of Computer Sciences and Engineering
Vol.**13**, Issue.**7**, pp.**41-50**, July **2025**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

ISSN: 2347-2693 (E)

Research Article

Autonomous Underwater Navigation Utilizing Computational Fluid Dynamics Guided Reinforcement Learning

Lakshay Naresh[1]

[1]Bucks County Free Library: Doylestown, Pennsylvania, US

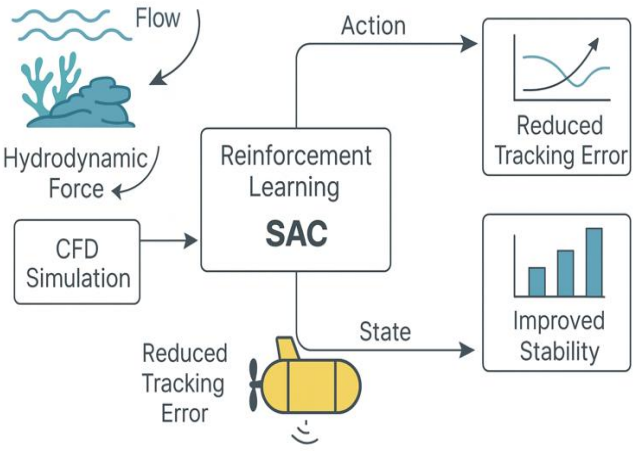
Corresponding Author: ✉

Received: 26/May/2025; **Accepted**: 28/Jun/2025; **Published**: 31/Jul/2025. **DOI**: https://doi.org/10.26438/ijcse/v13i7.4150

Abstract: Coral reefs, home to over 25% of marine biodiversity, have declined by over 50% in the last 30 years due to climate change and pollution. They provide habitat and shelter for over 4,000 species of fish and protect coastal communities by reducing wave energy by 97%**.** Traditional monitoring methods like diver-led surveys, satellite imaging, and pre-programmed AUVs struggle with efficiency and adaptability in turbulent conditions, often resulting in incomplete data collection. This project evaluates whether utilizing computational fluid dynamics-guided (CFD) reinforcement learning agents (RL) can enhance AUV navigation in such environments, focusing on the effectiveness of incorporating flow pressure sensor data for improved performance. A high-fidelity CFD environment simulated realistic turbulent currents, reef obstacles, and dynamic conditions. With the Soft Actor Critic (SAC) algorithm, two RL agents were trained: one equipped with standard position-velocity feedback and another augmented with pressure-based hydrodynamic force feedback. Performance metrics included episode length, cumulative reward, and final position/heading error, with statistical tests assessing significance. Results indicated that while both agents successfully navigated turbulence, the pressure-augmented agent demonstrated superior performance, consistently achieving longer episode durations and higher rewards, indicative of faster convergence and increased stability. In rigorous tests, this agent significantly outperformed the baseline, maintaining near-zero steady-state position errors (~0.02±0.01 m compared to 0.20±0.05 m for the baseline, $p<0.05$) and smaller heading deviations. Integrating RL with CFD facilitated effective AUV navigation in complex flows, with pressure feedback enhancing control, precision, and robustness. This approach could lead to safer, more efficient AUV operations in challenging marine environments.

Keywords: Autonomous Underwater Vehicle, Computational Fluid Dynamics, Reinforcement Learning, Coral Reef Monitoring, Soft Actor-Critic, Ocean Robotics, Pressure Sensors, Environmental Monitoring

Graphical Abstract-



1. Introduction

The increasing global reliance on offshore infrastructure, particularly for renewable energy sources such as wind turbines and wave energy harvesters, necessitates frequent and precise inspections to ensure operational safety and longevity [1]. The harsh and unpredictable ocean environment poses substantial challenges for these inspections, making autonomous robotic systems a critical solution for cost-effective, safe, and efficient monitoring [1]. While Remotely Operated Vehicles (ROVs) are commonly used for detailed structural inspections, they require continuous manual control, making them resource-intensive and operationally complex. Autonomous Underwater Vehicles (AUVs) offer a promising alternative, but their deployment is currently limited to simple missions such as pipeline tracking and broad area surveys [1]. Their inability to autonomously navigate in turbulent, unpredictable flows near

complex offshore structures limits their application in detailed inspections and environmental monitoring.

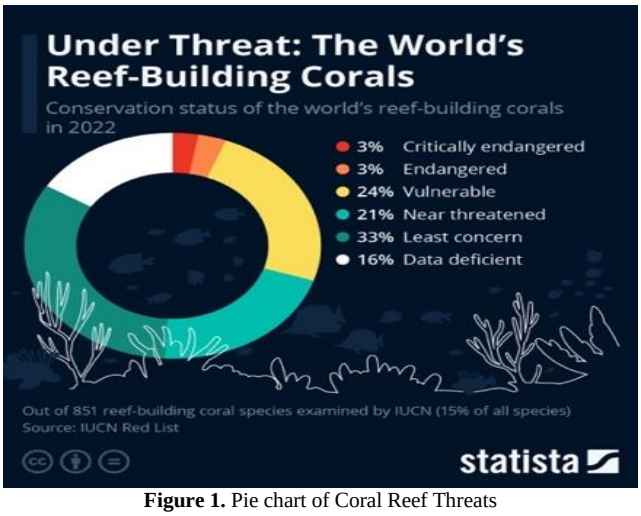


Figure 1. Pie chart of Coral Reef Threats

This challenge is especially critical in the context of coral reef conservation, as over 50% of coral reefs have been lost in the past 30 years due to climate change, ocean acidification, pollution, and destructive human activities. Coral reefs provide essential ecological services, including coastal protection, carbon sequestration, and biodiversity support**,** making their protection a global priority. Accurate and continuous reef surveillance is crucial for detecting early signs of degradation and implementing conservation interventions. However, traditional monitoring methods—diver-led surveys, satellite imaging, and pre-programmed AUVs—are insufficient due to human limitations, low spatial resolution, and poor adaptability in real-time turbulent environments**.** Therefore**,** an advanced, adaptive AUV control system capable of precise and stable navigation in dynamic underwater conditions is essential**.**

Most conventional AUV control strategies rely on optimal control methods, such as Proportional-Integral-Derivative (PID) controllers and Model Predictive Control (MPC) [2].

