symmetry

MDPI

Article

A Novel FBG Placement Optimization Method for Tunnel Monitoring Based on WOA and Deep Q-Network

Jiguo Liu [1,2], Ming Song [1,2,*], Heng Shu [1,2], Wenbo Peng [1,2], Longhai Wei [1,2] and Kai Wang [1,2]

[1] Research and Development Center on Tunnel and Underground Space Technology, China Communications Construction Company, Wuhan 430056, China
[2] CCCC Second Highway Consultant Co., Ltd., Wuhan 430056, China
* Correspondence: songming2@cccltd.cn

Abstract: By employing the whale optimization algorithm's (WOA) capability to reduce the probability of being stuck in a locally optimal solution, this study proposed an improved WOA-DQN algorithm based on the Deep Q-Network algorithm (DQN). Firstly, the mathematical model of Fiber Bragg Grating (FBG) sensor placement was established to calculate the reward of DQN. Secondly, the effectiveness and applicability of WOA-DQN were validated through experiments in nine cases. It indicated that the algorithm is far superior to other methods (Noisy DQN, Prioritized DQN, DQN, WOA), especially with the learning rate of 0.001, the initial noise 0.4, the hidden layer 3–512, and the updated frequency of 20. Finally, the FBG sensors were placed at [0°, 27°, 30°, 47°, 51°, 111°, 126°, 219°, 221°, 289°] to detect the accurate deformation of the tunnel with the maximum error 8.66 mm, which is better than the traditional placement. In conclusion, the algorithm provides a theoretical foundation for sensor placement and improves monitoring accuracy. It further shows great promise for deformation monitoring in tunnels.

Keywords: sensor placement; whale optimization algorithm; Deep Q-Network; tunnel monitoring

check for updates

Citation: Liu, J.; Song, M.; Shu, H.; Peng, W.; Wei, L.; Wang, K. A Novel FBG Placement Optimization Method for Tunnel Monitoring Based on WOA and Deep Q-Network. *Symmetry* **2024**, *16*, 1400. https://doi.org/10.3390/sym16101400

Academic Editor: Hsien-Chung Wu

Received: 15 September 2024
Revised: 2 October 2024
Accepted: 12 October 2024
Published: 21 October 2024

1. Introduction

Due to factors like changes in the geological environment, changes in hydrological characteristics, geological disasters, and vehicle loads in the tunnel, tunnel structural diseases such as cracks, deformation, and water leakage are prone to occur, which affect the structural health of the tunnel (Tan et al. [1]; Liu et al. [2]; Yang et al. [3]; Pan et al. [4]). If the tunnel structure is not maintained promptly, it could result in significant financial losses and casualties. The tunnel's structural deformation provides information about the tunnel's internal structural stress. One of the most efficient ways to guarantee the safety of the tunnel structure is to analyze deformation from vast field monitoring data (Huang et al. [5]; Xing et al. [6]; Duan et al. [7]). At present, a large number of researchers have carried out studies on the 3D deformation of tunnels. Simeoni & Zanei [8] proposed a procedure to evaluate the accuracy of convergence measurements by using distometers. The measurements with the distometers for 584 lines took eight days and two workers. This technique could seem time-consuming and uncomfortable. Puente et al. [9] described an automatic method for the detection of tunnel luminaires as well as easily obtaining their 3D spatial location using colored 3D point clouds. Wang et al. [10] proposed that the profile-image method, by which the profile of a tunnel can be determined by a laser-lit profile in an image, is tested in the study to validate its applicability. According to the theory of tunnel displacement as measured by a total station with 3D coordinate measurement and with remote distance measurement, Luo et al. [11] presented the formulas for tunnel crown settlement and horizontal displacement measurement. The results of these studies have resulted in a high-quality 3D reconstruction model of tunnel deformation. However, due to the accuracy of convergent meters, the measurement error cannot be reduced. The cost-prohibitive ground 3D laser scanning and total station rely on manual on-site data collection, so they cannot provide real-time tunnel deformation. Even though video image makes it possible to create a tunnel deformation model faster and in real time, the demand for detection precision cannot be met. Fiber grating sensing technology has been extensively applied in the field of tunnel health monitoring (Kinet et al. [12]; Minardo et al. [13]; Feng et al. [14]). It has stronger anti-interference and anti-corrosion capabilities and higher detection accuracy compared to traditional methods.

To ensure timely and accurate 3D reconstruction of tunnel deformation, it is necessary to optimize the placement of Fiber Bragg Grating (FBG) sensors. Numerous researchers have recently optimized sensor placement of specific monitoring targets using various methods to improve the effectiveness and accuracy of monitoring. Kammer et al. [15] obtained the best placement and the number of sensors by removing the degree of freedom with the least independent contribution. In the past, sensors were installed uniformly throughout the tunnel section (Lai et al. [16]). Due to the tunnel's particular design and the varied stresses at each site, sensors whose placement relies on subjective experience cannot provide useful monitoring data. Xia et al. [17] proposed a hunting underwater method based on level sets to optimize the placement of underwater sensors, which improved the monitoring results of the sensor network. Li Mei et al. [18] employed a genetic algorithm to optimize the configuration of fiber optic sensor networks with the signal attenuation of the sensor network as the objective function, and the research indicated that it functioned well in practical application. The sensor placement produced by these methods cannot provide the best accuracy for monitoring and a theoretical basis. Therefore, it is essential to keep researching optimization methods.

