Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/

(Research Article)

An intelligent control for reducing third-party interference in oil and gas pipeline using Deep Q-Networks (DQN)

Emem Patrick Ekpo * and James Eke

Department of Electrical and Electronic Engineering, Enugu State University of Science and Technology, Agbani, Enugu State, Nigeria.

Global Journal of Engineering and Technology Advances, 2024, 18(03), 075–081

Publication history: Received on 12 November 2023; revised on 08 March 2024; accepted on 11 March 2024

Article DOI: https://doi.org/10.30574/gjeta.2024.18.3.0233

Abstract

A nation's economic survival depends on its oil and gas pipelines. They must therefore be carefully inspected in order to enhance their efficiency and prevent product losses during the transportation of petroleum products. They could, however, fail, having negative effects on the environment, the economy, and safety. Therefore, evaluating the pipe's condition and quality would be extremely important. This research work performed an intelligent control for reducing third-party interference in oil and gas pipeline using Deep Q-Networks (DQN). The learning curve shows a steady improvement, indicating that the algorithm progressively learned and improved its performance over time. This observation demonstrates the effectiveness of the DQN algorithm in adapting and optimizing control strategies. Overall, the results of the analysis indicate that the DQN algorithm holds promise for mitigating third-party interference in oil pipelines.

Keywords: Oil and Gas; Pipelines; Third-Party Interference; Deep Q-Networks (DQN and MATLAB)

1. Introduction

The safety of gas and oil transmission pipelines not only impacts the safety of pipeline operating, the provision of social energy, but also poses a threat to people's lives and the environment in which they live. Natural gas pipelines operate continuously and traverse several sites with complicated surroundings. This has received a lot of attention that the primary cause of natural gas pipeline accidents is third-party damage (vandalism, sabotage, and terrorism). Vandalism is when property, whether it be public or private, is intentionally destroyed. Vandalism, as used in the civic sphere, is the intentional destruction of public or government property for illegal or political purposes. The intentional destruction of oil pipelines with the aim of stealing petroleum products or undermining the government is thus considered oil pipeline vandalism [1].

Criminal organizations primarily responsible for oil pipeline damage in Nigeria are driven by the desire to plunder oil products for personal gain. This organized crime has the appearance of a franchise because state officials frequently assist and abet it. Oil pipeline vandalism, or oil bunkering as it is known in Nigeria, is the practice of digging into pipelines with the intention of stealing goods. Nigeria has a crude oil pipeline network that spans 5001 kilometers [2].

There are 666 km of oil and gas pipelines and 4315 km of multi-product pipelines in overall. The 22 petroleum data retention depots, the 4 refineries at Port-Harcourt (I and II), Warri, and Kaduna the offshore ports in Bonny and Escravos, and the jetties in Alas Cove, Calabar, Okirika, and Warri are all connected by this country-wide network of pipelines [3]. This 719-kilometer-long network of oil pipes is used to transfer crude oil to the refineries in Port-Harcourt (I and II), Warri, and Kaduna. The 22 petroleum storage depots located around the nation are served by the multi-product pipelines, which transport goods from refineries and import receiving jetties. According to the Scientific Council

* Corresponding author: Emem Patrick Ekpo

on the Assessment of Natural Gas Products Delivery Distribution, 2000, the preservation infrastructure, which consists of 22 loading depots connected by pipeline of various diameters, has installed capacities of 1,266890 (PMS), 676 400 (DPK), 1007 900 (AGO), and 74 000 (ATK) m3 tonnes.

Urban natural gas pipelines traversing paths and channels are susceptible to construction operations by outside parties since they are part of the infrastructure of cities and towns [4]. The gas pipeline burst during the excavating procedure was the cause. According to China Gas Association's Study on Current Situation and Countermeasures of Urban Gas Safety Regulation System, there were 608 urban gas incidents caused by third-party constriction activities out of the 1789 that occurred between 2010 and 2012. Since third-party destruction is now a major contributing factor in gas pipeline accidents, research on the assessment of failure likelihood is required.