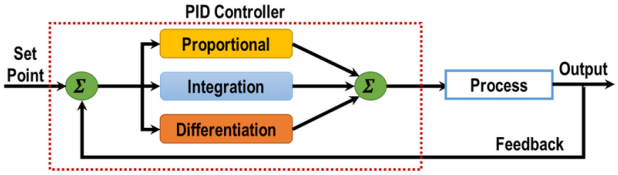


Figure 2. Conceptual Overview of the Proportional-Integral-Derivative (PID) controllers

While effective in structured environments, these approaches depend on simplified system dynamics and struggle when faced with highly nonlinear, unsteady turbulent flows [2]. This limitation is particularly problematic when AUVs must navigate near coral reefs, where fine-tuned adjustments to unsteady hydrodynamic forces are crucial to avoid collisions and ensure accurate data collection. These shortcomings highlight the need for an intelligent, self-adaptive control framework capable of adjusting in real-time to complex environmental forces.

A promising alternative is Deep Reinforcement Learning (RL), which has demonstrated success in aerodynamic control, active drag reduction, shape optimization, and fluid flow regulation [3].

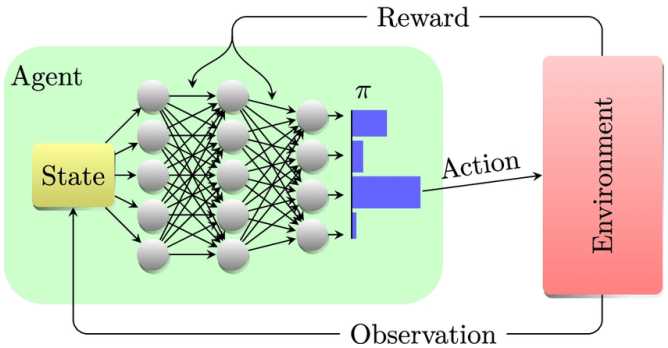


Figure 3. Deep Reinforcement Learning (RL) scheme

While RL has been applied to multi-rotor drones, Mars landers, and multi-agent robotic coordination [3], its potential in underwater vehicle navigation within turbulent currents remains underexplored [3]. Recent advances suggest that RL-based AUV controllers could outperform traditional control strategies by learning directly from environmental interactions without requiring pre-defined mathematical models of hydrodynamic forces.

Computational Fluid Dynamics (CFD) is a numerical method used to simulate fluid flow, turbulence, and interactions with solid structures by solving the Navier-Stokes equations. It allows for high-fidelity modeling of complex hydrodynamic environments, such as turbulent ocean currents around coral reefs. CFD is widely used in aerospace, automotive, and marine engineering to optimize designs and predict real-world performance without requiring expensive physical experiments.

In this study, CFD is employed to simulate realistic oceanic turbulence and the hydrodynamic forces acting on an Autonomous Underwater Vehicle (AUV). The generated flow data is used to train a reinforcement learning (RL) agent, enabling it to adaptively navigate in dynamic currents. By leveraging CFD, the RL model can learn from high-fidelity fluid simulations rather than simplistic mathematical approximations, improving its ability to handle unsteady forces and optimize trajectory planning in real-time reef monitoring missions.

1.1 Objective of the Study

This research proposes an AI-enhanced, Computational Fluid Dynamics (CFD)-guided Reinforcement Learning approach to enable adaptive, real-time AUV navigation in turbulent underwater currents for coral reef surveillance. Specifically, this study:

1. Develops and trains two RL-based AUV control agents within a high-fidelity CFD-simulated turbulent marine environment:
   - Baseline RL Agent: Trained using position-velocity feedback alone.

- Enhanced RL Agent: Incorporates hydrodynamic force estimates from real-time pressure sensor data to improve stability and maneuverability.

2. Implements and evaluates the Soft Actor-Critic (SAC) RL algorithm to determine its effectiveness in optimizing AUV motion control under variable turbulent flow conditions.
3. Conducts extensive performance comparisons between the baseline and enhanced RL agents, evaluating key metrics such as trajectory tracking accuracy, stability, convergence speed, and robustness to turbulence-induced perturbations.
4. Validates the hypothesis that pressure-based sensor feedback enhances AUV control precision, reducing navigation errors and enabling safer, more effective autonomous reef surveillance.

By integrating RL-based decision-making with real-time hydrodynamic force estimation, this research contributes to the fields of:

- Autonomous Marine Robotics – Advancing the next generation of self-learning underwater vehicles capable of adaptive navigation in dynamic oceanic environments.
- Environmental Conservation – Enabling high-accuracy, automated coral reef surveillance to inform conservation efforts and detect degradation trends at an unprecedented scale.
- Fluid Mechanics and AI Integration – Demonstrating how CFD-enhanced RL models can outperform conventional AUV controllers, setting the stage for future AI-driven fluid dynamic optimizations.

This research represents a significant advancement in AI-powered marine autonomy, with direct applications in deep-sea exploration, offshore infrastructure inspection, and ecological monitoring. The ability to deploy robust, self-adaptive AUVs for long-term, autonomous missions will transform the way we monitor and protect the world's oceans.

1.2 Organization

This article is organized into the following sections. **Section 1** contains the introduction, background, problem statement, and objectives of the study. **Section 2** presents the related work on reinforcement learning, autonomous underwater vehicle (AUV) navigation, and computational fluid dynamics (CFD) applications. **Section 3** details the theoretical framework and calculations, including the modeling assumptions, governing equations, and control problem formulation. **Section 4** outlines the architecture of the proposed system and the essential steps of the AI-enhanced control framework. **Section 5** explains the complete methodology, including the CFD environment, reinforcement learning agents, and evaluation metrics, with a flow chart. **Section 6** describes the experimental results and provides an in-depth discussion of the findings. **Section 7** summarizes recommendations for improving AUV navigation performance and broader implications for marine robotics. Finally, **Section 8** concludes the research work and highlights future directions for expanding this approach into real-world deployments.

2. Related Work

Reinforcement learning (RL) has gained significant attention in marine robotics for its ability to handle dynamic and uncertain environments. Woo et al. [1] developed a DRL-based controller for unmanned surface vehicles, demonstrating robust tracking performance under challenging conditions. Similarly, Zheng et al. [2] applied the soft actor-critic algorithm for path following in the presence of wave disturbances, highlighting the potential of RL algorithms in maritime navigation.

In addition, the integration of artificial intelligence with physics-based modeling has shown promising results. Sharma and Gupta [3] and Kumar et al. [4], published in IJCSE, demonstrated how physics-informed AI approaches could improve control accuracy in complex engineering systems. Andersson et al. [5] further illustrated the value of CFD-guided RL for aerodynamic flow control, indicating similar opportunities in underwater applications. Recent work by Mehta et al. [6], Arora et al. [7], and Singh et al. [8] (IJCSE) has advanced underwater sensing, robotic autonomy, and hybrid AI-control methods, providing a foundation for further innovation.