In recent studies on the whale optimization algorithm (WOA), Huang et al. [19] conducted a comprehensive systematic review of the theoretical foundation, improvement strategies, and hybrid algorithms of WOA, revealing its improved optimization performance. Sun et al. [20] proposed an improved WOA based on nonlinear parameters and feedback mechanisms, significantly enhancing its ability to solve high-dimensional optimization problems. Meanwhile, Habib et al. [21] used the improved WOA algorithm to enhance the stability and transient response performance of an automatic voltage regulator system, demonstrating its effectiveness in control systems. Furthermore, Li et al. [22] applied an improved WOA based on a hybrid strategy to the problem of locating electric vehicle charging stations, demonstrating the algorithm's application prospects in complex decision-making. Deep reinforcement learning (DRL) has performed well in optimization problems in various fields. The agent in RL continuously interacts with the environment to find new optimal strategies. Currently, RL can be divided into three types, including DRL based on value function (Luo et al. [23]; Sun et al. [24]), policy function (Meng et al. [25]; Liu et al. [26]), and Actor-Critic framework (Huang et al. [27]). Deep Q-Network (DQN) is one of the DRL algorithms based on value function. The exploration method of the DQN is to use the ε-greedy policy to select the exploratory behavior. Although it can theoretically explore the environment globally, it is limited by the experience replay and cannot achieve global exploration for actual storage applications. The single DQN has the drawbacks of a long convergence time or even failure to converge.

In order to improve the shortcoming, this study introduces the Whale Optimization Algorithm (WOA) which is a swarm intelligence algorithm proposed by Mirjalili et al. [28]. It is divided into three stages including encircling prey, bubble-net attacking, and searching for prey. All individuals will, with a certain probability, either attack with the bubble-net method or encircle the prey when trying to find the best solution. Due to its special searching method, it is utilized frequently in a variety of optimization fields (He et al. [29]; Lou et al. [30]; Pan et al. [31]).

WOA enhances the DQN network by leveraging its powerful global search capabilities, helping to overcome the limitations of DQN in sensor placement problems such as local optimality, limited motion space, and insufficient exploration capabilities. Specifically, WOA can automatically optimize DQN hyperparameters, improve exploration efficiency, and dynamically adjust action decisions. This enables DQN to find optimal solutions in complex sensor placement scenarios, thereby improving monitoring coverage and overall system performance. Therefore, this study proposed an innovative method WOA-DQN based on DQN and WOA to enable DQN to effectively and quickly obtain the optimal FBG placement in tunnel. The adaptability of WOA-DQN was verified by the tunnel numerical model and the optimal FBG placement was obtained which can get accurate tunnel deformation. It offers a theoretical basis for the sensor placement in tunnel health monitoring and a new idea for sensor placement in other fields.

In order to verify the actual effect of the sensor layout, multiple experimental comparative tests were conducted, and the results showed that a completely symmetrical sensor layout cannot achieve the expected monitoring effect in complex environments. Because in practical environments, it is rare to fully achieve this symmetry. For example, geological structure, construction errors, material properties, air flow, traffic loads, and other factors can all lead to the asymmetry of the actual tunnel environment. Therefore, sensor layouts based on symmetry design may not be able to handle these subtle changes. The experimental results indicate that in nonideal tunnel environments, symmetric layouts often overlook important local variations. An asymmetric layout can flexibly respond to these practical situations and avoid blind spots in data collection in symmetrical design. In addition, experiments have shown that asymmetric layouts can better cope with dynamically changing environmental conditions, especially in long-term monitoring, exhibiting stronger robustness and adaptability.

2. Mathematical Model of FBG Sensor Placement

2.1. Problem Description

This study focused on the tunnel deformation during the elastic stage, excluding the other situations during the plastic deformation stage. Based on the relationship among axial force, bending moment, and curvature on the section of the tunnel, the objective in optimizing FBG sensor placement (OSP) is to reconstruct the precise deformation. Root Mean Square Error (RMSE) is used to measure the performance of reconstruction in this study. In engineering, the maximum reconstruction error of a monitoring point is below 20 mm. Therefore, the error is constrained as RMSE $\leq 20 \cdot N$ mm when N sensors are installed and is defined as Equation (1).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{N}\left(\left(x_i - x'_i\right)^2 + \left(y_i - y'_i\right)^2\right)}{N}} \tag{1}$$

where (x_i, y_i) is the real coordinate of monitoring point, and (x'_i, y'_i) is the obtained coordinate by the principle of curve reconstruction in Section 2.2.

2.2. Principle of Tunnel Cross-Section Curve Reconstruction

Assuming that $X = [\theta_1, \ldots, \theta_N]$ is the chosen sensor placement. According to the rules of the circle, the angle θ_i between two FBG positions is constrained as $0 \leq \theta_1 < \theta_2 < \cdots < \theta_N < 360$, where N is the number of the used sensors.

