

Adaptive Vibration Control of Smart Structure Using Deep Reinforcement Learning

Shinya Honda,^{a,*}, Yuta Imura^b, Katsuhiko Sasaki^c, Ryo Takeda^d

^aFaculty of Engineering, Hokkaido University. Email: honda@eng.hokudai.ac.jp

^bFaculty of Engineering, Hokkaido University. Email: imu0004@eis.hokudai.ac.jp

^cFaculty of Engineering, Hokkaido University. Email: katsu@eng.hokudai.ac.jp

^dFaculty of Engineering, Hokkaido University. Email: r.takeda@eng.hokudai.ac.jp

Abstract

In this research, the authors developed an adaptive control method using deep reinforcement learning which is a kind of machine learning to suppress the vibration of smart structures. This method just requires information about the control response and input, and does not require numerical models for the controlled object to design the controller. We experimented to verify the effectiveness of this method. In this experiment, a smart structure fabricated by an aluminum plate and a piezoelectric actuator was used as a controlled object. Three kinds of reinforcement learning algorithms are employed, Deep Q Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Twin Delayed DDPG (TD3), and the control performance is compared. As a result, we succeeded in reducing the H_∞ norm of the frequency response to impulse disturbance by up to about 40 dB compared to the uncontrolled case. This demonstrates the applicability of the control method using deep reinforcement learning to adaptive vibration control.

Keywords: Adaptive control; reinforcement learning; smart structure; vibration control

1. Introduction

Vibration control of mechanical structures is an important technology to prevent unexpected behavior and damage caused by vibration. In recent years, it has become difficult to achieve the desired vibration characteristics only by structural design due to the increasing performance of mechanical structures, and the importance of active vibration control technology has increased. Therefore, various control theories, such as classical control theory and modern control theory, have been developed and used in practice. However, many of these methods have problems such as the need to model the control target precisely and the deterioration of control performance when the surrounding environment or the system to be controlled changes.

As the potential use of artificial intelligence (AI) is widely discussed in many research fields [1], control utilizing machine learning, a type of AI technology, has also been attracting attention as a new control method that solves these problems. Machine learning is an analytical technique that uses a computer to learn a large amount of data to find useful patterns in the data [2]. By utilizing this technology for vibration control, it is possible to control an unknown control target by learning, without modeling the target. In addition, continuous learning is expected to suppress the degradation of control performance due to

changes in the environment, including changes in the surrounding environment and the control target. The following are examples of research on vibration control methods using machine learning. Yang et al. performed vibration control of smart structures using multi-neural networks [3]. Honda et al. conducted adaptive vibration control using self-organizing maps [4].

In this study, we investigated vibration control using deep reinforcement learning in machine learning. Reinforcement learning is a technique in which a computer learns and acquires the optimal behavior for performing a task from experience by interacting with a dynamic environment in a trial-and-error manner [5]. In particular, deep reinforcement learning, which utilizes deep learning for conventional reinforcement learning, is characterized by its ability to handle complex problems with nonlinearities and has demonstrated performance that can beat human players in the field of games such as chess [6]. However, although there are some examples of applications of deep reinforcement learning in control, such as control of an inverted pendulum by reinforcement learning [7], [8] and parameter tuning of PID control [9], [10], there are few examples of vibration control of continuous bodies by reinforcement learning. Mu et. al. proposed flutter suppression method based on machine learning [11], and Samaitis et. al., evaluated adhesive bond quality based on machine learning and pulse-echo immersion data [12]. Therefore, to develop an adaptive vibration control method using deep reinforcement

*Corresponding author. Tel.: +81-11-706-6415

Kita-13, Nishi-8, Kita-ku
Sapporo, Hokkaido, Japan, 060-8628

learning, this study proposed and compared several vibration control methods using different reinforcement learning algorithms.

