Journal of Theoretical and Applied Information Technology
15th January 2024. Vol.102. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SYSTEM IDENTIFICATION AND FUZZY CONTROLLER DESIGN OF TETHERED UNMANNED UNDERWATER VEHICLES (TUUV) USING DEEP REINFORCEMENT LEARNING CONTROLLER TO AVOID COLLISION

DEVJANI BHATTACHARYA[1], PUTTAMADAPPA C[2]

[1]Designation, Dayananda Sagar University, Department of Electronics and Communication Engineering, Bengaluru, India

[2]Designation. Dayananda Sagar University, Department of Electronics and Communication Engineering, Bengaluru, India

E-mail: [1]devjanibhattacharya.res-soe-ece@dsu.edu.in, [2]puttamadappa@dsu.edu.in

ABSTRACT

In this paper, the main aim is to develop an effective method to avoid collisions underwater. Therefore, a novel method on fuzzy logic controller (FLC) (Fuzzy type 2) method and system identification have been designed with the standard vehicle depth and pitch control dynamics parameters, along with the equations of tethered Unmanned underwater vehicles (TUUV) are elaborately discussed in order to avoid movable obstacles found in underwater. Moreover, the fuzzy controllers help to accumulate appropriate information from the sonar system. Hence utilizing the obtained sonar data or information, the fuzzy controller defines the attack angle and speed using the movement captured through the underwater vehicle. The current research has used Deep Reinforcement Learning Controller for avoiding collision and normally helps to acquire knowledge in accordance with the success rate. The information gathered was utilized to provide an effective outcome in the process. The simulation results of the current study with the appropriate definition of variables and the performance evaluation will be estimated with the obtained output in terms of controlling parameters, including Depth for both positive and Negative, Depth rate, Pitch response, Trajectory, and Tracking for TUUV has been explained. The intelligent framework characteristics of the system identification and fuzzy controller design of TUUV utilizing a novel method - Deep Reinforcement Learning Controller to avoid collision has the capability to procure efficient and better results in avoiding movable obstacles found underwater. Eventually, the simulation results of the proposed novel technique allow the underwater vehicle to securely navigate by avoiding the obstacles in the desired path.

Kkeywords: *Tethered Unmanned Underwater Vehicles, Deep Reinforcement Learning, Fuzzy Logic Controller, Collision Avoidance, Pitch response, Trajectory, and Tracking*

1. INTRODUCTION

Recently, underwater vehicles are gaining huge popularity among the researchers because of their ability to operate in the unstructured oceanic environment [1]. Underwater vehicles help researchers to explore marine life and carry out research investigations with respect to ocean engineering [2]. This research emphasizes on tethered unmanned underwater vehicles.

(TUUV) which belongs to the category of remotely operated vehicles. Autonomous control of TUUVs is a highly complex and challenging task due to the nonlinearities and dynamic characteristics associated with these systems. The variations in the behavior of the UUVs create disturbances in water which leads to volatile hydrodynamic effects. Several researchers have proposed different control approaches to overcome these issues. A H∞ controlling approach was proposed and simulated by [3]. Different sliding mode control (SMC) techniques have been proposed in existing literary works for handling uncertainties [4] [5] [6]. These works have validated the efficacy of SMC based methods in terms of different evaluation metrics such as adaptive control, trajectory tracking, stability analysis, and ability to cope with external disturbances. Advanced and intelligent control techniques based on artificial intelligence and reinforcement learning are implemented in [7] [8] [9]. These techniques enable automated control of UUVs and allow the learning ability of the vehicles based on simple user inputs. They also improve the navigation, tracking, and dynamic positioning of UUVs. Adaptive control methods such as model-based [10], model- reference based method [11], and fuzzy adaptive control methods [12] are used commonly in the controlling of UUVs. The adaptive controllers can effectively handle the dynamics of vehicle parameters and the susceptibility of these parameters to variations. In addition to the existing problems, there are other factors which deteriorate the performance of the controllers. The salinity of underwater affects the buoyancy, the sensors and actuators used for sensing the changes affect the mass parameters, and the presence or obstacles and algae affects the damping parameters. These issues can be handled effectively by the adaptive controllers. However, there are certain limitations which restrict the implementation of the adaptive controllers. These controllers possess highly undesirable frequency characteristics, make contradictory assumptions, and exhibit poor transient behavior [13]. Several researchers have attempted to overcome the drawbacks of these controllers. Most of the research works have focused on increasing the controller gain which results in undesirable outcomes and increase the risk of parameter divergence. Hence it is suggested to limit the controller gain by compromising on adaptation and convergence. The drawbacks of existing controllers motivate this research to address the problem. Since it is essential to avoid collision in underwater vehicles, there is a great need to understand and address this problem.