The curvature of tunnel section can be calculated by the center wavelength variation obtained (Parent et al. [32]; Zhang et al. [33]), allowing for the tunnel cross-section curve to be reconstructed. As shown in Figure 1, O_1 is the center of the arc, Q_0Q_1, θ_1 is the center angle and r_1 represents the radius. Regardless of the curvature being 0, the coordinates of $Q_1(x_1, y_1)$ can be calculated as Equation (2).

$$\begin{cases} x_1 = x_0 + r_1 \cdot \sin\theta_1 \\ y_1 = y_0 + r_1 \cdot (1 - \cos\theta_1) \end{cases} \tag{2}$$

where κ_1 is the curvature of Q_1, $x_0 = 0$, $y_0 = 0$, $r_1 = 1/\kappa_1$.

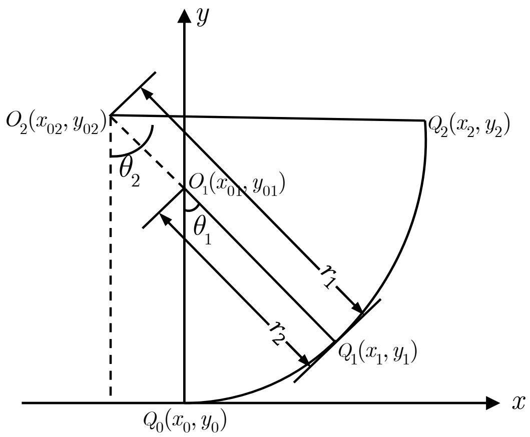

Figure 1. Reconstruction of tunnel cross-section curve.

The arc Q_1Q_2 starts at $Q_1(x_1, y_1)$ and ends at $Q_2(x_2, y_2)$, the center is $O_2(x_{02}, y_{02})$, θ_2 denotes the center angle, and r_2 represents the radius, so the Q_2 can be calculated as Equation (3).

$$\begin{cases} \theta_2 = \theta_1 + \theta'_2 \\ x_2 = x_1 - r_2 \cdot \sin\theta_1 + r_2 \cdot \sin\theta_2 \\ y_2 = y_1 + r_2 \cdot \cos\theta_1 - r_2 \cdot \cos\theta_2 \end{cases} \tag{3}$$

where κ_2 is the curvature of Q_2, and $r_2 = 1/\kappa_2$.

Similarly, the arc segment Q_iQ_{i+1} starts at $Q_i(x_i, y_i)$ and ends at $Q_{i+1}(x_{i+1}, y_{i+1})$, the center of $Q_{i+1}Q_i$ is O_{i+1}, r_{i+1} denotes the radius, θ_{i+1} represents the central angle, so the Q_{i+1} can be obtained as Equation (4).

$$\begin{cases} \theta_{i+1} = \theta_i + \theta'_{i+1} \\ x_{i+1} = x_i - r_{i+1} \cdot \sin\theta_i + r_{i+1} \cdot \sin\theta_{i+1} \\ y_{i+1} = y_i + r_{i+1} \cdot \cos\theta_i - r_{i+1} \cdot \cos\theta_{i+1} \end{cases} \tag{4}$$

where $r_{i+1} = 1/\kappa_{i+1}$, κ_{i+1} denotes the curvature of Q_{i+1}.

After analysis, the cases which curvature is negative or zero are shown in Figure 2. Therefore, the $Q_{i+1}(x_{i+1}, y_{i+1})$ can be obtained as Equation (5).

$$\begin{cases} x_{i+1} = x_i + sign(-\kappa_{i+1}) \cdot r_{i+1} \cdot \sin\theta_i + sign(\kappa_{i+1}) \cdot r_{i+1} \cdot \sin\theta_{i+1} \\ y_{i+1} = y_i + sign(\kappa_i \cdot \kappa_{i+1}) \cdot r_{i+1} \cdot \cos\theta_i - r_{i+1} \cdot \cos\theta_{i+1} \end{cases} \tag{5}$$

where $f(x)$ and θ_{i+1} is calculated as Equations (6) and (7).

$$sign(x) = \begin{cases} 1\,, & if\ x \geq 0 \\ -1\,, & if\ x < 0 \end{cases} \tag{6}$$

$$\theta_{i+1} = \begin{cases} \theta_i + \theta'_{i+1}\,, & if\ \kappa_i \cdot \kappa_{i+1} \geq 0 \\ \pi - \theta_i + \theta'_{i+1}\,, & if\ \kappa_i \cdot \kappa_{i+1} < 0 \end{cases} \tag{7}$$

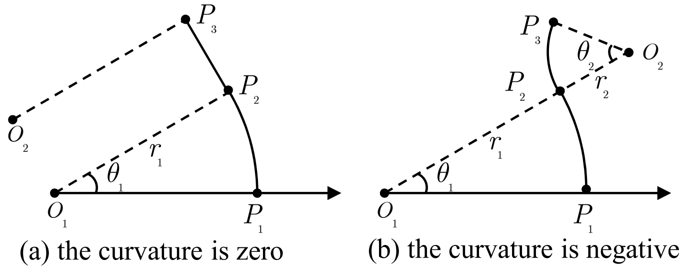

Figure 2. Special cases of curve fitting.