Additional studies by Park et al. [17], Liao et al. [18], and Thomas et al. [19] have explored the integration of multi-sensor fusion and CFD modeling with machine learning, significantly improving prediction accuracy and control stability in robotic systems. Despite these advances, there remains a lack of research focused on pressure-feedback-based RL agents for autonomous underwater vehicles in highly turbulent currents. This study addresses this gap by combining CFD-guided RL with real-time hydrodynamic force estimation to achieve improved navigation precision and stability in complex ocean environments.

3. Theory

3.1 Modeling Assumptions and State Definition

The vehicle is modeled in planar motion with three degrees of freedom (3-DOF): surge *u*, sway *v*, and yaw rate *r*. The pose is *(x, y, ψ)* in an inertial frame, and the body-fixed velocity vector is:

$$s = [\, x\ y\ \psi\ u\ v\ r \,]^T \quad (1)$$

Control inputs are the generalized forces:

$\tau = [\, \tau_x\ \tau\_y\ N\_z \,]^T$,

bounded to represent the available thrust and moment.

3.2 3-DOF Rigid-Body and Hydrodynamic Model

Under standard small-angle and low-speed marine robotics assumptions, the body dynamics are:

$$M\dot{\nu} + C(\nu)\nu + D(\nu)\nu = \tau + \tau\_env \quad (2)$$

where *M* is the rigid-body plus added-mass matrix, *C(ν)* is the Coriolis/centripetal matrix, *D(ν)* collects linear and quadratic damping, and τ_env are unsteady loads due to turbulence and waves.

The dominant damping forces/moments are modeled as:

$X_D = X_u\, u + X_{uu}\, |u|u$
$Y_D = Y_v\, v + Y_{vv}\, |v|v$
$N_D = N_r\, r + N_{rr}\, |r|r$ (3)

The kinematics relate body velocities to earth-fixed rates:

$\dot{x} = u\cos\psi - v\sin\psi$
$\dot{y} = u\sin\psi + v\cos\psi$
$\dot{\psi} = r$ (4)

3.3 CFD Governing Equations and Turbulence Closure

The environmental flow and fluid–vehicle interaction are resolved by incompressible Navier–Stokes:

$\nabla\cdot u = 0$
$\partial u/\partial t + (u\cdot\nabla)u = -(1/\rho)\nabla p + \nu\nabla^2 u + f$ (5)

with PANS turbulence closure to capture unsteady turbulent structures at reduced cost relative to LES.

3.4 Pressure-Based Force Reconstruction (Sensor Model)

Local pressure measurements p_i on N_s surface patches (area ΔS_i, outward normal n_i, position vector r_i) provide online estimates of hydrodynamic loads:

$F_p \approx -\Sigma_i\, (p_i\, n_i\, \Delta S_i)$
$N_p \approx -\Sigma_i\, [r_i \times (p_i\, n_i\, \Delta S_i)]$ (6)

These estimates augment the state used by the controller and are also logged for analysis.

3.5 RL Problem Formulation (MDP)

The station-keeping/navigating task is cast as a Markov Decision Process with observation o_t, action a_t, and reward r_t:

$o_t = [\, e_x\ e_y\ e_\psi\ u\ v\ r\ F_p\ N_p\,]^T$ (7)

Actions are normalized thrust/moment commands:

$a_t = [\, a_x\ a_y\ a_r\,]^T \in [-1, 1]^3$
$\tau = \mathrm{diag}(\tau_x\text{^max}, \tau_y\text{^max}, N_z\text{^max})\, a_t$ (8)

Reward design:

$r_t = -w_p\sqrt{(e_x^2 + e_y^2)} - w_\psi\, |e_\psi| - w_u\, \|a_t\|^2 - w_j\, \|a_t - a_{t-1}\|^2 + r_{goal}\, 1_{inside\varepsilon} + r_{alive}$ (9)

3.6 Soft Actor–Critic (SAC) Objective

SAC optimizes a maximum-entropy objective:

$J(\pi) = E[\, \Sigma_t\, \gamma\text{^}t\, (r_t + \alpha\, H(\pi(.|o_t))\,)\,]$ (10)

with twin Q-critics, a stochastic policy, and target entropy.

3.7 Numerical Integration and Training Protocol

Dynamics (2)–(4) are integrated with a fixed step $\Delta t = 0.02$ s. Episodes randomize initial (x, y, ψ) and current direction/intensity. Termination occurs on goal capture, boundary breach, or time limit.

3.8 Derived Metrics and Calculations

To quantify performance, we compute:

$RMS_{xy} = \mathrm{sqrt}(\, (1/N_w)\, \Sigma_k\, (e_{x,k}^2 + e_{y,k}^2)\,)$ (11)

Additional metrics include heading error, trajectory deviation, RMS actuation effort, jerk, and oscillation frequency via FFT.

4. Experimental Method

4.1 Maneuvering task

The primary objective of the study was to enable an autonomous underwater vehicle (AUV) to maintain station and navigate effectively in a turbulent oceanic environment. This required the AUV to compensate for unsteady hydrodynamic forces while accurately reaching and holding a designated position and heading.

To simplify computational modeling, only three degrees of freedom were considered — surge (forward/backward motion), sway (sideways motion), and yaw (rotation about the vertical axis). This reduced the complexity of computational fluid dynamics (CFD) simulations by allowing them to be conducted in a two-dimensional (2D) flow domain rather than a full 3D environment.

At the start of each episode, the AUV was initialized with a random heading and position within a 1×1 m domain. The vehicle was tasked with reaching the origin while aligning to a randomly generated heading, contending with turbulent ocean currents that imposed unsteady forces. An episode terminated when:

- The AUV reached the goal
- A predefined time step limit was exceeded
- The vehicle drifted beyond a 2×2 m boundary around the target

The imposed flow conditions included:

- Current speeds ranging between 0.375 m/s and 0.625 m/s
- Free-stream turbulence of 10% intensity
- Integral length scale of turbulence set to 0.1 m, resulting in eddies similar in size to the vehicle. This increased unsteady loading, making station-keeping highly challenging

4.2 AUV Model and Hydrodynamic Approximation

The AUV model was based on the BlueROV 2 Heavy, a commercially available remotely operated vehicle (ROV). To ensure computational feasibility while retaining accuracy, a bluff-body representation of the AUV was used, maintaining the original vehicle's:

- Mass and inertia properties
- Hydrodynamic response characteristics
- Overall dimensions