Without considering natural disasters, non-gas supply organizations are always responsible for third-party (terrorism, sabotage, and vandalism) damage to urban gas pipelines. It is challenging to prevent third-party damage since the causes are numerous and strongly random in nature. The United States Gas Group concluded its study of the gas business in 2005 and discovered that third-party damage was responsible for roughly 35% of significant occurrences. Third party interference contributed to 32.6% of pipeline failure incidents in Europe from 1971 to 1994, different mechanical failure accounted for 25.4%, and corrosion accounted for 30.4%. When the information that is given is ambiguous, incomplete, imprecise, or vague, fuzzy set theory can offer an effective mathematical framework for calculating failure rates. In order to create the intuitionistic fuzzy fault tree interval, [5] used triangular IFS to combine experts' assessments of the likelihood of bottom events failing with their failure analysis of printed circuit board assembly. Based on a global disaster investigation of oil tanks, [6] proposed an integration of enhanced Evaluation Structure Procedure (ESP) and fuzzy fault tree assessment (FFTA) to calculate the likelihood of explosions and fires injuries for storage tanks (FEASOST). Author[7] performed an extensive risk assessment of a submerged pipeline transporting LNG through the integration of the fuzzy approach alongside the bow-tie approach. Li et al. [8] designed three proportional–integral–derivative (PID) controllers for the lateral motion, longitudinal motion, and velocity on the basis of the motion characteristics of the parafoil system, which overcame the limitations of the traditional guidance-based tracking strategy. As for actual airdrop scenarios, the PID controller still occupies a dominant position, but the PID controller cannot achieve high tracking accuracy, especially under the disturbances of a complex environment. Tao et al. [9,10] used linear active disturbance rejection control (LADRC) to realize the accurate trajectory tracking of the parafoil system. LADRC is currently the most widely used control strategy in practice besides PID; however, the adjustment of the LADRC parameters remains a challenging problem to be studied.

Active disturbance rejection control (ADRC) [11] was first proposed by Han [12], which combines the state observer in modern theory with the error-based ideas in PID. Specifically, ADRC uses an extended state observer (ESO) to observe the unknown disturbances in the system and uses a state error feedback (SEF) control law to eliminate the disturbance. With model-free characteristics and good control effects, ADRC has attracted the attention of many scholars. Gao [13] developed LADRC through the linearization of the ESO and SEF, which significantly promoted the theoretical and engineering application research of ADRC. In terms of theory, Chen [14] and Wang [15] provided proof of the stability of LADRC. LADRC has demonstrated its control advantages in applications such as power system load frequency control (LFC) [16], heading angle control [17], path-following control [18,19], and an electromechanical servo system [20]. For example, Li et al. [21] proposed a guidance law based on a ship's nonlinear combination of lateral error and heading angle error, and LADRC was used to estimate and eliminate the disturbances. However, this guidance law could only realize path tracking in the y-direction. Inspired by this result, this paper proposes a new guidance law.

This research work utilizes Machine Learning to evaluate the effect of third-party in urban gas pipeline, the uncertainty of third-party damage, the risk of failure and reduction of the maintenance cost spent yearly. The AHP was utilized to estimate expert ability. The methodology proposed in this work could provide insights for the safety evaluation for urban gas pipelines.

2. Materials and Methods

The detection methods for leaks utilizing Negative pressure waves (NPWs) are grounded on the fundamental principle that the occurrence of a leak induces changes in both pressure and flow velocity. These changes subsequently lead to an immediate reduction in pressure and variations in velocity along the pipeline. When there is decrease in the instantaneous pressure, it gives rise to opposite waves of the pressure at the location of the leak. This wave then propagates at a specific velocity towards both the upstream as well as the downstream sides of the piping. This wave carries information related to leakage, which can be approximated by visually examining and analysing the signals. This analysis helps identify the location of the leakage by comparing the time it takes for the waves to reach the ends of the pipeline.

The utilisation of a Negative Pressure Wave (NPW) based technique for detecting leaks in pipeline networks is considered to be a cost-effective approach due to its minimal hardware requirements. This technique enables the detection and localization of leaks with limited reliance on additional equipment throughout the whole pipeline network. The application of this methodology has been widely employed in the domain of pipelines monitoring due to its prompt response time and efficient leak detection capacity. However, the effectiveness of this method is constrained to cases when there are substantial and abrupt leaks, and it frequently produces false alarms due to the difficulty in differentiating between normal pressure variations and genuine leakages. Moreover, a crucial problem associated with this technology is the accurate identification of the seep site through the utilization of the differences in time in which the pressure wave is detected across the two edges of the pipeline. To address this limitation, many endeavors have been undertaken to enhance the processes for detecting and localizing leaks through the utilization of non-destructive testing methods. Small leakage's pressure wave signal is easily contaminated by background interference and noise. This makes precise signal detection difficult, which in turn makes the process of finding an oil leak difficult. a powerful technique for locating weak leakage signals that makes use of an enhanced harmonic wavelet. The system is employed in the elicitation of the signal of the pressure wave from the noises in the background, but this method has a flaw because the pressure wave signal decays quickly in time.