3. Proposed WOA-DQN Optimization Algorithm

3.1. Definition of WOA-DQN

Based on Deep Q-Network (DQN) and Whale Optimization Algorithm (WOA), this study proposes an intelligent optimization method named WOA-DQN for placement optimization of FBG sensors. The agent policy network and the cumulative reward value in DQN is used as the population individual and objective function in WOA, respectively. After iterating many times in DQN, WOA-DQN carries out a WOA iteration. In this case, the WOA individual uses the total reward instead of the single-step state to calculate the fitness. Finally, the optimal FBG placement solution can be obtained.

3.2. Behavior Description of WOA-DQN

3.2.1. Markov Decision Process

In this study, each selected FBG position is defined as an independent state, and the process of selecting and changing position is redefined as a Markov Decision Process (Luo et al. [23]; Sun et al. [24]). The state s_j is defined as the current sequence of FBG positions as Equation (8).

$$s_i = (P_1, P_2, \ldots, P_k) \tag{8}$$

where $i = 0,\ 1,\ 2,\ \ldots,\ N_a$. N_a is the maximum number of actions in a round, $k = 1, 2, \ldots, m$. m is the maximum number of optional positions, and the value of P_k is related to the angle. The initial state s_0, where sensors are installed uniformly in the tunnel section is expressed as Equation (9).

$$s_0 = \left(\frac{360 \times 0}{m}, \frac{360 \times 1}{m}, \ldots, \frac{360 \times j}{m} \right) \tag{9}$$

where $j = 1, 2, \ldots, m-1$. After the agent starts at s_0, it chooses the next action a_i by decision-making control to get the next state s_{i+1} and the obtain the reward r_{i+1} until the termination condition is met. The agent state formula is expressed as Equation (10).

$$s_{i+1} = \varphi(s_i, a_i) \tag{10}$$

where $i = 0,\ 1,\ 2,\ \ldots,\ N_a - 1$, $\varphi(s_i, a_i)$ is the state transition function; thus, the MDP of the algorithm is presented as Equation (11).

$$s_0 \overset{a_0}{\rightarrow} (s_1, r_1) \overset{a_1}{\rightarrow} (s_2, r_2) \cdots, (s_{N_s-1}, r_{N_s-1}) \overset{a_{N_s-1}}{\rightarrow} (s_{N_s}, r_{N_s}) \tag{11}$$

3.2.2. Policy

The policy is the way that the agent chooses the FBG position in current state, requiring the agent to learn continuously. The policy probability function p of choosing a action in state s is expressed as Equation (12).

$$p(a|s) \doteq \prod (s, a)\ s \in S, a \in A \tag{12}$$

where $\Pi(\cdot)$ is the mapping policy function, Q is the action-value function, S and A are the state set and action set, respectively. The policy π is defined as $\pi = \{\pi_1, \pi_2\}$., π_1 is the value of position that the agent chooses, π_2 is the action after the agent choosing π_1, and π is expressed as Equations (13) and (14).

$$\pi_1(a|s|) = po\ po = 0, 1, \ldots, n-1 \tag{13}$$

$$\pi_2(a|s, \pi_1(a|s)) = \begin{cases} -1 \\ 1 \end{cases} \tag{14}$$

With adding Gauss noise to the DQN network, the agent automatically adjusts the noise to explore unknown states as much as possible and effectively. The noise formula is defined as Equation (15).

$$\omega = \mu + \sum \odot \varepsilon \tag{15}$$

where ω is the weight of the noise, μ and Σ are the adaptive noises of the Noisy layer, and ε is the zero-mean Gauss noise.

3.2.3. Action Value Function

The action-value function Q is the expectation of the cumulative reward that the agent obtained in the environment. The agent chooses the action in the way of maximizing the total reward. Due to the optimum reward being $R_i = \sum_{i'=1}^{N_a} \gamma^{i'-i} r_{i'}$. The optimal function Q^* is Equation (16).

$$Q^*(s,a) = \max_\pi \mathbb{E}[R_i | s_i = s, a_i = a, \pi] \tag{16}$$

Since $Q^*(s,a)$ obeys the Bellman equation, the updated formula is expressed as Equation (17).

$$Q_{i+1}(s,a) = \mathbb{E}_{s'}[r + \gamma \max_{a'} Q_i(s',a') | s,a] \tag{17}$$

where s', a' represent the next state and action respectively.

In the DQN, the Q-network that is composed of multiple hidden layers is used to gradually approach the Q-function. After the Q-network receives the state, the value of Q for each possible action will be calculated and output. The loss is expressed as Equation (18).

$$L(\omega) = \mathbb{E}[(Q(s',a') - r + \gamma \max \hat{Q}(s',a';w'))^2] \tag{18}$$

where ω' is the weight of the objective network $\hat{Q}$, and ω is the weight of the network Q. Then, the stochastic gradient descent method is used to update the weight as Equation (19).

$$\omega \leftarrow \omega + \lambda \cdot \nabla_\omega L(\omega) \tag{19}$$

where λ is the learning rate of the neural network.

3.2.4. Reward

The reward value r_i in the current action is calculated as Equation (20). And the reward is determined by root mean square error (RMSE). In the reward equation, the reconfiguration error (e) corresponds to RMSE (Equation (1)) and is further stated in Equation (20).

$$r_i = \frac{e_i - e_0}{e_i + e_0}, e_i \geq 0, e_0 \geq 0, i = 1, \ldots, N_a \tag{20}$$

where e_i is the total error of the fitting curve in the current state s_i, e_0 is the total error of the fitting curve in the current state s_0, and r_i is limited in $(-1, 1]$.