To demonstrate that the reinforcement learning methods is effective vibration control method for the smart structure with continuous body, a smart structure consisting of a thin aluminum plate and piezoelectric actuators was employed as the control target. We applied a control method using deep reinforcement learning to this structure and conducted vibration control experiments to verify the effectiveness of this method. As a result, the controller was successfully constructed by learning without using any information about the control target, thus demonstrating the applicability of the method to adaptive vibration control.

2. Controller Configuration using Deep Reinforcement Learning

2.1. Reinforcement learning overview

Reinforcement learning is a machine learning technique. In this method, a computer learns the best way to solve a problem by collecting information through interaction with its environment. A schematic diagram of reinforcement learning is shown in Fig. 1. In reinforcement learning, the subject that interacts with the environment is called an agent, which selects an action according to its own action selection rule (policy) based on the observed state of the environment (state) and acts against the environment. The agent then receives an evaluation of its action as a reward from the environment. The goal of reinforcement learning is to maximize this reward by learning and improving the strategy through interaction with the environment and to obtain the optimal strategy.

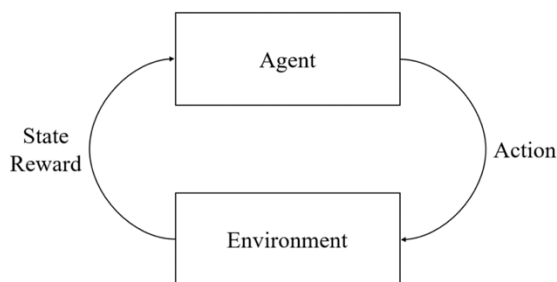


Figure 1. Diagram of reinforcement learning

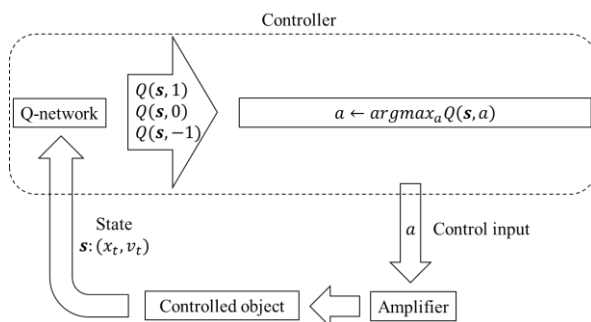


Figure 2. DQN controller

By applying reinforcement learning methods to vibration control, it is possible to adjust appropriate control inputs through learning, making model-free control possible. It is also possible to construct a system that can respond to changes in the control target and environment. However, since the control system is not designed using a detailed dynamic model of the controlled object, its control performance may be inferior to that of conventional models.

The method that utilizes deep learning for reinforcement learning is called deep reinforcement learning. This method uses Deep Neural Networks (DNNs) to process information about states observed in the environment and to represent policies, and DNNs make it possible to store and update these models more efficiently. Widely used three types of deep reinforcement learning algorithms were used in this study: Deep Q Network (DQN), Deep Deterministic Policy Gradient (DDPG), and Twin Delayed DDPG (TD3).

DQN is a deep reinforcement learning algorithm developed by DeepMind, Inc. in 2013 [13]. The algorithm uses a Q-network to learn an action-value function (Q-function), which represents the value of each possible action in each state of the environment. After learning, the agent selects the action that maximizes the value of the Q-network according to the observed state and maximizes the reward. Due to the algorithm of selecting the action that maximizes the Q-function from multiple actions, the number of actions handled is finite, so the action space becomes a discrete value.

DDPG is proposed by Lillicrap et al. in 2015 [14]. This algorithm uses two networks for learning: an Actor-network, which outputs actions based on the state observed from the environment as input, and a Critic network, which evaluates the Actor-network. The agent's behavior is determined by the output of the Actor-network. Therefore, unlike DQN, the action space becomes continuous.

TD3 is an improved method of DDPG [15]. The basic method is the same as DDPG, but three additional methods called Clipped Double Q learning, Target Policy Smoothing, and Delayed Policy Update are added to stabilize the learning process.