As inferred from existing literary works, the response of TUUV reduces due to the strong modeling uncertainties and other complexities such as disturbance in the waves or buoyancy changes. The nonlinearity associated with the hydrodynamic elements affect the accuracy of the pitch control techniques. Hence it is essential to design a robust controlling strategy for improving the functioning of the TUUVs. In this work, an advanced controller design is proposed for TUUV using deep reinforcement learning based type 2 FLC. The FLC is simulated by considering the TUUV parameters such as standard vehicle depth and pitch control dynamics. The novel tracking controller proposed in this work creates optimized fuzzy rules automatically using a reinforcement learning (RL) mechanism. In addition, the DRL algorithm acquires the fuzzy rules by estimating the value of each state action pair.

The prominent contributions of this paper are as follows:

- This paper designs and simulates a TUUV model considering the depth and pitch dynamics parameters.
- A novel DRL based controller is designed for avoiding collision and improving the functioning of TUUV.
- The TUUV system is modelled by considering the system parameters such gravity, buoyancy, effects of hydrodynamic forces, and influence of sensors and actuators.
- The pitch response, trajectory and tracking for TUUV is estimated using the proposed pitch controlling strategy based on Type-2 FLC.
- The controller performance is validated by comparing the results with existing controllers.

2. RELATED WORKS

Several controlling strategies are available for achieving autonomy in an unmanned underwater vehicle. Conventional Proportional Integral Derivative (PID) controllers are the frequently employed methods for controlling the location, and orientation of UUVs. These controllers offer robots performance with better simplicity [14] [15]. However, the performance of PID controllers are highly affected when they are implemented for the control of nonlinear plants which are time varying and have significant time delay. These limitations of PID controllers can be resolved by using nonlinear PD or PID controllers i.e., conventional PD and PID controllers are integrated with other nonlinear techniques and learning models [16]. In addition to this, PD/PID controllers are combined with anti-windup design [17] with nonlinear functions. A novel control approach for trajectory tracking in AUVs using nonlinear PID (NLPD) techniques was presented by [18]. This work discusses the advantages of the nonlinear PID strategy compared to traditional PID control approaches under different operating constraints. Results of the experimental analysis shows that the nonlinear PID control techniques exhibit excellent trajectory performance. However, the performance of the NLPD approach deteriorated due to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

uncontrolled variations and external disturbances in the system, observed while determining depth trajectory tracking. In addition, the tracking error in nonlinear techniques is too high and this is mainly due to the buoyancy in the submarines. To overcome this problem, NLPD techniques are combined with integral components which can reduce the tracking error. Correspondingly, a wide range of potentially strong controllers are proposed for resolving the path tracking problem in UUVs. Some of the extensively used techniques are based on FLC [19], sliding mode control (SMC) [20] [21], neural network methods [22] [23], high order SMC (HOSMC) [24], adaptive controllers [25] [26] [27] [28] etc. These techniques are advantageous in terms of providing better path tracking under dynamic conditions. However, they also suffer from certain drawbacks. FLC based techniques are cost effective and simple, but it is difficult to tune the controlling process in underwater systems since they are highly unstable. Besides, FLCs are not suitable for unknown systems i.e., system without any information. Neural networks are considered as a potential alternative to overcome this problem, Neural networks have an ability to learn from previous instances and does not require any information for training the system. But, models based on neural networks have high computation time which is not suitable for real-time applications. SMC based models are robust enough to handle the uncertainties and external disturbances and have finite time convergence. However, the behavior of the SMC technique can have an unfavorable effect on the system which is mainly due to the chattering effect. The chattering effect can be resolved by implementing a hyperbolic tangent function instead of using a signum function [29]. Another way of reducing the chattering effect in SMC is to implement a high order SMC which uses a quasi-continuous control mechanism [30] [31]. For achieving better execution in terms of path tracking and dynamic control, the gain of the controller is adjusted using an adaptive law. It can be summarized from the literature review that nonlinear controllers are more advantageous compared to conventional non-robust techniques. But the design of these controllers involves a lot of complexities and hence the traditional PID controllers and FLCs are integrated with learning based models such as reinforcement learning for improving their performance. The fusion of these techniques is more suitable for estimating the uncertainties of UUV parameters and for achieving better performance under external disturbances.

The drawbacks of existing controllers are as follows: Adaptive controllers might struggle to adapt quickly to rapidly changing or nonlinear dynamics, leading to potential errors in collision avoidance. It is challenging to achieve adaptability since it might affect the stability. In dynamic environments, adaptive controllers prioritize collision avoidance without emphasizing stability. Hence, these controllers fail to achieve a balanced trade-off between collision avoidance and stability" As inferred from existing literary works, the response of TUUV reduces due to the strong modelling uncertainties and other complexities such as disturbance in the waves or buoyancy changes. The nonlinearity associated with the hydrodynamic elements affect the accuracy of the pitch control techniques. Hence it is essential to design a robust controlling strategy for improving the functioning of the TUUVs.