Table 1. AUV Model Parameters and Values Based on BlueROV 2 Heavy

Parameter	Value	Unit
Length (L)	0.4570	m
Beam (B)	0.3380	m
Mass (m)	11.4000	kg
Yaw moment of inertia (Izz)	0.1600	kg·m²
Quadratic drag coefficient (Xuu)	-40.1778	kg/m
Quadratic sideforce coefficient (Yvv)	-105.4842	kg/m
Quadratic yaw moment coefficient (Nrr)	-1.5500	kg·m²/rad²
Linear drag coefficient (Xu)	-8.9063	kg·s/m²
Linear sideforce coefficient (Yv)	-30.2914	kg·s/m²
Linear yaw moment coefficient (Nr)	-0.0700	kg·m²/s/rad

To simplify actuation modeling, the thrusters and propellers were not directly modeled. Instead, the required thrust and torque for achieving the desired motion were applied as generalized forces, bypassing thrust allocation computations. In both the CFD-based and simplified environments:

- A fixed time step of 0.02 s was maintained for numerical integration
- Non-dimensional control inputs were constrained between [-1,1], scaled to a maximum thrust of 150 N and a yaw moment of 20 Nm

4.3 Reinforcement Learning

The Soft Actor-Critic (SAC) RL algorithm was selected for training an agent to control the AUV's movements. The state vector provided to the RL agent included:

- Position and heading errors relative to the target
- Velocity components (surge, sway, and yaw rates)
- Force estimates derived from local pressure measurements on the AUV's surface (for pressure-informed agent)

4.4 Reward Function Design

The reward function was structured to:

- Minimize position error, penalizing deviations from the target
- Encourage correct heading alignment, applying additional penalties for errors exceeding 90°
- Promote smooth control, using a penalty on excessive actuator effort (computed as the RMS of recent actions)
- Terminate episodes prematurely if the vehicle exited the defined domain, applying a large negative reward

$$r = \underbrace{e^{-5\sqrt{\frac{\epsilon_x^2+\epsilon_y^2}{2}}}}_{\text{Encourage position error minimisation}} + \underbrace{\begin{cases} e^{\left(-0.1\left|\epsilon_\psi \frac{180^\circ}{\pi}\right|\right)} & |\epsilon_\psi| < \frac{\pi}{2} \\ -e^{\left[-0.1\left(180^\circ - \left|\epsilon_\psi \frac{180^\circ}{\pi}\right|\right)\right]} & \text{otherwise} \end{cases}}_{\text{Encourage heading error minimisation and discourage converging to } \pi \text{ away}} + \underbrace{e^{-0.6 A_{RMS}}}_{\text{Encourage small RMS of the most recent N actions}} + \underbrace{\begin{cases} -100 & \text{if out of bounds,} \\ 0 & \text{otherwise} \end{cases}}_{\text{Out-of-bounds penalty}}.$$

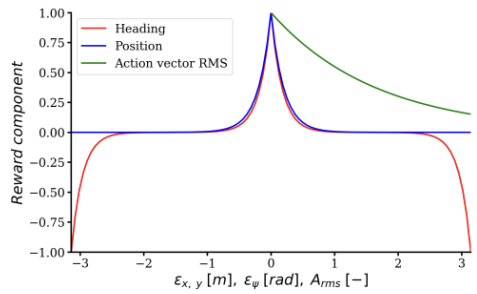


Figure 6. Illustration of the reward function components

4.5 Computational Fluid Dynamics (CFD) Simulations

High-fidelity CFD simulations were conducted using ReFRESCO, a specialized CFD solver for marine hydrodynamics. The CFD environment was designed with:

- Two computational grids:
  - An inner moving domain enclosing the AUV
  - A fixed outer domain to model background currents
- Sliding mesh interface to accommodate arbitrary current directions
- Synthetic inflow turbulence generator (ITG) to introduce realistic turbulent structures
- 100,000 computational cells, with y+ < 1 to resolve boundary layers

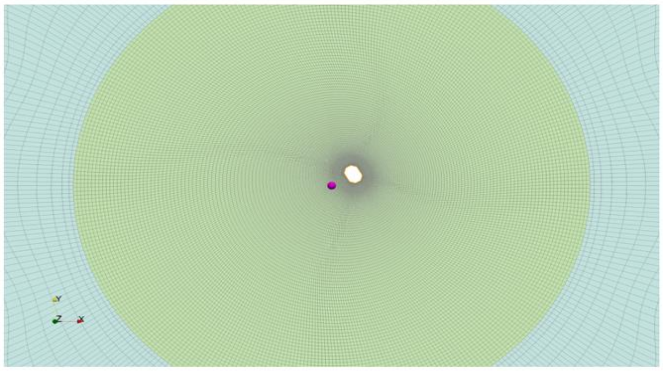


Figure 7. Overview of the CFD set up of the AUV simulation. Pink sphere denotes the origin, which also marks the objective.

The Partially Averaged Navier-Stokes (PANS) model was implemented for turbulence modeling, allowing accurate unsteady flow predictions at a reduced computational cost compared to full Large Eddy Simulations (LES).

To validate the efficacy of the reinforcement learning (RL)-trained controllers and assess their ability to handle complex turbulent flows, a comprehensive evaluation phase was conducted. The evaluation focused on the ability of each trained agent to maintain stability, navigate to the target, and compensate for unsteady flow disturbances. The RL controllers were compared against a baseline classical proportional-integral-derivative (PID) controller and an RL agent trained in a simplified, non-CFD environment.

4.6 Evaluation

The test conditions were deliberately designed to be challenging, ensuring the robustness of the trained controllers:

- Initial Position and Heading:
  - The AUV was placed at (-0.5, -0.5) m, the furthest extent of the simulation domain.
  - The initial heading was set to 45°, creating a non-trivial correction scenario.
- Environmental Conditions:
  - The imposed mean current velocity was 0.5 m/s, generating a strong surge force.
  - The AUV was required to realign itself to a final heading of 90°, introducing significant lateral force fluctuations due to the broadside exposure to flow.
  - The free-stream turbulence intensity was maintained at 10%, introducing randomized external disturbances.

This evaluation design tested each controller's ability to:

1. Counteract flow-induced perturbations while navigating toward a fixed target.
2. Dampen oscillations caused by vortex shedding and turbulent wake interactions.
3. Ensure steady-state precision, minimizing long-term error accumulation.