2.2. Controller configuration with DQN

The smart structure that is the control object in this study is measured by its velocity and displacement by non-contact sensors. Since a piezoelectric actuator is attached as the control device, the control input becomes the voltage. The controller shown in Fig. 2 was configured using a Q-network for vibration control by DQN. The state \mathbf{s} given to the Q-network is set to $\mathbf{s} = (x_t, v_t)$ where x_t is the displacement observed by the sensor and v_t is the velocity. The action space A , which represents the set of actions that can be selected by the agent, is set to $A = \{-1, 0, 1\}$. The observed state $\mathbf{s} = (x_t, v_t)$ and the selectable actions $a \in A$ are used as inputs to the Q-network, and the value of the Q function $Q(\mathbf{s}, a)$ is computed for all a . Then, the value a that maximizes

$Q(s, a)$ is input to the actuator through an amplifier as the control input voltage.

The reward r given to the agent during learning is set to $r = -|x_{t+1}| - 0.05|a_t|$ where a_t is the agent's control input and x_{t+1} is the displacement at the next time step after the control input. The constant 0.05 was determined by trial and error by numerical experiment. From the first term, the reward increases as the absolute value of x_{t+1} decreases. This means that the agent learns to select control inputs that reduce the absolute value of x_{t+1} , and as a result, vibration suppression can be expected. In addition, the second term is smaller than the first term, so energy-saving and efficient control can also be expected. When learning and updating the controller, the Q-network is learned and updated based on the DQN algorithm using the above $s, a,$ and r .

2.3. Controller configuration with DDPG and TD3

The controller shown in Fig. 3 is constructed using the Actor network for DDPG/TD3. The state s input to the network is set to $s = (x_t, v_t)$ as in the DQN case. The controller inputs s observed from the target to the Actor network and calculates the output $\mu(s)$. In this case, $-1 < \mu(s) < 1$ becomes continuous values. The output a from the controller is $a = \mu(s)$. This a is input to the actuator through an amplifier as the control input voltage. The reward r given to the agent during learning is defined as $r = -|x_{t+1}| - 0.05|a_t|$ as in the DQN case. The above $s, a,$ and r are used to train and update the Actor-network based on the DDPG/TD3 algorithm, and the controller is trained and updated.