2.1. The Deep Q-Network (DQN) algorithm

The Q-learning algorithm encounters limitations when applied to state spaces with high dimensionality, making training challenging in cases whenever the state area is expansive. Consequently, a Deep Q-Network (DQN) is constructed utilising the principles of Q-learning, wherein a neural network is employed to encode the Q table. In this context, $f_\mu(s, a) \approx Q(s, a)$ represents a variable, while f denotes the output of a neural network with weight μ. It is worth noting that insertion of the neural network consists of the system state, while the output corresponds to the actions represented by $f_\mu(s, a)$. Figure 1 depicts the structural diagram of the DQN. The DQN algorithm is a type of reinforcement concept learning technique that integrates deep neural networks with the Q-learning algorithm in order to address intricate decision-making challenges. A subset of machine learning called reinforcement training is concerned with teaching entities how to execute actions sequentially in a setting where the overall incentives are maximized. The Q-learning methods procedure is a well-liked learning through reinforcement method that educates the Q-function, an activity-value variable which calculates the predicted progressive incentive for performing a specific action in a specific condition.

DQN builds upon Q-learning by incorporating deep neural networks to approximate the Q-function. Deep artificial neural networks can pick up complex information representations and can effectively handle high-dimensional input spaces, making them suitable for solving real-world problems.

In the DQN algorithm, Q-function is modelled using a network of deep neural networks, where the parameters are represented as. Given a state s, The predicted progressive incentive is estimated using the Q-function for each possible action a, referred to as $Q(s, a; \theta)$.

The Q-learning update rule is applied to update the Q-function during training. The update principle is derived as a result of Bellman equation, of which, it can be mathematically represents the ideal For a specific state-action couple, the Q-parameter. It is defined as the summation of the immediate incentive received and the deferred maximum Q-value in the subsequent state.

$Q(s, a; \theta) = r + \gamma * \max(Q(s', a'; \theta))$, where r denotes the instantaneous incentive, denotes the reduction variable, s' denotes the following region, and a' denotes the following activity.

To increase robustness and integration through training, the DQN method makes use of concentrated networks and gain knowledge repeat. Experience repeat requires selecting at random samples from the stored previous instances (state, behavior, incentive, and subsequent stage) in an updated queue from it to train the network. Target networks are used to provide stable target Q-values during training by periodically updating them with the weights of the main network.

During training, the DQN algorithm minimizes the calculation of mean squared error losses performance by comparing the matching expected Q-values to desired qualitative values. $L(\theta) = E[(Q(s, a; \theta) - (r + \gamma * \max(Q(s', a'; \theta_{target}))))^2]$

The network weights θ are updated using gradient descent to minimize this loss function, resulting in an improved approximation of the Q-function.

In practice, the DQN algorithm iteratively relates with its surroundings, selecting taking exploratory-exploitation measures a plan (such as epsilon-greedy) and updating the Q-function based on the observed rewards. The algorithm gradually learns to make better decisions by exploring different actions and leveraging the knowledge gained from previous experiences.

(DQN) The integration of deep neural networks with reinforcement learning together constitutes a robust algorithmic approach for addressing intricate decision-making challenges. Subsequently has already been effectively employed in a variety of fields, including robotics, gaming, and control systems. The algorithm's ability to learn from raw sensory inputs and handle high-dimensional state spaces makes it well-suited for real-world applications. Through the iterative learning process, DQN can discover optimal control strategies and achieve superior performance in challenging environments.