3. DYNAMIC MODELING OF THE TUUV

The TUUV system is modeled by considering the system parameters such gravity, buoyancy, effects of hydrodynamic forces, and influence of sensors and actuators. The dynamic model of the TUUV is expressed using a matrix, using the SNAME notation [32]. The notations and expressions used for modeling is expressed in below equations:

$$\acute{\eta} = J(\eta)\, v \tag{1}$$

$$M\dot{v} + C(v)v + D(v)v + g(\eta) = \tau + \omega_d \tag{2}$$

Where, v is the vector of the matrix which represents the velocity in the fixed frame and is defined as v = [u, v, w, p, q ,r]T , η defines the position and is given as η = [x, y, z, φ, θ, †]T, g denotes the gravitational force, τ is the control input vector and wd is the vector that represent disturbances in the external environment.

A Jacobian transformation matrix denoted as J(η) ∈ R6x6 is used to map the body fixed frame to the earth fixed frame. The matrix consists of three main parameters namely mass (M), Coriolis centripetal (C), and inertia (D). In this work, the TUUV is assumed to have slow dynamics and hence the velocity of the vehicle in terms of Coriolis centripetal becomes negligible (C(v) ≈ 0). The dynamics of the TUUV is described using six degrees of freedom which is formulated by considering 3 different translations and orientations. The input vector denoted as τ ∈ R6

Journal of Theoretical and Applied Information Technology
15th January 2024. Vol.102. No 1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

uses the six actions for controlling the system. These actions are initially represented in the form of body frame which is further remodeled and converted into the earth frame by utilizing the kinematic equations and the model parameters as illustrated in equation 3.

$\acute{\eta} = (5)\, v$

$5^{\cdot} = (5)\, v * + J * (5)\, v$

$M^{*}(5) = J^{-}(5) M J^{-1}(5)$

$D^{*}(v, 5) = J^{-}(5)\, D(v) J^{-1}(5)$

$g^{*}\,(5) = J\text{^}(-T)\,(5) g(5)$

$r^{*} = J^{-}(5)\, r$

$\omega_d{}^{*} = J^{-}(\eta)\omega_d$ (3)
d

Based on the transformation, equation 2 can be transformed into an earth frame and is expressed using equation 4.

$M*(5)5** + D*(v, 5)5* + g*(5) = r* + \omega_d*$ (4)

This work investigates the dynamics of the TUUV and analyses its translational motion along the z axis and its orientation is calculated based on the pitch angle. Hence, the parameters can be transformed as follows: M*(η), D*(η) ∈ R2x2 and g*, τ*, ω* ∈ R2. The term τ* is defined as the control input and is given as;

$\tau^{*} = J^{-T} T\ Ku$ (5)

Where, u ∈ R2 denotes the vector of the control inputs.

4. DESIGN OF TUUV USING PID AND FUZZY TYPE 2 CONTROLLER

The proposed TUUV model is designed by considering the depth and pitch dynamics parameters of the system. The controller in this research combines PID and the type 2 FLC for target tracking and path planning. The motion equation for the TUUV is defined as;

$m\,(\dot{\omega} - u0q) = Z$ (6)

$Iyq = M$ (7)

Where Z is defined as the heave external force, m is the pitch moment consisting of hydrodynamics, q is the pitch.

Assuming that the TUUV is operating in a vertical plane with a constant speed and a smaller pitch angle, the following relation can be derived as shown in equation 8.

$\dot{\theta} = q$ (8)

$\dot{z} = -u_0 \sin\theta + \omega\cos\theta = -\theta\, u_0 + \omega$ (9)

The external forces and momentum acting upon the system is evaluated by determining the relation between different parameters such as linear damping, dynamics of stern plane deflection, and hydrodynamic mass. The effect of moment due to the vertical distance between the centre of gravity and buoyancy is also considered while modeling the system. The equations for modeling the TUUV are given below:

$Z = Z\dot{\omega}\, w + Z\dot{q}\dot{q} + Zww + Zqq + Z\eth\delta s$

$M = M\dot{\omega}\, w + M\dot{q}\dot{q} + M$

$M \approx M\dot{\omega}\, w + M\dot{q}\dot{q} + \dot{M}$

$ww + Mqq - (zG - zB)in\theta + M\eth\delta s\ \omega\omega + Mqq - \overline{W}\overline{BGz}\theta + M\eth\delta s$ (10)

The transfer function for the system considering the depth parameter is given in equation 11.

$$\frac{Z}{\delta s}(s) = \frac{b_1 s^2 + (b_2 a_{12} - b_1 a_{22} - b_2 u_0) + (b_2 u_0 a_{11} - b_1 a_{21} u_0 - b_1 a_{23} + b_2 a_{13})}{[s^3 - (a_{11} + a_{22})^2 + (a_{11}a_{12} - a_{23} - a_{21}a_{12}) + (a_{11}a_{23} - a_{21}a_{13})]} \quad (11)$$

The transfer function for the control design is given as:

$$\frac{\theta(s)}{\eth_s(s)} = \frac{K_\theta}{s^2 + 2\pounds_\theta \omega_\theta s + \omega^2{}_\theta} \quad (12)$$

$$\frac{z(s)}{\eth_s(s)} = -\frac{u_0\ \theta(s)}{s\ \eth_s(s)} \quad (13)$$

Where K_θ is the gain constant of the controller, ω_θ is the natural frequency, and $\pounds_\theta$ is the relative damping ratio.