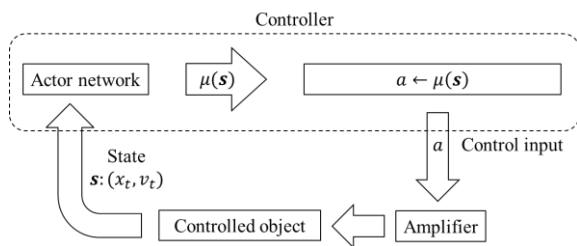


Figure 3. DDPG and TD3 controller

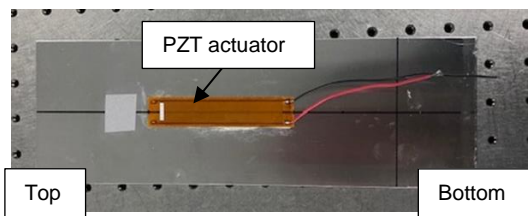


Figure 4. Smart structure composed of Al plate and PZT actuator

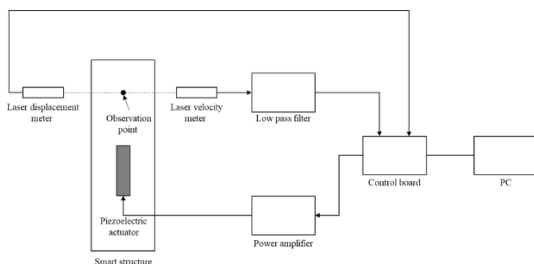


Figure 5. Diagram of the experimental system

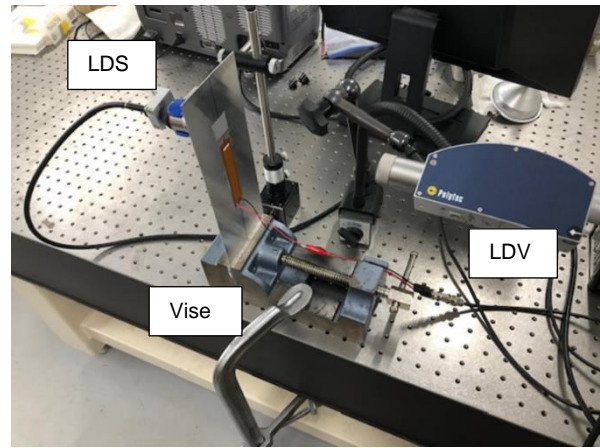


Figure 6. Experimental set-up

3. Experiments

3.1. Experimental setup

The smart structure shown in Figure 4 was employed as the control object. In this study, we define a structure equipped with actuators that can generate control forces as a smart structure. Therefore, the structure consists of a piezoelectric actuator (PZT) attached to a flat aluminum plate. The dimensions of the aluminum plate are 300 mm long, 100 mm wide, and 1 mm thick. The piezoelectric actuator was glued to one side of the plate from 80 mm to 180 mm from the top edge on the center line. The piezoelectric actuator M8514-P1 (Smart Materia) is employed here.

A schematic diagram and picture of the experimental setup is shown in Figures 5 and 6. In this system, the smart structure is clamped by a vise attached to a vibration-isolation table up to 50 mm from the bottom edge. A laser displacement sensor (LDS) and a laser doppler vibrometer (LDV) were installed in front of and behind the smart structure to observe the displacement and velocity in the thickness direction at the center of the structure at a position 60 mm from the top edge. The observed displacement and the velocity are sent to the control PC through a low-pass filter and control board. The control PC calculates the control input according to the observed displacement and velocity and sends the signal to the power amplifier through the control board. The control input voltage is amplified by the power amplifier and applied to the piezoelectric actuator for control. The sampling frequency and the frequency of the control input are set at 100 Hz.

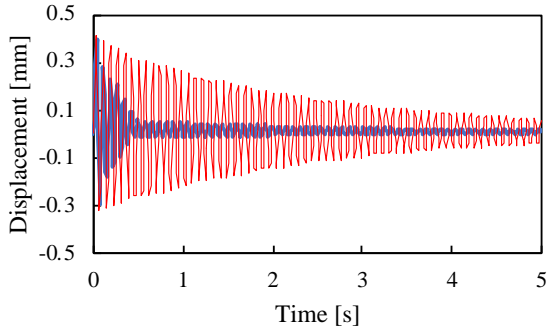
3.2. Vibration control for impulse disturbance

A voltage was applied to the piezoelectric actuator for 0.01 s (one-time step) to simulate impulse excitation and vibrate the structure. After that, vibration control was performed using a controller for 5 seconds. Next, the controller was trained and updated based on the DQN, DDPG, and TD3 algorithms using measured displacement and velocity data. Vibration control experiments were then conducted again using the updated controllers. This was repeated 100 times for one episode. After that, vibration control experiments were conducted using the controllers

that had been trained throughout the 100 episodes to verify the control performance.

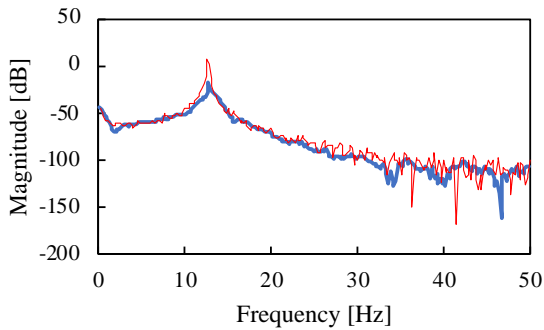
3.3 Vibration control for random disturbance

Vibration control experiments were also conducted for random disturbances to confirm the control performances under more realistic environments.