In order to enable the agent to explore the environment and affect the latter actions, the total reward and evaluation function R_i are expressed as Equation (21).

$$R_i = r_i + \sum_{j=1}^{K} \gamma^j r_{i+j} \tag{21}$$

where γ is the attenuation rate of reward, K is the total action number, r_i is the reward of the action a_i.

3.2.5. Experience Replay

In this study, based on prioritized experience replay, the experience memory of each state is stored into the buffer pool and the experience of the current agent that does not perform well is prioritized to learn continuously. The priority weight is updated is as Equation (22).

$$\delta_i = |R_i + \gamma Q_t(s_{i-1}, a_{i-1}) - Q(s_{i-1}, a_{i-1})| \tag{22}$$

where $Q_t(\cdot)$ is the objective function and $Q(\cdot)$ will be trained in the WOA-DQN.

When there are n FBG sensors that would be installed in m optional positions, the agent of DQN can only move once to a new state closer that is closer to the current state in each round. Besides, there are various interference states between two local optimal states, the maximum number of actions N_a should be limited to avoid wasting too much computation resources. Therefore, N_a is constrained as $N_a \leq (m - n) \times n$. The procedure of WOA-DQN algorithm is as follows, and the Flowchart of the algorithm is shown in Figure 3.

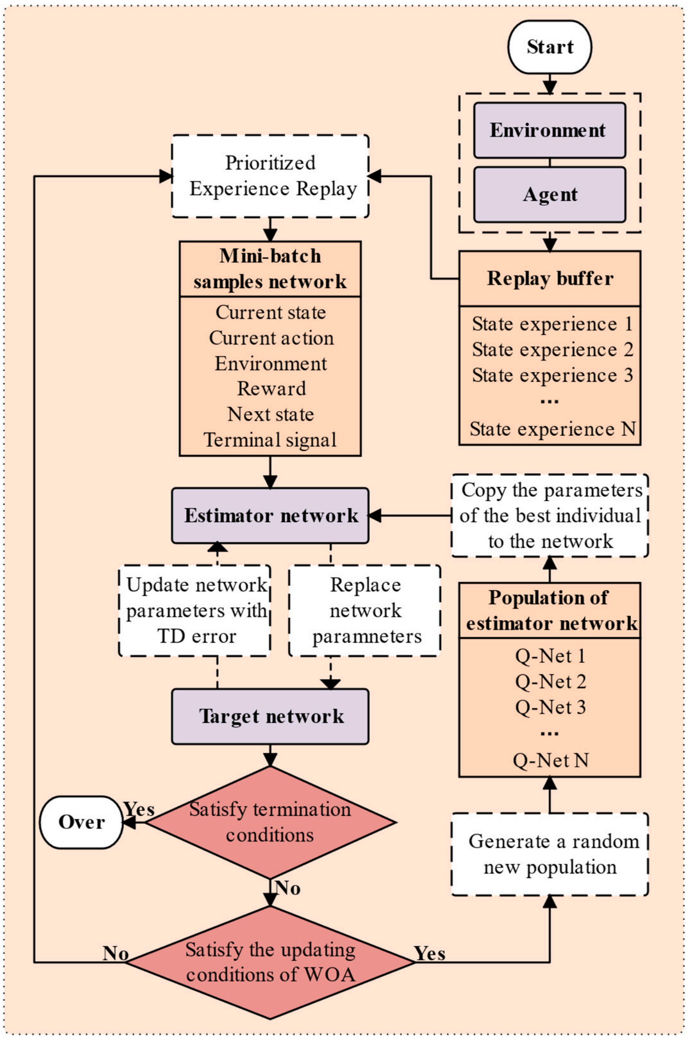

Figure 3. Flowchart of the WOA-DQN algorithm.

(1) The agent interacts with the environment in order to get enough experience replay to restore in the experience pool.
(2) The agent extracts a small batch of sample from the experience pool to train the policy network.
(3) After lots of training, the WOA begins to update the policy network.
(4) The WOA population is initialized to randomly generate many policy networks.
(5) The individual interacts with the environment to explore the better policy network.
(6) Based on the fitness of the individual, the optimal policy network is updated in the experience pool.
(7) If the termination condition is met, the agent will terminate, otherwise, repeat steps 2 to 6.

4. Verification Experiment and Analysis

4.1. Experiment Settings

In order to verify the effectiveness of WOA-DQN for FBG sensor placement optimization, the crossing-river shield tunnel experiment was carried out. The segment of the tunnel is 0.65 m thick, 15.5 m in outer diameter, and 14.2 m in inner diameter. The simulation models of ANSYS in nine loading cases were considered to build datasets. ANSYS [2021 R2] offers a comprehensive suite of analysis tools, a robust solver, extensive material libraries, and a user-friendly interface. Its versatility, accuracy, and industry acceptance make it a popular choice for engineering analyses, enabling researchers to simulate complex systems efficiently. The loading cases are shown in Table 1. The pressures of the shield tunnel include water pressure and earth pressure. Finally, the deformation datasets of the tunnel in each loading case were obtained. WOA-DQN was trained by deformation datasets based on a Dell (Dell Technologies, headquartered in Round Rock, TX, USA.) PowerEdge 740 server. The server was configured with two Intel (Intel Corporation, headquartered in Santa Clara, CA, USA.) Xeon Gold 6248 CPUs, an NVIDIA (NVIDIA Corporation, headquartered in Santa Clara, CA, USA.) Quadro RTX-8000 graphics card, and the experimental environment was python3.9-tensorflow2.6.