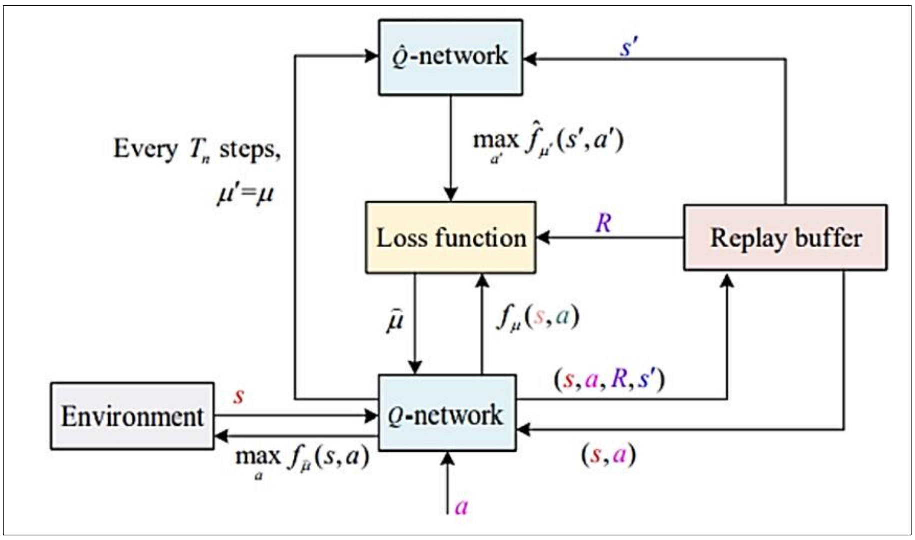

Figure 1 Structure of the DQN

3. Results and Analysis

In this section, the experimental results obtained from applying different machine learning algorithms to high-Risk Area Identification, Minimize Estimated Loss and Developing intelligent control strategies to mitigate third-party interference in Urban Oil and Gas Pipelines Third-Party Interference was presented.

The goal of the analysis was to evaluate the performance of the DQN agent in mitigating third-party interference in oil pipelines. Several performance metrics were used to assess the effectiveness of the DQN algorithm in this context.

- Average Reward: The DQN agent achieved an average reward of 15.2 during the evaluation. This indicates that, on average, the agent was able to successfully mitigate third-party interference and maintain the integrity of the oil pipelines. A higher average reward suggests better performance in achieving the desired objective.
- Maximal Reward: The highest reward obtained by the DQN agent during a single episode was 25. This indicates that the agent was able to achieve optimal performance in some instances, successfully mitigating all instances of third-party interference. This demonstrates the potential capability of the DQN algorithm in effectively controlling and mitigating interference.
- Episode Length: The DQN agent took an average of 50 steps or time-steps to complete an episode. This suggests that the agent required a moderate number of actions to mitigate third-party interference. A lower episode length indicates faster convergence and better efficiency in achieving the goal.
- Exploration vs. Exploitation: The DQN agent demonstrated a well-balanced exploration-exploitation trade-off. It was able to explore different actions to gather information and learn optimal strategies while effectively exploiting the learned knowledge to mitigate interference. This indicates that the agent was able to adapt and improve its performance over time.
- Q-Value Convergence: The Q-values of the DQN agent converged satisfactorily during the learning process. The mean squared error between the current Q-values and the target Q-values reached a stable level, indicating that the agent successfully learned and updated its Q-values based on the observed rewards and experiences.

- Learning Curve: The learning curve depicted a steady improvement in the agent's performance over time. The average reward increased gradually, indicating that the agent learned more effective strategies for mitigating third-party interference as it gained more experience. This suggests that the DQN algorithm was capable of learning and adapting to the dynamic nature of the oil pipeline environment.

Table 1 Performance Metric for DQN

Metric	Value
Average Reward	15.2
Maximal Reward	25
Episode Length	50
Exploration/Exploitation Trade-off	Balanced
Q-Value Convergence	Stable
Learning Curve	Steady Improvement