4.1 PID and Fuzzy Type 2 Controller

Fuzzy logic controllers with a self-tuning PID controller are used for parameter tuning of the TUUV. Fuzzy-based PID controllers increase stability and provide efficient speed control. Figure 1 depicts the schematic of the closed loop controller.

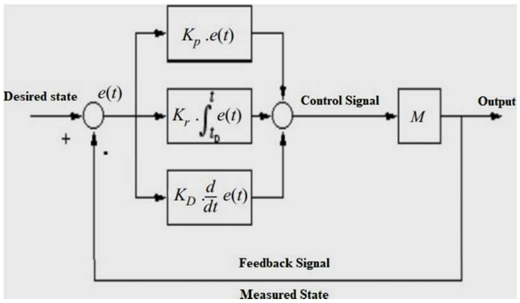

Figure 1 Simulink Model Of The Plant With PID Controller.

The controller is provided with a feedback system, which compares the output with a set point (desired state) and estimates the error e(t). Based on the error generated, the controller adjusts the output generated by the control unit 'M'. The process is repeated until the error becomes zero or negligible. The term 'P' (proportional) in the PID controller is directly proportional to the actual error value i.e., if the error is high then correspondingly, the control output will also be high and vice versa. The response speed of the controller is proportional to the gain factor (K_p) and the response speed can be increased by increasing K_p. However, the value of Kp should not be increased beyond the threshold value since it makes the system unstable. The integral controller 'I' is responsible for reducing the steady state error by integrating the actual error value with respect to time. The integration will reduce the error value to a negligible value or near to zero. Unlike the proportional gain (K_p), the gain of the integral controller (Ki) is inversely proportional to the response speed of the controller. Combining proportional and integral controller (PI controller) will strengthen the steady state response of the system. Derivative controller 'D' is usually implemented as a combined form of PD or PID and is not implemented as a single unit since it might result in zero output for a non-zero error. The controller gain K_d is directly proportional to the error generated.

The transfer function illustrating the gain and time delay of the controller is defined as:

$$C(S) = K_P + \frac{K1}{S} + K_d S = K_P \left[1 + \frac{1}{T_1 s} + T_d s\right] \quad (14)$$

Where, K_P Ki and K_d are defined as the controller gain for the Proportional, Integral and Derivative respectively, and Ti and Td are the integral time and derivative time respectively.

FLC is applied for the UUV, for maintaining constant speed irrespective of uncertainties such as time varying parameters and external disturbances. In this research, a type 2 FLC is used which can modulate the FLC parameters. The inference rule in a fuzzy controller is designed by aggregating the input and output parameters which are tuned using a DRL technique. The fuzzy inference system (FIS) is tested using a trial and error technique to optimize the controlling process. The tuned parameter values are converted into fuzzy values considering the membership functions and the respective values of the parameters. The design parameters of the FLC are selected based on the controller's input and output. Four main parameters such as base, inference engine, fuzzification and defuzzification are used in the design of FLC. The design parameters of the FLC have a direct impact on the fuzzy inference systems. The error (e) is calculated by determining the defuzzified output. If the obtained error is positive then the corresponding output is also positive and vice versa. The membership rule between two variables is shown in equation 15.

$$\mu\_R\ (x, y) = min \quad (15)$$

where, $\mu\_R\ (x,\ y)$ defines the membership rule between the variables.

4.2 Deep Reinforcement Learning

The DRL technique is used for improving the functioning of TUUV and for avoiding collision. The DRL technique employs an efficient Q-learning mechanism which allows the system parameters to make decisions automatically without requiring any previous knowledge of the environment [33]. Reinforcement learning algorithms employ a goal- directed learning approach which extracts necessary information by interacting with the systems unlike other approaches which are trained using only the training data. The dynamic learning behavior of the DRL algorithms not only decides the actions to be taken, but also discovers which actions provide greatest rewards for achieving the goal [34] and validates the action using a trial and error method.

Typically, the problem of RL is computed as a Markov Decision Process (MDP) and is defined as a tuple denoted by (S, A, R, P, γ). The terms S and A are defined as the set of all possible states and available actions for each state respectively, R signifies the reward function, P is defined as the probability of transition, and γ represents the discount factor. The main objective of the agent in the RL interface is to identify a policy π (a|s) which takes a random action (a $\in$ A) for the current state (s $\in$ S) for maximizing the expected (E), and cumulative reward as shown in equation 19:

$$max\ E|_{(T)}| \quad (16)$$

Where, R(T) = $\sum c \quad \gamma tr(at, st)$; $0 \leq \gamma \leq 1$

The primary components of the RL model are policy, reward shaping, value function, and model [35]. The preliminary objective of the DRL algorithm is to obtain an optimal policy π^* that represents the respective control actions (ut) for all states of the system (xt) based on the short and long term rewards.