Table 1. Loading parameters of different cases.

Loading Cases	Water Pressure [MPa]	Earth Pressure [MPa]	Elastic Resistance Coefficient [MPa/m]
Case 1	2.00	0.50	0.25
Case 2	2.00	1.00	0.50
Case 3	2.00	1.50	0.75
Case 4	2.00	2.00	1.00
Case 5	4.00	0.50	0.25
Case 6	4.00	1.00	0.50
Case 7	4.00	1.50	0.75
Case 8	4.00	2.00	1.00
Case 9	5.00	3.00	1.50

4.2. Verification and Analysis

The WOA-DQN model was trained by using the deformation datasets of Case 1. The parameters of the model are shown in Table 2. The other initial algorithm parameters are shown in Table 3. As shown in Figure 4, the total reward became increasingly larger when the agent interacted with the environment continuously. The result shows that the value of the total reward increased rapidly in the early stages of exploration, and the WOA-DQN was convergent to the feasible globally optimal solution in the stable stage gradually.

Table 2. Initial algorithm parameters of WOA-DQN.

Parameters	Value
Learning rate λ	0.001
Reward attenuation rate γ	0.95
Number of training sample n_e	256
Capacity of experience replay pool B_e	500,000
Maximum training times N_e	60,000
Number of hidden layers Hl	3
Number of neurons in hidden layer Hn	512
Initial noise parameter ε_0	0.4
Update frequency of objective network K_d	20
Number of WOA population n_w	200
Max number of iteration in WOA N_w	5000
Logarithmic spiral constant of WOA b	1
Update frequency of WOA in objective network K_w	100
Number of FBG sensors	10

Table 3. Other initial algorithm parameters.

Algorithm	Parameters
Noisy DQN	$\lambda = 0.001, \gamma = 0.95, n_e = 256, B_e = 500{,}000, N_e = 60{,}000$
Prioritized DQN	$\lambda = 0.001, \gamma = 0.95, n_e = 256, B_e = 500{,}000, N_e = 60{,}000$
DQN	$\lambda = 0.001, \gamma = 0.95, n_e = 256, B_e = 500{,}000, N_e = 60{,}000$
WOA	$n_w = 200, N_w = 5000, b = 1, K_w = 100$

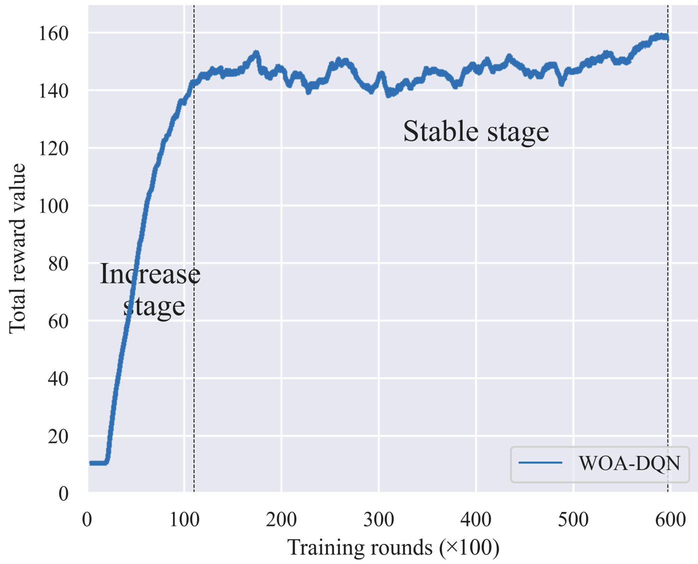

Figure 4. Training process and total reward of Case 1.

Finally, the optimal positions of sensors in Case 1 are [0°, 20°, 27°, 29°, 29°, 50°, 126°, 221°, 222°, 290°]. The maximum error is 2.26 mm and the total error is 9.87 mm when the model is applied to deformation reconstruction of Case 1, as shown in Figure 5. This meets engineering code and verifies that the algorithm is effective.

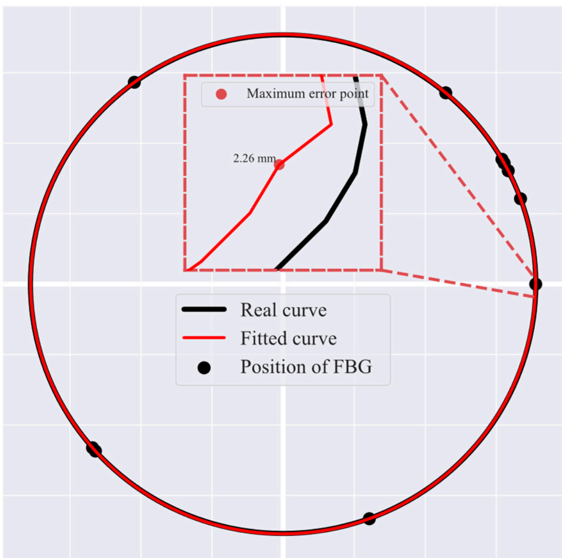

Figure 5. Tunnel reconstruction result of Case 1.