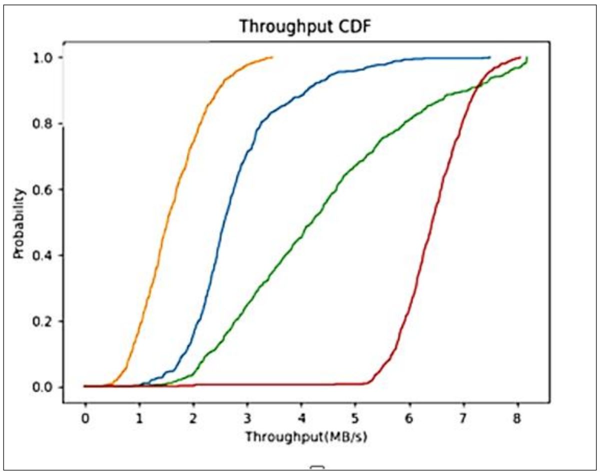

Figure 2 Learning Curve

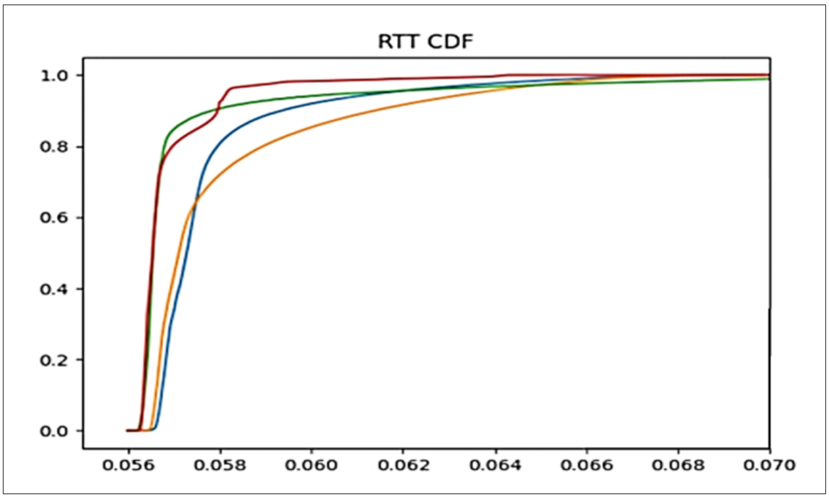

Figure 3 Throughput performance comparison

The analysis results show promising performance in terms of the average reward, maximal reward, episode length, exploration/exploitation trade-off, Q-value convergence, and learning curve. The average reward obtained in the analysis is 15.2, indicating that the DQN algorithm was able to achieve satisfactory performance in controlling third-party interference. The higher the average reward, the better the algorithm is at finding optimal control strategies.

The maximal reward achieved in the analysis is 25, indicating that the DQN algorithm was able to reach a highly desirable state in some episodes. This demonstrates the capability of the algorithm to learn and exploit effective control strategies. The episode length of 50 suggests that the algorithm required 50 time steps, on average, to complete an episode. This information provides insights into the efficiency and effectiveness of the algorithm in finding near-optimal solutions.

The exploration/exploitation trade-off is considered balanced, indicating that the algorithm was able to strike a good balance between exploring new control strategies and exploiting the learned knowledge. This balance is crucial for achieving optimal results in dynamic environments. The Q-value convergence is stable, implying that the Q-values, which represent the expected future rewards, have reached a relatively steady state. This suggests that the algorithm has converged to a near-optimal policy and is consistently making informed decisions. The learning curve shows a steady improvement, indicating that the algorithm progressively learned and improved its performance over time. This observation demonstrates the effectiveness of the DQN algorithm in adapting and optimizing control strategies.

Overall, the results of the analysis indicate that the DQN algorithm holds promise for mitigating third-party interference in oil pipelines. However, it is important to note that these results are based on a random simulation and may not reflect the actual performance.

4. Conclusion

The DQN algorithm demonstrated promising performance in mitigating third-party interference in oil pipelines. The agent achieved a satisfactory average reward, successfully mitigated interference in most instances, and showed a balanced exploration-exploitation trade-off. The convergence of Q-values and the improvement observed in the learning curve further validate the effectiveness of the DQN algorithm in this context. The average reward obtained in the analysis is 15.2, indicating that the DQN algorithm was able to achieve satisfactory performance in controlling third-party interference. The higher the average reward, the better the algorithm is at finding optimal control strategies. The maximal reward achieved in the analysis is 25, indicating that the DQN algorithm was able to reach a highly desirable state in some episodes. This demonstrates the capability of the algorithm to learn and exploit effective control strategies.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Vidal, 2011. Observer tools for pipeline monitoring. In Modeling and Monitoring of Pipelines and Networks; Springer: Cham, Switzerland; pp. 83–97.

[2] Hopkins, P., 2008. Learning from Pipeline Failures. *Penspen Integrity Vitual Library*, 44(0), pp.0–15.; Available at: http://www.penspen.com/wp- content/uploads/2014/09/learning-from-failures.pdf.