— DQN control — Without control

(a) Time response



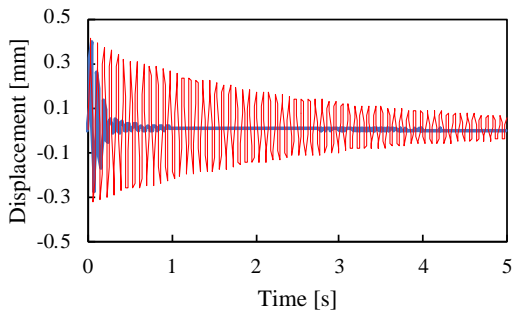
— DQN control — Without control

(b) Frequency response

Figure 7. Control response to impulse disturbance (DQN).

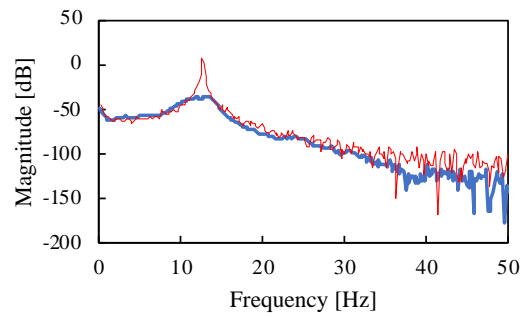
Table 1. Control result for impulse disturbance.

	Displacement [mm]	H_{∞} norm [dB]
DQN	11.95	-18.67
DDPG	5.27	-35.55
TD3	5.47	-33.85
W/O control	51.02	5.41



— DDPG control — Without control

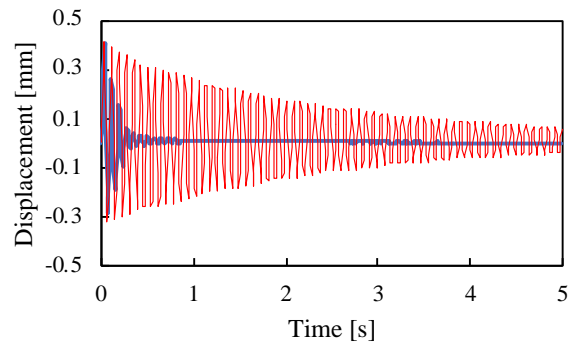
(a) Time response



— DDPG control — Without control

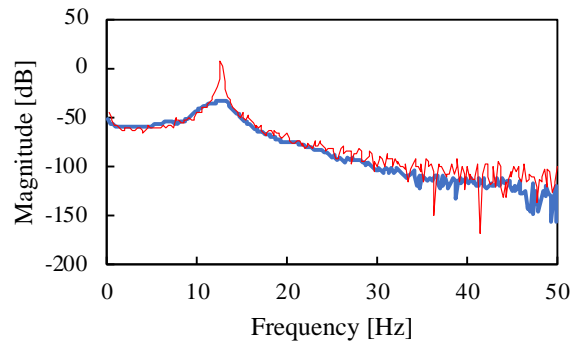
(b) Frequency response

Figure 8. Control response to impulse disturbance (DDPG).