The configuration of the TUUV model has six degrees of freedom and hence the control system must control the output of 6 thrusters to obtain better dynamic references. In addition, the control system must also handle the nonlinearities associated with 6 DOF (degrees of freedom) in a dynamic and volatile environment [36]. In this research, a deterministic policy based on actor-critic learning is used for approximating the behavior of the model. In this process, the model parameters are updated periodically through which the adaptability is achieved. The policy parameters are updated continuously in every iteration by learning from previous interaction between the TUUV and the external environment.

5. SIMULATION RESULTS

The Simulink model of the controller for TUUV is illustrated in figure 2

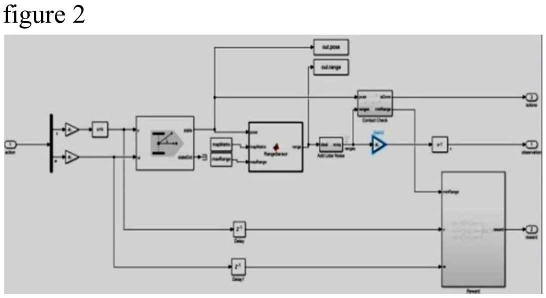

Figure 2 Simulink Model Of The Controller

The proposed controller was simulated using MATLAB, which is a high performance oriented simulation software that uses technical programming language for generating simulink models. The SIMULINK model consists of an integrator, time-delay blocks, an amplifier and output blocks. The simulation model integrates computation, visualization, and programming in a systematic manner and the problems and solutions are expressed using appropriate mathematical expressions. In this research, simulation was performed using Math and computation Algorithm which helps in modelling, simulating and prototyping the data. During the simulation, all blocks are connected in such a way that the data from one block is sent to another block without losing any important information. The signals are generated and are given to the blocks as input and are analysed as different functions. The processed and transferred data is in discrete form since all computerized systems process discrete functions. An appropriate simulation time set up was determined by the fastest dynamics in the simulated system and the maximum and minimum range for the input and output of the model is predetermined. Simulation is performed to validate the working of the controller in terms of different performance parameters which are discussed in the below sections.

5.1 Rotational Velocity and Translational Velocity

The PID controller communicates with TUUV and external environments. The parameters of the PID controllers are tuned using a type 2 FLC and the output graph is illustrated below:

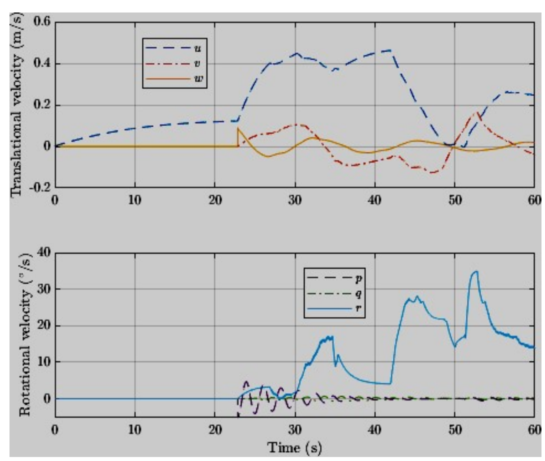

Figure 3 Rotational Velocity And Translational Velocity Of The System

The outcome of the controlling action depicting the potential capability of the model is shown in figure 3. These two parameters are computed for the given waypoints and their variation with the changes in time is analysed. It can be inferred from the results that the control parameters can be obtained by computing the rotational and translational velocity. The coordinates u, v, w define the translational displacement and p, q, r denotes rotational displacement points. These coordinates are evaluated from the center of mass of the TUUV.

5.2 Rotation and Displacement

The rotation and displacement of the vehicle model using type 2 FLC is shown in figure 3

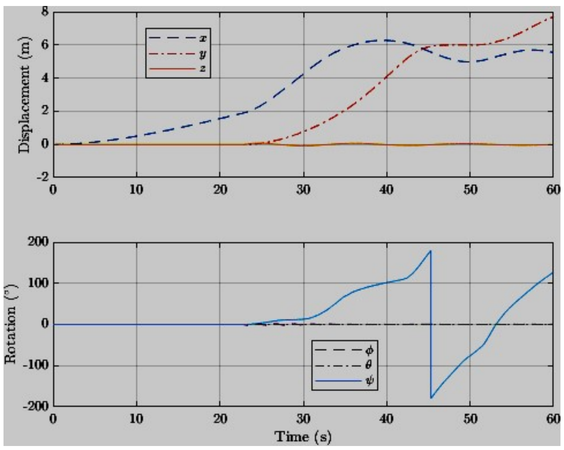

Figure 4 Rotation And Displacement Of The Model Using Type 2 FLC

The outcome of the controlling action depicting the potential capability of the model is shown in figure 4. These two parameters are computed for the given waypoints and their variation with the changes in time is analysed. It can be inferred from the results that the control parameters can be obtained by computing the rotational and translational velocity. The coordinates u, v, w define the translational displacement and p, q, r denotes rotational displacement points. These coordinates are evaluated from the center of mass of the TUUV.