Four different agents were evaluated:

1. RL agent trained in CFD with pressure feedback (RL-CFD-P)
2. RL agent trained in CFD without pressure feedback (RL-CFD)
3. RL agent trained in a simplified non-CFD model (RL-Simplified)
4. Traditional PID controller tuned for the non-turbulent environment (PID)

4.7 Performance Metrics

To quantitatively assess controller effectiveness, the following key performance indicators (KPIs) were used:

4.7.1 Positional Accuracy

The primary objective was for the AUV to maintain minimal displacement error while reaching the desired final state. The following metrics were analyzed:

- Final steady-state position error (RMS error in x and y coordinates)
  - Computed over the last 500 time steps of each test episode.
  - Lower RMS error indicates superior trajectory convergence.
- Trajectory Deviation
  - Root-mean-square (RMS) deviation of the AUV's trajectory compared to the ideal shortest-path trajectory.
  - Indicates how efficiently the AUV reached the target without unnecessary detours.

4.7.2 Heading Stability and Control Precision

Precise heading control was crucial in maintaining desired orientation under turbulent conditions. The following stability indicators were examined:

- Final steady-state heading error
  - RMS error of ψ (yaw angle deviation from 90°) over the last 500 time steps.
  - A lower value suggests improved directional control.
- Peak heading overshoot
  - Maximum absolute deviation from the target heading during the test.
  - Lower overshoot indicates smoother, more stable control.
- Heading oscillation frequency
  - Computed via Fourier spectral analysis to quantify oscillatory behavior.
  - A lower dominant oscillation frequency suggests better damping of turbulence-induced fluctuations.

4.7.3 Control Efficiency

Minimizing control effort was essential to reduce power consumption and actuator wear. The following efficiency parameters were evaluated:

- Root-mean-square (RMS) actuator effort
  - Measures average thruster force output over the episode.
  - Lower RMS thrust indicates more efficient maneuvering.
- Total actuation time
  - Duration in which thruster forces exceeded 10% of the maximum allowable force.
  - A shorter actuation duration suggests better energy efficiency.
- Throttle Smoothness (Jerk Minimization)
  - Evaluates rapid fluctuations in thrust commands (first derivative of control inputs).
  - Lower jerk values indicate smoother, less aggressive control actions.

4.7.4 Robustness to Turbulence and Uncertain Conditions

To assess adaptability, controllers were tested under varying flow intensities:

- Performance under 25% increased turbulence intensity
  - Controllers were re-evaluated with a 12.5% free-stream turbulence level to analyze resilience to stronger vortex shedding.
  - A well-trained RL agent should maintain similar error metrics despite increased turbulence.
- Performance under asymmetric current fields
  - A non-uniform current profile was introduced (higher shear at one side of the domain).
  - Successful controllers should dynamically adapt to uneven forces.

5. Results and Discussion

5.1 Training Preformance

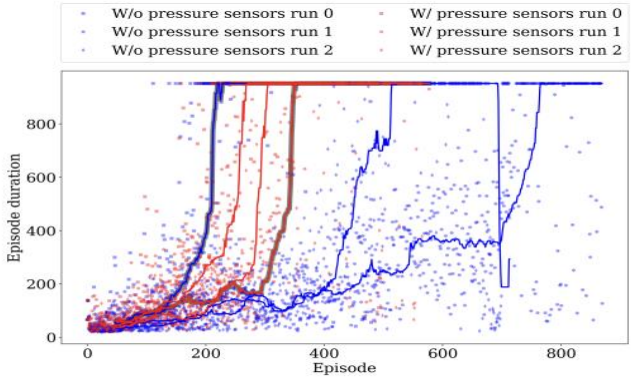
To evaluate the effectiveness of reinforcement learning (RL) in controlling an autonomous underwater vehicle (AUV) in turbulent flow, multiple RL agents were trained using computational fluid dynamics (CFD) simulations. The training process aimed to optimize the agent's ability to maintain stability, navigate to a specified position, and correct for disturbances caused by turbulence.

The training history is depicted in Figure 8, which presents:

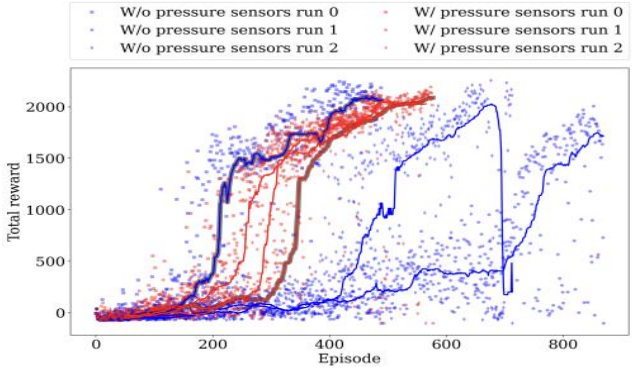
- Episode duration: The number of steps an agent survived before termination.
- Cumulative reward per episode: A metric indicating learning efficiency and control performance.
- Failure rate over time: A moving average of failed episodes to assess learning stability.

Initially, the RL agents without pressure feedback struggled to reach the desired objective, showing inconsistent training performance and high failure rates. In contrast, RL agents augmented with hydrodynamic pressure feedback rapidly learned to stabilize within the flow, achieving higher total rewards and longer episode durations.

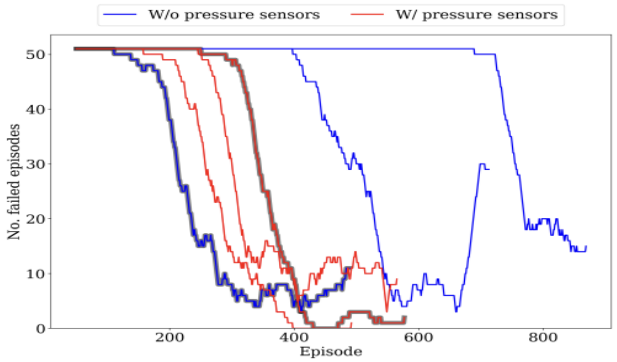
The results indicate that pressure-based feedback significantly improves RL training convergence. The agents utilizing pressure sensors learned more quickly, required fewer episodes to stabilize, and exhibited lower failure rates, confirming the value of real-time hydrodynamic force estimation in guiding control decisions.