— TD3 control — Without control

(a) Time response

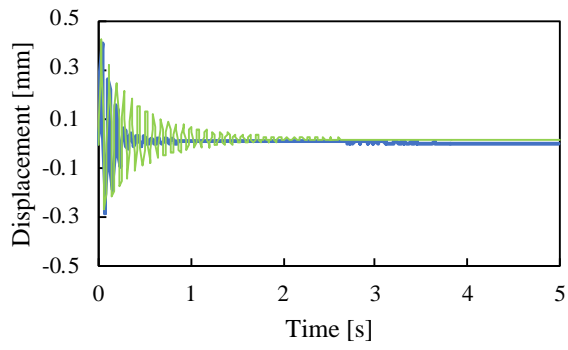


— TD3 control — Without control

(b) Frequency response

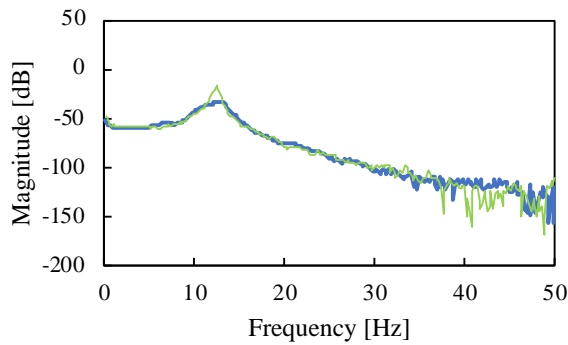
Figure 9. Control response to impulse disturbance (TD3).

In this experiment, TD3 was used as the algorithm for control and controller training. To apply the random disturbance, a piezoelectric actuator for the disturbance was attached to the back side of the smart structure. The voltage signal following uniform random numbers was applied to the actuator for a period of 5 seconds. The controller was then trained for 100 episodes as in the vibration control experiment for impulse disturbances.



— TD3 control — Proportional control

(a) Time response



— TD3 control — Proportional control

(b) Frequency response

Figure 10. Comparison of TD3 with proportional control

Table 2. Control results for random disturbance

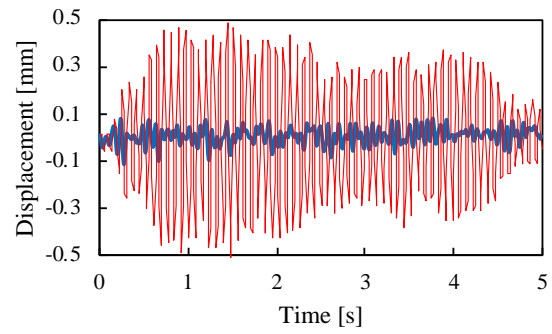
	Displacement [mm]	H_{∞} norm [dB]
TD3	13.82	-34.78
W/O control	102.91	20.34

4. Result and Discussion

4.1 Results for impulse disturbance

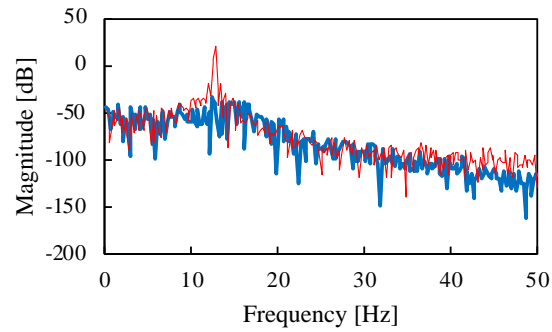
Figures 7-9 show the time and frequency responses of displacement when using controllers configured with the DQN, DDPG, and TD3 algorithms, respectively. The amplitude of vibration for each method is much smaller than that for the uncontrolled method, indicating that the vibration was successfully controlled by the control method. The total displacement and H_{∞} norm of the displacement are listed in Table 1. From the table, the control performance using DDPG and TD3 are higher than that using DQN. This is because the control input of DQN is discrete, whereas DDPG/TD3 has a continuous control input, providing a more optimal control input.

On the other hand, there was no significant difference in control performance between DDPG and TD3. This is because the two algorithms are similar, and the problem treated in this study was relatively simple, so the advantages of TD3, such as its high learning stability, were not fully utilized.



— TD3 control — Without control

(a) Time response



— TD3 control — Without control

(b) Frequency response

Figure 11. Control response to random disturbance

Figure 10 compares the time and frequency responses of displacement when using the controller with TD3 and using the simple proportional controller. The gain of the proportional controller was set so that the maximum value of the control input matches that of the controller using reinforcement learning. From Figure 10, it is known that the TD3 control method achieves faster vibration damping and lower peak than the proportional control method. Therefore, it can be said that the TD3 control method indicates higher control performance than the simple proportional control.