4.2.1. Quantity of FBG Sensors

Due to the excessive sensors, the agent cannot explore the environment fully to sink into a locally optimal solution as shown in Figure 6. Therefore, the optimal number of sensors needs to be discussed. This study discusses the influence of FBG sensor quantity based on the datasets of all cases. The minimum average error and standard deviation of different sensors were obtained as shown in Table 4. Obviously, the minimum average error is the smallest when the number of sensors is 10. In the following, the number of sensors is set at 10 for discussion and analysis.

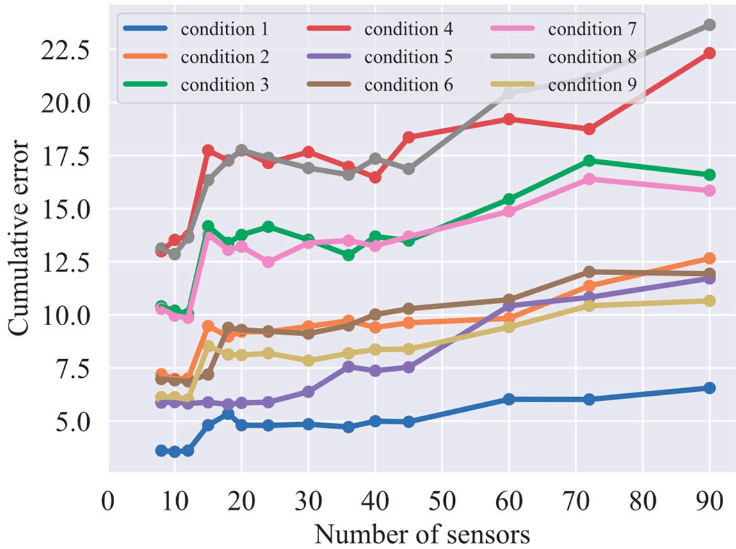

Figure 6. Performance of different numbers of sensors in all cases.

Table 4. The performance comparison of different sensors in Case 1.

Number of Sensors	Minimum Average Error [mm]	Standard Deviation [mm]
8	85.11	33.32
10	84.48	33.71
12	85.13	35.35
15	108.84	47.30
18	109.59	45.09
20	110.82	47.79
24	109.43	46.04
30	110.17	45.61
36	110.63	41.76
40	112.14	42.09
45	114.68	44.39
60	129.34	48.31
72	137.99	48.33
90	146.61	55.49

4.2.2. Analysis of the Traditional Placement

After the discussion above, the sensor number for each case was set as 10 within the range [8, 90]. The WOA-DQN was used to search for the optimal placement of all cases. By analysis of the result shown in Table 5, FBG sensors were installed near the spandrel and skewback which were the key position on the reconstruction accuracy of the deformation in the tunnel. Considering the placement of FBG sensors in each case, the position of FBG sensors is finally [0°, 27°, 30°, 47°, 51°, 111°, 126°, 219°, 221°, 289°], as shown in Figure 7.

Table 5. The optimal placement with 10 sensors for different cases.

Cases	Ours		Tradition		FBG Placements [°]
	Total Error [mm]	Max Error [mm]	Total Error [mm]	Max Error [mm]	
Case 1	9.87	2.26	62.74	9.42	[0, 20, 27, 29, 30, 50, 126, 221, 222, 290]
Case 2	19.40	4.47	126.54	18.74	[0, 3, 10, 27, 47, 111, 122, 126, 220, 289]
Case 3	28.32	6.62	189.49	27.90	[0, 12, 43, 47, 48, 51, 124, 126, 219, 289]
Case 4	37.58	7.67	237.73	35.24	[39, 40, 125, 127, 215, 216, 217, 219, 223, 311]
Case 5	16.37	3.17	153.08	21.07	[41, 146, 156, 159, 225, 229, 232, 236, 310, 358]
Case 6	19.22	4.95	129.67	19.10	[0, 27, 36, 47, 110, 118, 126, 219, 220, 289]
Case 7	27.67	6.11	191.47	28.41	[0, 1, 2, 21, 23, 24, 49, 53, 221, 290]
Case 8	35.71	8.66	244.18	35.98	[0, 21, 30, 45, 108, 114, 125, 127, 219, 289]
Case 9	17.01	4.21	110.50	16.65	[0, 1, 34, 43, 47, 111, 126, 217, 218, 288]

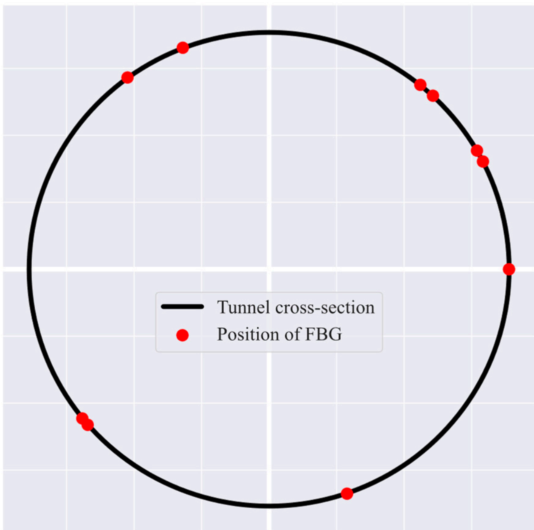

Figure 7. Final FBG sensors placement.