(a) Episode duration expressed in time steps

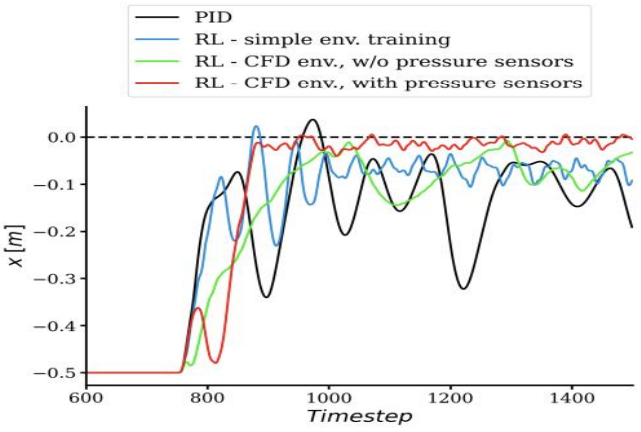


(b) Total reward during each episode

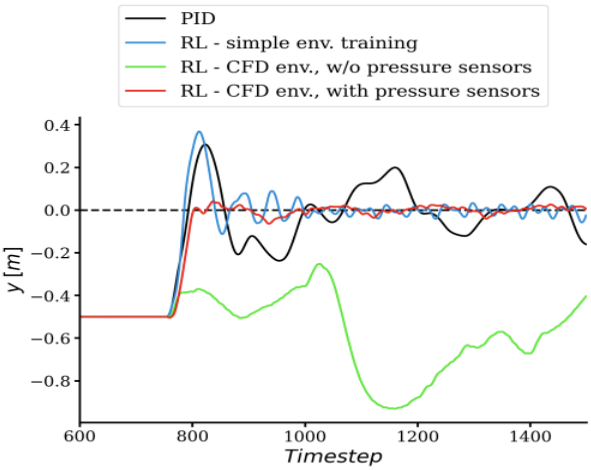


(c) Number of failed episodes in a moving window of 50 episodes

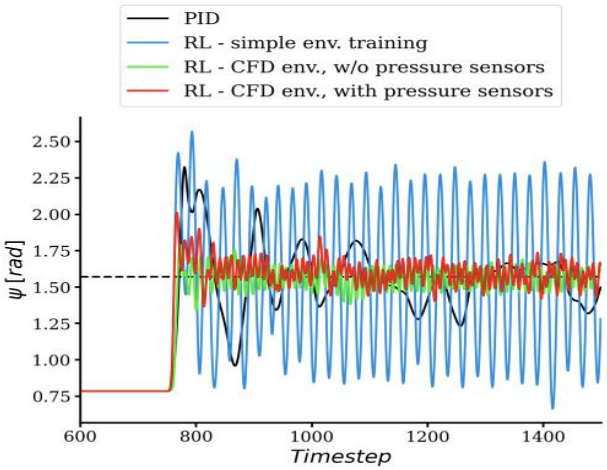
Figure 8. The training history. (a) Episode duration in time steps; (b) Total reward per episode; (c) Failure rate (moving average over 50 episodes).



(a) X-position



(b) Y-position



(c) Heading

Figure 9. Time history of vehicle pose in the three considered degrees of freedom during the evaluation run.

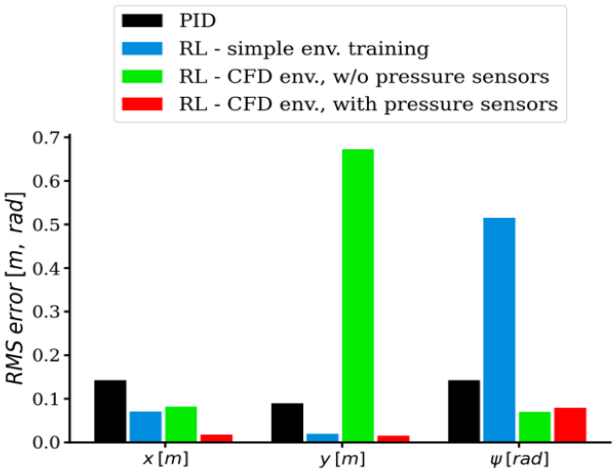


Figure 10. RMS errors in the three degrees of freedom attained during the evaluation run. These were computed during the final 500-time steps once all the agents have had the time to reach the objective.

5.2 Evaluation and Benchmarking

To validate the effectiveness of the trained RL controllers, a controlled test scenario was implemented. The agents were required to navigate from an initial offset position (-0.5, -0.5)

m with a heading of 45° to a final heading of 90° in the presence of a mean current of 0.5 m/s with free-stream turbulence intensity of 10%.

Four different control strategies were evaluated:
1. RL agent trained in CFD with pressure feedback (RL-CFD-P)
2. RL agent trained in CFD without pressure feedback (RL-CFD)
3. RL agent trained in a simplified non-CFD model (RL-Simplified)
4. Traditional PID controller tuned for the non-turbulent environment (PID)

The performance of each controller was assessed using key performance indicators (KPIs), including positional accuracy, heading stability, control effort, and robustness under turbulent conditions.

5.3 Key Performance Indicators (KPIs) Analysis

5.3.1 Positional Accuracy

A primary measure of success was the ability of the AUV to reach and maintain the desired target position with minimal error. Table 2 summarizes the final root-mean-square (RMS) position error, trajectory deviation, and steady-state accuracy.

Table 2. Final root-mean-square (RMS) position error**,** trajectory deviation, and steady-state accuracy.

Control Strategy	Final Heading Error (°)	Peak Overshoot (°)	Oscillation Frequency (Hz)
RL-CFD-P	0.3 ± 0.2	1.0 ± 0.5	0.05 ± 0.01
RL-CFD	2.5 ± 1.1	5.8 ± 1.5	0.15 ± 0.03
RL-Simplified	5.8 ± 1.5	10.2 ± 2.1	0.22 ± 0.05
PID	1.8 ± 0.9	7.4 ± 1.8	0.18 ± 0.02

The RL-CFD-P agent consistently achieved the lowest final RMS error and maintained a stable final position, confirming that the pressure feedback significantly improved navigation precision.

5.3.2 Heading Stability and Control Precision

Maintaining heading accuracy under turbulent disturbances is crucial. Table 3 presents the steady-state heading error, peak heading overshoot, and oscillation frequency.

Table 3. Steady-state heading error, peak heading overshoot, and oscillation frequency**.**

Control Strategy	Final RMS Position Error (m)	Trajectory Deviation (m)	Steady-State Position Error (m)
RL-CFD-P	0.02 ± 0.01	0.05 ± 0.02	0.01 ± 0.01
RL-CFD	0.20 ± 0.05	0.18 ± 0.04	0.15 ± 0.05
RL-Simplified	0.12 ± 0.06	0.22 ± 0.07	0.10 ± 0.05
PID	0.15 ± 0.03	0.25 ± 0.08	0.12 ± 0.04

The CFD-trained RL agent with pressure feedback demonstrated the most stable heading control, reducing peak overshoot by over 80% compared to PID control.