4.2 Results for random disturbance

The time and frequency responses of the displacements with the TD3 controller, after learning against the random disturbance was completed, are shown in Figure 11. The total displacement and the H_{∞} norm of the displacement are listed in Table 2. The vibration amplitude is significantly suppressed compared to the uncontrolled case, and the H_{∞} norm is reduced by about 55 dB. This indicates that the present controller is effective to suppress the vibration caused by random disturbances.

5. Conclusions

In this study, we proposed a controller configuration method for an unknown control target based on three different deep reinforcement learning algorithms. Vibration control experiments were conducted on a smart structure consisting of an aluminum plate and

piezoelectric actuators to verify the learning and control performance of these methods. Since the method is model-free, it is applicable to other continuum smart structures.

Vibration control experiments against impulse disturbances were conducted using controller configuration methods based on the DQN, DDPG, and TD3 algorithms. As a result, the H_∞ norm of the response of the controller with the DQN was about 24 dB lower than that without control, and the controllers with the DDPG and TD3 algorithm were about 40 dB lower than that without control.

Vibration control experiments against random disturbances were also conducted to verify the performance assuming under more realistic conditions. The controller based on TD3 was used. As a result, the H_∞ norm of the frequency response was successfully reduced by about 55 dB after the controller completed training compared to the uncontrolled case. This result confirms that the proposed method is applicable to random disturbances.

References

- [1] M. Soori, B. Arezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics, a review," *Cogn. Robot.*, vol. 3, pp. 54–70, 2023.
- [2] S. Raschka, *Python Machine Learning, 2nd Edition*. Packt Publishing, 2017.
- [3] S. M. Yang and Y. J. Lee, "Vibration Control of Smart Structures by Using Neural Networks," *J. Dyn. Syst. Meas. Control*, vol. 119, no. 1, pp. 34–39, 1997.
- [4] S. Honda, Y. Narita, and N. Kida, "Adaptive control for vibration suppression by using self-organization map," *10th Asian Control Conf.*, 2015.
- [5] Y. Li, *Deep Reinforcement Learning: An Overview*. 2017.
- [6] D. Silver *et al.*, "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play," *Science*, vol. 362, no. 6419, pp. 1140–1144, 2018.
- [7] W. Linglin, L. Yongxin, and Z. Xiaoke, "Design of reinforce learning control algorithm and verified in inverted pendulum," *34th Chinese Control Conf.*, 2015.
- [8] G. Puriel-Gil, W. Yu, and H. Sossa, "Reinforcement Learning Compensation based PD Control for Inverted Pendulum," *15th Int. Conf. Electr. Eng. Comput. Sci. Autom. Control*, 2018.
- [9] X. Wang, Y. Cheng, and W. Sun, "A Proposal of Adaptive PID Controller Based on Reinforcement Learning," *J. China Univ. Min. Technol.*, vol. 17, no. 1, pp. 40–44, 2007.
- [10] Y. Qin, W. Zhang, J. Shi, and J. Liu, "Improve PID controller through reinforcement learning," *IEEE CSAA Guid. Navig. Control Conf.*, 2018.
- [11] X. Mu, R. Huang, Q. Zou, and H. Hu, "Machine learning-based active flutter suppression for a flexible flying-wing aircraft," *J. Sound Vib.*, vol. 529, 2022.
- [12] V. Samaitis, B. Yilmaz, and E. Jasiuniene, "Adhesive bond quality classification using machine learning algorithms based on ultrasonic pulse-echo immersion data," *J. Sound Vib.*, vol. 546, 2023.
- [13] V. Mnih *et al.*, *Playing Atari with Deep Reinforcement Learning*. 2013.
- [14] T. P. Lillicrap *et al.*, *Continuous control with deep reinforcement learning*. 2015.
- [15] S. Fujimoto, H. van Hoof, and D. Meger, *Addressing Function Approximation Error in Actor-Critic Methods*. 2018.