5.3 Rotation and Displacement

The rotation and displacement of the vehicle model using type 2 FLC is shown in figure 5

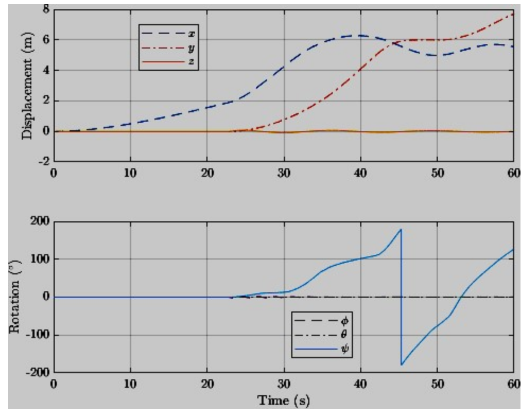

Figure 5 Rotation And Displacement Of The Model

The curves in figure 5 shows that the vehicle model communicates through the controller which is tuned using the type 2 fuzzy controller. The communication is carried out through the set of tracking points.

The vehicle path is determined by the type 2 FLC based PID controller for 3-way points as shown in figure 6

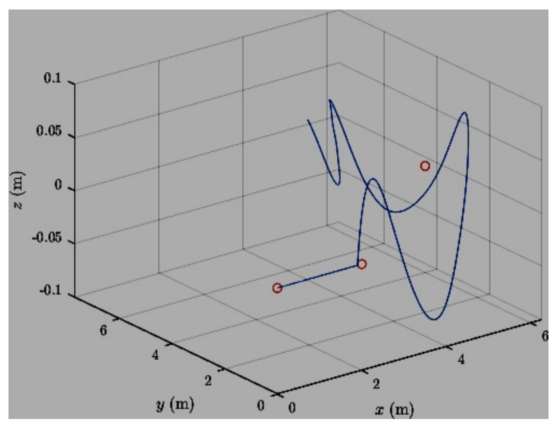

Figure 6 Vehicle Path Determined By The Proposed Controller For 3-Way Point.

It can be inferred from the above figure 5.5 that the 3-waypoints are analyzed for varying time instances and the vehicle moves randomly without colliding with the obstacle. A lag between the movement of the vehicle and the way points can be observed.

5.4 Comparative Analysis

The performance of the proposed approach is validated by comparing the results with other existing approaches such as genetic algorithm based approach, conventional fuzzy logic control (Chen et al., 2018) [37], and Type 2 FLC. The results of the comparative analysis are discussed in the table below.

Table 1. Comparative analysis

Controller Methods	Maximum error (e)	Standard deviation (m)
Genetic Algorithm (GA)	0.63	0.91
Conventional FLC	1.1	0.38
Type 2 FLC	0.24	0.09
Proposed Technique	0.17	0.07

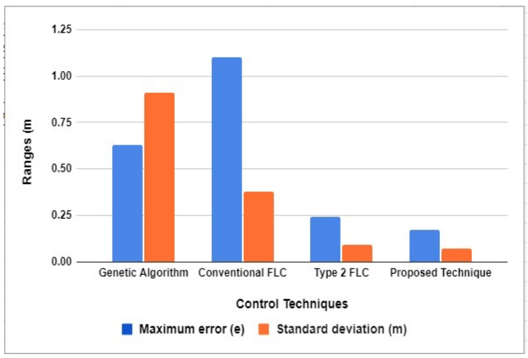

FIGURE 7 COMPARATIVE ANALYSIS

The results of the comparative analysis show that the proposed approach achieves a very minimum error of 0.17 which is lower than other control strategies and the proposed approach provides better error and standard deviation values. This validates the fact the proposed approach is quite effective and robust in controlling the vehicle path under nonlinear and dynamic conditions.

6. CONCLUSION

A comprehensive evaluation of the DRL based fuzzy type 2 controller is discussed for controlling the vehicle path in a tethered unmanned underwater vehicle. The proposed controller strategy is designed incorporating the static and dynamic functionalities of the TUUV. The depth control and pitch parameters were considered while modeling the dynamic system. The parameters of the PID controller were tuned using a type 2 FLC and the functioning of the controller was optimized using a DRL technique. The DRL technique automatically learns the nonlinearities associated with the external environment and decides the corresponding action. The proposed model was simulated using MATLAB and results validated the superior performance of the proposed control strategy. A comparative analysis was conducted wherein the errors and standard deviation values of the proposed approach were compared with existing techniques. The performance of the DRL based type 2 FLC was significantly better than the conventional techniques. For future research, this work can be extended to explore the coordination and communication between multiple TUUVs. This work intends to develop collaborative strategies among several tethered vehicles to collectively avoid collisions while accomplishing tasks.