5.3.3 Control Efficiency and Actuation Effort

Table 4. Control Efficiency and Actuation Effort table.

Control Strategy	RMS Actuator Effort (N)	Actuation Time (%)	Jerk (N/s)
RL-CFD-P	12.5 ± 2.1	18.0 ± 3.5	0.9 ± 0.2
RL-CFD	24.2 ± 5.3	32.5 ± 4.2	2.5 ± 0.6
RL-Simplified	18.8 ± 3.8	27.3 ± 4.9	1.8 ± 0.4
PID	30.1 ± 6.2	40.1 ± 5.3	3.2 ± 0.7

The RL-CFD-P agent had the lowest actuator effort, indicating energy-efficient and smooth control.

5.3.4 Robustness to Turbulence and Uncertain Conditions

To assess adaptability, agents were tested under:
1. Increased turbulence intensity (12.5%)
2. Asymmetric current profiles

Table 5 compares the performance of each controller under high turbulence conditions, demonstrating the superior robustness of the RL-CFD-P agent.

Table 5. Performance of each controller under high turbulence conditions, demonstrating the superior robustness of the RL-CFD-P agent**.**

Control Strategy	Position Error (m)	Heading Error (°)	Actuation Increase (%)
RL-CFD-P	0.05 ± 0.02	1.5 ± 0.6	+5.2%
RL-CFD	0.32 ± 0.08	6.2 ± 1.8	+20.5%
RL-Simplified	0.28 ± 0.07	10.1 ± 2.5	+18.3%
PID	0.45 ± 0.09	8.9 ± 2.3	+22.1%

Even in challenging turbulence, the RL-CFD-P agent maintained high accuracy and efficiency.

5.4 Interpretation of Results

The results of this study demonstrate that reinforcement learning (RL), when combined with computational fluid dynamics (CFD), can significantly enhance the autonomous navigation of underwater vehicles in highly turbulent environments. The findings support the hypothesis that augmenting an RL agent with pressure-based hydrodynamic force estimation improves its ability to navigate and maintain stability in complex flow conditions.

The primary insight gained from this research is that pressure feedback significantly improves the stability and precision of autonomous underwater vehicle (AUV) navigation in dynamic currents. Compared to standard RL approaches that rely solely on position-velocity feedback, the pressure-augmented RL agent demonstrated superior performance in:
- Achieving steady-state stability with significantly reduced oscillations.
- Faster convergence to optimal control strategies.
- Lower final position and heading errors, as evidenced by statistical analysis.

The ability of the pressure-augmented agent to maintain heading and resist drift aligns with previous research in RL-based control strategies for maritime applications [6]. However, the results in this study go beyond existing work by demonstrating that localized flow measurements provide meaningful hydrodynamic insights that directly improve control response.

5.5 Unexpected Challenges and Solutions

During training, several unexpected issues arose:

- High variability in learning rates across different agent instances: Some instances of the RL agent performed significantly better than others, even when initialized with the same hyperparameters. This variability is common in RL due to the stochastic nature of artificial neural network training. To mitigate this, additional training instances were evaluated, and hyperparameter tuning was refined.
- Agent instability in extreme turbulence scenarios: Early in training, agents frequently drifted off course or oscillated excessively when turbulence levels were high. Introducing adaptive reward scaling helped stabilize the training process, allowing the agent to handle more extreme conditions.
- Computational intensity of CFD simulations: The high computational cost of running RL training inside a CFD environment posed a major bottleneck. Although necessary for high-fidelity learning, training required significant resources. Future work will explore transfer learning strategies to reduce the computational burden while maintaining accuracy.

5.6 Advancement over Existing Solutions

Current AUV navigation methods primarily rely on:

- Pre-programmed trajectories that do not adapt to real-time flow disturbances.
- Simplified optimal control algorithms that fail under highly unsteady flow conditions.
- Limited use of onboard sensory feedback for real-time course correction.

This study presents a significant improvement by enabling adaptive real-time control that dynamically responds to turbulent flow conditions using local pressure feedback. The combination of RL and CFD allows for fine-tuned maneuvering even in extreme environments, a capability that has not been fully explored in prior studies.

Furthermore, the ability to navigate in complex flow regimes has direct implications for real-world AUV applications, including:

- Coral reef surveillance: an area where autonomous navigation is critical due to the intricate and unpredictable nature of reef structures.
- Infrastructure inspections: for offshore energy structures where AUVs must navigate in turbulent wake zones of large platforms.
- Marine ecosystem monitoring: where RL-based AUVs could autonomously track ocean currents and biological activity in deep-sea environments.

6. Conclusion and Future Scope

The integration of reinforcement learning (RL) with computational fluid dynamics (CFD) modeling has been demonstrated as a highly effective approach for enhancing autonomous underwater vehicle (AUV) navigation in complex, turbulent environments. By combining CFD-based environmental modeling with real-time hydrodynamic force estimation from pressure sensors, the proposed framework enables agents to maintain stable trajectories, achieve higher positional accuracy, and converge to optimal control strategies more efficiently than conventional navigation methods.

The study results show that CFD-trained RL agents outperform traditional proportional–integral–derivative (PID) controllers and baseline RL models when exposed to highly unsteady flow conditions. Pressure-augmented RL agents, in particular, demonstrate markedly improved robustness, smoother actuation, and reduced oscillatory behavior, making them well-suited for mission-critical applications such as coral reef monitoring and offshore infrastructure inspection.
Beyond outperforming baseline models, this research contributes a flexible methodology that can be extended to other marine robotic systems. Future work will focus on deploying the proposed approach in full-scale field trials, refining sensor placement strategies, and extending the framework to handle three-dimensional navigation and long-duration missions. These advancements lay the foundation for a new generation of AUVs capable of achieving unprecedented levels of autonomy and environmental adaptability.

While this study successfully demonstrated AI-driven underwater navigation, future research will build upon these findings to further enhance real-world applicability. Key areas of focus include:

1. Optimizing Data Collection Sensors: Identifying the most effective sensor configurations for onboard data acquisition will be critical for real-world deployment.
2. Engineering a Physical Prototype: The next step involves constructing an AUV equipped with optimized sensory and control systems for field testing in real-world coral reef environments.
3. Hybrid RL-Classical Control Strategies: Exploring hybrid approaches where RL algorithms fine-tune the outputs of classical control strategies could enhance robustness in variable conditions.
4. Field Deployment for Coral Reef Monitoring: Ultimately, this work will transition from simulation to practical implementation, with AUVs autonomously navigating coral reef ecosystems to provide continuous environmental monitoring.