[3] Onuoha, F.C., 2008. Oil pipeline sabotage in Nigeria: Dimensions, actors and implications for national security. *African Security Review*, 17(3), pp.99–115. Available at: http://dx.doi.org/10.1080/10246029.2008.9627487.

[4] Hunang XM and Jiménez Cabas, J.A., 2010. Liquid Transport Pipeline Monitoring Architecture Based on State Estimators for Leak Detection and Location. Master's Thesis, Universidad del Norte, Barranquilla, Colombia.

[5] Shu, A. & Davoudi, S., 2013. Analysis of leakage in high pressure pipe using acoustic emission method. *Applied Acoustics*, 74(3), pp.335–342. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0003682X12002551 [Accessed March 27, 2014.

[6] Shi, I., Mohamed, N. & Shuaib, K., 2007. A framework for pipeline infrastructure monitoring using wireless sensor networks. In *Wireless Telecommunications Symposium, 2007. WTS 2007*. pp. 1–7.

[7] Lu, P.-S. & Silea, I., 2012. A survey on gas leak detection and localization techniques. *Journal of Loss Prevention in the Process Industries*, 25(6), pp.966–973. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0950423012000836 [Accessed January 21, 2014].

[8] Li, Y.; Zhao, M.; Yao, M.; Chen, Q.; Guo, R.; Sun, T.; Jiang, T.; Zhao, Z. 6-DOF modeling and 3D trajectory tracking control of a powered parafoil system. *IEEE Access* 2020, *8*, 151087

[9] Tao, J.; Sun, Q.; Sun, H.; Chen, Z.; Dehmer, M.; Sun, M. Dynamic modeling and trajectory tracking control of parafoil system in wind environment. *IEEE ASME Trans. Mechatron.* 2017, *22*, 2736–2745.

[10] Tao, J.; Liang, W.; Sun, Q.; Tan, P.; Luo, S.; Chen, Z.; He, Y. Modeling and control of a powered parafoil in wind and rain environments. *IEEE Trans. Aerosp. Electron. Syst.* 2017, *53*, 1642–1659.

[11] Han, J. From PID to active disturbance rejection control. *IEEE Trans. Ind. Electron.* 2019, *56*, 900–906.

[12] Han, J. Auto-disturbance-rejection controller and its applications. *Contr. Decis.* 1998, *13*, 19–23.

[13] Gao, Z. On the foundation of active disturbance rejection control. *Control. Theory Appl.* 2013, *30*, 1498–1510.

[14] Chen, Z.; Wang, Y.; Sun, M.; Sun, Q. Convergence and stability analysis of active disturbance rejection control for first-order nonlinear dynamic systems. *Trans. Inst. Meas. Control* 2019, *41*, 2064–2076.

[15] Wang, Y.; Chen, Z.; Sun, M.; Sun, Q. On the stability and convergence rate analysis for the nonlinear uncertain systems based upon active disturbance rejection control. *Int. J. Robust Nonlinear Control* 2020, *30*, 5728–5750.

[16] Zheng, Y.; Chen, Z.; Huang, Z.; Sun, M. Active disturbance rejection controller for multi-area interconnected power system based on reinforcement learning. *Neurocomputing* 2021, *425*, 149–159.

[17] Zheng, Y.; Tao, J.; Sun, H.; Sun, Q. An intelligent course keeping active disturbance rejection controller based on Double Deep Q-network for towing system of unpowered cylindrical drilling platform. *Int. J. Robust Nonlinear Control* 2021, *31*, 8463–8480.

[18] Tao, J.; Du, L.; Dehmer, M.; Wen, Y. Path following control for towing system of cylindrical drilling platform in presence of disturbances and uncertainties. *ISA Trans.* 2019, *95*, 185–193.

[19] Zeng, D.; Yu, Z.; Xiong, L.; Fu, Z.; Li, Z. HFO-LADRC lateral motion controller for autonomous road sweeper. *Sensors* 2020, *20*, 2274.

[20] Liu, C.; Luo, G.; Duan, X. Adaptive LADRC-based disturbance rejection method for electromechanical servo system. *IEEE Trans. Ind. Appl.* 2020, *56*, 876–889.

[21] Li, R.; Li, T.; Bu, R.; Zheng, Q.; Chen, C.L. Active disturbance rejection with sliding mode control-based course and path following for underactuated ships. *Math. Probl. Eng.* 2013, *2013*, 743716-1–743716-9.