REFERENCES:

[1] Shi, J. G., Li, D. J., & Yang, C. J, Design and analysis of an underwater inductive coupling power transfer system for autonomous underwater vehicle docking applications. *Journal of Zhejiang University SCIENCE C*, *15*(1), 2014, pp.51-62.

[2] Li, Y., Cui, R., Li, Z., & Xu, D. Neural network approximation based near-optimal motion planning with kinodynamic constraints using RRT. *IEEE Transactions on Industrial Electronics*, 65(11), 2018,pp.8718-8729.

[3] Roche, E., Sename, O., Simon, D., & Varrier, S. A hierarchical varying sampling H∞ control of an auv. IFAC *Proceedings Volumes*, *44*(1), 2011,pp.14729-14734.

[4] Xu, J., Wang, M., & Qiao, L. Dynamical sliding mode control for the trajectory tracking of under actuated unmanned underwater vehicles. *Ocean engineering*, *105*, 2015,pp. 54-63.

[5] Wang, Y., Gu, L., Gao, M., & Zhu, K. Multivariable output feedback adaptive terminal sliding mode control for underwater vehicles. Asian Journal of Control, 18(1), 2016,pp. 247- 265.

[6] Yan, Z., Wang, M., & Xu, J. Robust adaptive sliding mode control of underactuated autonomous underwater vehicles with uncertain dynamics. *OceanEngineering*, *173*, 2019,pp. 802-809.

[7] Cui, R., Yang, C., Li, Y., & Sharma, S. Neural network based reinforcement learning control of autonomous underwater vehicles with control input saturation. *IEEE, UKACC International Conference on Control (CONTROL),* July 2014*,* pp. 50-55.

[8] Yu, R., Shi, Z., Huang, C., Li, T., & Ma, Q. Deep reinforcement learning based optimal

trajectory tracking control of autonomous underwater vehicle. *IEEE 36th Chinese control conference (CCC), July* 2017*,* pp. 4958-4965.

[9] Abrosimov, V., & Panteley, E. Machine Learning for Recognition Learning of Control Systems for Autonomous Unmanned Underwater Vehicles of Events in Hostile Environments. Springer, Cham International Conference on Distributed Computer and Communication Networks, 2021, September, pp. 293-306.

[10] Hassanein, O., Anavatti, S. G., Shim, H., & Ray, T. Model-based adaptive control system for autonomous underwater vehicles. Ocean Engineering, 127, 2016, pp. 58-69.

[11] Makavita, C. D., Nguyen, H. D., Jayasinghe, S. G., & Ranmuthugala, D. (2016, April). Command governor adaptive control for unmanned underwater vehicles with measurement noise and actuator dead-zone. In *2016 Moratuwa Engineering Research Conference (MERCon)* (pp. 379-384). IEEE

[12] Yu, C., Xiang, X., Wilson, P. A., & Zhang, Q. Guidance-error-based robust fuzzy adaptive control for bottom following of a flight-style AUV with saturated actuator dynamics. *IEEE Transactions on Cybernetics*, *50*(5), 2019, pp. 1887-1899.

[13] Maalouf, D., Creuze, V., & Chemori, A. A novel application of multivariable *f* 1 adaptive control: From design to real-time implementation on an underwater vehicle. *IEEE/RSJ international conference on intelligent robots and systems,* October 2012 pp. 76-81.

[14] Maalouf, D., Tamanaja, I., Campos, E., Chemori, A., Creuze, V., Torres, J., & Lozano, R. From pd to nonlinear adaptive depth-control of a tethered autonomous underwater vehicle. *IFAC Proceedings Volumes*, *46*(2), 2013, pp. 743-748.

[15] Salim, M. A., Noordin, A., & Jahari, A. N. A robust of fuzzy logic and proportional derivative control system for monitoring underwater vehicles. IEEE *Second International Conference on Computer Research and Development,* 2010, May*,* pp. 849-853.

[16] Havenstrøm, S. T., Sterud, C., Rasheed, A., & San, O. Proportional integral derivative controller assisted reinforcement learning for path following by autonomous underwater vehicles. *arXiv preprint arXiv:2002.01022*, 2020

[17] Guerrero, J., Torres, J., Creuze, V., & Chemori, A. Trajectory tracking for autonomous underwater vehicle: An adaptive approach. *Ocean Engineering*, *172*, 2019, pp. 511-522.

[18] Campos, E., Chemori, A., Creuze, V., Torres, J., & Lozano, R. Saturation based nonlinear depth and yaw control of underwater vehicles with stability analysis and real-time experiments. *Mechatronics*, *45*, pp. 49-59.