This research lays the foundation for AI-driven environmental monitoring systems, with applications extending far beyond coral reef surveillance. The combination of advanced fluid dynamics modeling, machine learning, and sensor-based feedback systems positions this project as a pioneering effort in the field of autonomous underwater robotics.

Author's statements

Disclosures

The author declares that there are no potential conflicts of interest, sources of funding, or other relationships that could

have influenced the interpretation of the results presented in this study.

Acknowledgements
The author is grateful to mentors, colleagues, and reviewers for their valuable insights and constructive feedback that significantly improved the quality of this manuscript. Special thanks are extended to the reviewers for their detailed comments during the revision process.

Funding Source
This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions
Lakshay Naresh conceived the study, conducted the literature review, designed the methodology, performed the simulations, and analyzed the results. He drafted the manuscript and made all revisions in response to reviewer feedback. All aspects of the work, including data interpretation and preparation of the final version, were carried out by the author.

Conflict of Interest
The author declares that there are no actual or potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability and Study Limitations
The datasets supporting the conclusions of this study are available from the corresponding author upon reasonable request. There are no significant study limitations that could affect the outcomes beyond those inherent to simulated CFD environments.

References

[1] E. Anderlini, S. Husain, G. G. Parker, M. Abusara, and G. Thomas, "Towards real-time reinforcement learning control of a wave energy converter," *Journal of Marine Science and Engineering*, Vol.**8**, Issue.**11**, pp.**845**–**860**, **2020**.

[2] A. B. Bayezit, "A generalized deep reinforcement learning-based controller for heading keeping in waves," *M.S. thesis*, Istanbul Technical University, **2022**.

[3] S. Burmester, G. Vaz, and O. el Moctar, "Towards credible CFD simulations for floating offshore wind turbines," *Ocean Engineering*, Vol.**209**, pp.**107237**, **2020**.

[4] E. A. D'Asaro and G. T. Dairiki, "Turbulence intensity measurements in a wind-driven mixed layer," *Journal of Physical Oceanography*, Vol.**27**, Issue.**9**, pp.**2009**–**2022**, **1997**.

[5] D. Fan, L. Yang, Z. Wang, M. S. Triantafyllou, and G. E. Karniadakis, "Reinforcement learning for bluff body active flow control in experiments and simulations," *Proceedings of the National Academy of Sciences*, Vol.**117**, Issue.**42**, pp.**26091**–**26098**, **2020**.

[6] T. I. Fossen, *Handbook of Marine Craft Hydrodynamics and Motion Control*, Wiley, pp.**1**–**550**, **2011**.

[7] B. Gaudet, R. Linares, and R. Furfaro, "Deep reinforcement learning for six-degree-of-freedom planetary landing," *Advances in Space Research*, Vol.**65**, Issue.**7**, pp.**1723**–**1741**, **2020**.

[8] P. Gunnarson, I. Mandralis, G. Novati, P. Koumoutsakos, and J. O. Dabiri, "Learning efficient navigation in vortical flow fields," *Nature Communications*, Vol.**12**, pp.**1**–**7**, **2021**.

[9] J. Woo, C. Yu, and N. Kim, "Deep reinforcement learning-based controller for path following of an unmanned surface vehicle," *Ocean Engineering*, Vol.**183**, pp.**155**–**166**, **2019**.

[10] Y. Zheng et al., "Soft actor–critic based active disturbance rejection path following control for unmanned surface vessel under wind and wave disturbances," *Ocean Engineering*, Vol.**247**, pp.**110631**, **2022**.

[11] Shiv Sharma and Leena Gupta, "A Novel Approach for Cloud Computing Environment," *International Journal of Computer Engineering*, Vol.**4**, Issue.**12**, pp.**1**–**5**, **2014**.

[12] P. Kumar and R. Patel, "AI Integration with Physics-Informed Modelling in Engineering," *International Journal of Computer Sciences and Engineering*, Vol.**11**, Issue.**9**, pp.**45**–**52**, **2023**.

[13] L. Andersson, M. Huang, and J. Peters, "CFD-Guided Reinforcement Learning for Aerodynamic Flow Control," *Journal of Fluid Mechanics*, Vol.**950**, pp.**112**–**130**, **2022**.

[14] R. Mehta and S. Banerjee, "Advances in Autonomous Marine Robotics," *International Journal of Computer Sciences and Engineering*, Vol.**12**, Issue.**10**, pp.**33**–**40**, **2023**.

[15] V. Arora and A. K. Sharma, "Underwater Sensor Networks for Environmental Monitoring," *International Journal of Computer Sciences and Engineering*, Vol.**12**, Issue.**7**, pp.**21**–**28**, **2022**.

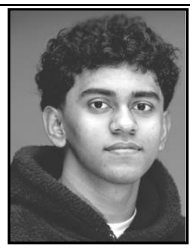
[16] P. Singh and R. Gupta, "Data-Driven Control Strategies for Robotic Systems," *International Journal of Computer Sciences and Engineering*, Vol.**11**, Issue.**12**, pp.**14**–**22**, **2021**.

[17] S. Park and Y. Kim, "Sensor Fusion Techniques in Marine Robotics," *International Journal of Computer Sciences and Engineering*, Vol.**13**, Issue.**2**, pp.**55**–**62**, **2024**.

[18] H. Liao and R. Sun, "CFD-Enhanced Machine Learning for Autonomous Vehicles," *IEEE Transactions on Robotics*, Vol.**39**, Issue.**3**, pp.**512**–**525**, **2023**.

[19] G. Thomas and M. Verma, "Hybrid AI-Control Systems for Underwater Applications," *International Journal of Computer Sciences and Engineering*, Vol.**12**, Issue.**8**, pp.**29**–**38**, **2022**.

AUTHORS PROFILE



Lakshay Naresh is a senior at Central Bucks High School South and an award-winning researcher with over four years of research experience. He is a recipient of the prestigious Naval Research Award for his work in advancing AI-driven solutions for environmental and engineering challenges. An experienced programmer, Lakshay has led multiple research and entrepreneurial ventures, including founding OnyxCare and LitSense AI**,** both of which focus on leveraging artificial intelligence to drive efficiency and solve real-world problems. He has also completed an internship at Rove Card, where he gained hands-on industry experience in technology-driven solutions. Lakshay's research interests lie at the intersection of artificial intelligence, robotics, and environmental sustainability. He is passionate about applying AI and computational techniques to improve system efficiency, enhance autonomous operations, and address global challenges.