[19] Nguyen, D. A., Tien, N. T., & Anh, P. V. Fuzzy controller design for autonomous underwater vehicles path tracking. *IEEE International Conference on System Science and Engineering (ICSSE),* July 2019*,* pp. 592-598.

[20] González-García, J., Narcizo-Nuci, N. A., García-Valdovinos, L. G., Salgado-Jiménez, T., Gómez-Espinosa, A., Cuan-Urquizo, E., & Cabello, J. A. E. Model-free highorder sliding mode control with finite-time tracking for unmanned underwater vehicles. *Applied Sciences*, *11*(4), 1836, 2021

[21] Lin, H., Chen, K., & Lin, R. Finite-time formation control of unmanned vehicles using nonlinear sliding mode control with disturbances. *Int. J. Innov. Comput., Inf. Control*, *15*(6), 2019, pp.2341-2353.

[22] Shojaei, K. Three-dimensional neural network tracking control of a moving target by underactuated autonomous underwater vehicles. *Neural Computing and Applications*, *31*(2), 2017, pp. 509-521

[23] Abrosimov, V., & Panteley, E. Machine Learning for Recognition Learning of Control Systems for Autonomous Unmanned Underwater Vehicles of Events in Hostile Environments. In *International Conference on Distributed Computer and Communication Networks,* 2021, September pp. 293-306. Springer, Cham.

[24] Khalid, M. U., Ahsan, M., Kamal, O., & Najeeb, U.Modeling and trajectory tracking of remotely operated underwater vehicle using higher order sliding mode control.*16th IEEE International Bhurban conference on applied sciences and technology (IBCAST),* 2019, January pp. 855-860..

[25] Yu, C., Xiang, X., Zhang, Q., & Xu, G. Adaptive fuzzy trajectory tracking control of an under-actuated autonomous underwater vehicle subject to actuator saturation. *International Journal of Fuzzy Systems*, *20*(1), 2018, pp.269-279

[26] Wang, N., Su, S. F., Yin, J., Zheng, Z., & Er, M. J. Global asymptotic model-free trajectory-independent tracking control of an uncertain marine vehicle: An adaptive universe-based fuzzy control approach. *IEEE Transactions on Fuzzy Systems*, *26*(3), 2017 pp.1613-1625.

[27] Liang, H., Fu, Y., Gao, J., & Cao, H. Finite-time velocity-observed basedadaptive output-feedback trajectory tracking formation control for underactuated unmanned underwater vehicles with prescribed transient performance. *Ocean Engineering*, *233*, 109071, 2021.

[28] Huang, B., Zhou, B., Zhang, S., & Zhu, C. Adaptive prescribed performance tracking control for underactuated autonomous underwater vehicles with input quantization. *Ocean Engineering*, *221*, 108549., 2021

[29] Kim, J., Joe, H., Yu, S. C., Lee, J. S., & Kim, M. Time-delay controller design for position control of autonomous underwater vehicle under disturbances. *IEEE Transactions on Industrial Electronics*, *63*(2), 2015, pp. 1052-1061.

[30] Ismail, Z. H., & Putranti, V. W. Second order sliding mode control scheme for an autonomous underwater vehicle with dynamic region concept. *Mathematical Problems in Engineering*, *2015*.

[31] Guerrero, J., Torres, J., Antonio, E., & Campos, E. Autonomous underwater vehicle robust path tracking: Generalized super-twisting algorithm and block backstepping controllers. *Journal of Control Engineering and Applied Informatics*, *20*(2), 2018, pp.51-63.

[32] Fossen, T. I. Marine control systems–guidance. navigation, and control of ships, rigs and underwater vehicles. *Marine Cybernetics, Trondheim, Norway, Org. Number NO 985 195 005 MVA, www. marinecybernetics. com, ISBN: 82 92356 00 2*, 2002

[33] Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, *34*(6), 2017, pp.26-38.

[34] Mohammed, M. Q., Chung, K. L., & Chyi, C. S. Review of deep reinforcement learning-based object grasping: Techniques, open challenges, and recommendations. *IEEE Access*, *8*, 2020, pp. 178450-178481.

[35] Ding, Z., Huang, Y., Yuan, H., & Dong, H. Introduction to reinforcement learning. In *Deep reinforcement learning Springer, Singapore*,2020, pp. 47-123)..

[36] Carlucho, I., De Paula, M., Wang, S., Petillot, Y., & Acosta, G. G. Adaptive low- level control of autonomous underwater vehicles using deep reinforcement learning. *Robotics and Autonomous Systems*, *107*, 2018, pp.71-86.

[37] Chen, J., Zhu, H., Zhang, L., & Sun, Y. Research on fuzzy control of path tracking for underwater vehicle based on genetic algorithm optimization. *Ocean Engineering*, *156*, 2018, pp.217-223.