4.3. Sensitivity Analysis of Parameters in WOA-DQN

The speed of WOA-DQN network convergence is influenced by the learning rate, multi-step reward, initial noise, hidden layer in the network, and other hyperparameters. Therefore, the reasonable selection range of those parameters was discussed to optimize the algorithm. The initial algorithm parameters of WOA-DQN are shown in Table 2.

4.3.1. Learning Rate

The learning rates for the experiment were set as 0.1, 0.01, 0.005, 0.002, and 0.0001 to discuss the influence on the WOA-DQN. The change in total reward in each case is shown in Figure 8. Obviously, the algorithm was convergent to the feasible globally optimal solution when the learning rates were set as 0.01, 0.005, 0.002, and 0.001. The agent converged fastest when the learning rate was 0.01. However, the total reward fluctuated the most after convergence. Especially when the learning rate was 0.001, the total reward had less fluctuation after convergence; the robustness was the best with a slower convergence speed than 0.01.

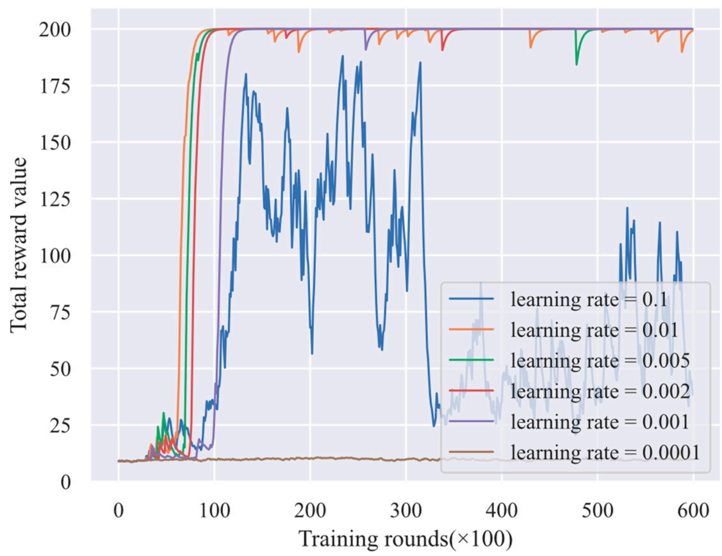

Figure 8. Performance of different learning rates.

4.3.2. Noise Parameter

The initial noise parameters were set as 0.1, 0.3, 0.4, 0.5, 0.6 and 0.9. It was observed that the agent would converge when the initial noise parameter was 0.3 or 0.4, as shown in Figure 9. After a long period of training, the agents would be close to the optimal solution in other cases. However, they would converge to the locally optimal solution or fail to converge finally. Cases where the agent was convergent are shown in Table 6. Therefore, the initial noise parameter is set as 0.4 to get the best performance in the placement optimization.

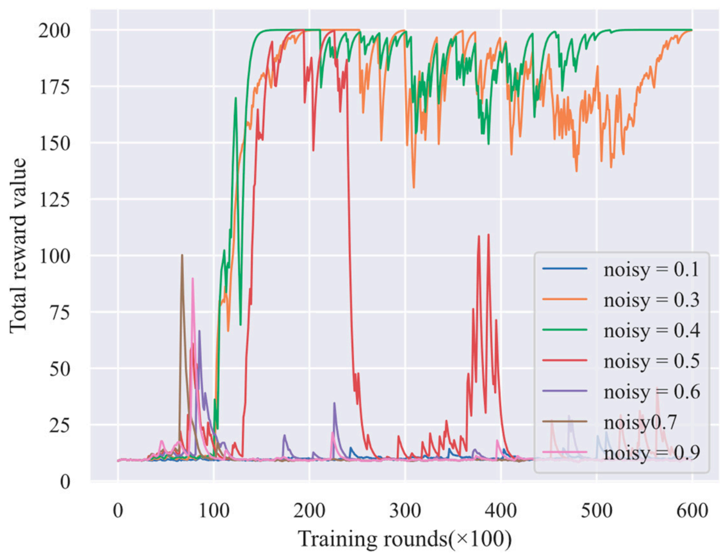

Figure 9. Performance of different initial noise parameters.

Table 6. Performance of different noise parameters.

Noise Parameter	Vault Error [mm]	Left Waist Error [mm]	Right Waist Error [mm]	Total Error [mm]	Required Steps
0.3	15.7947	15.8178	15.6888	293.9521	2827
0.4	15.3458	14.1637	14.1785	294.2378	2678

4.3.3. Hidden Layer

The numbers of hidden layer were 1, 3 and 10, while the numbers of node were 64, 128, and 512 respectively. The compared results of different parameters are illustrated in Figure 10. Both convergence speed and performance are considered, the hidden layer number was set as 3 and the node number was set as 512 respectively.

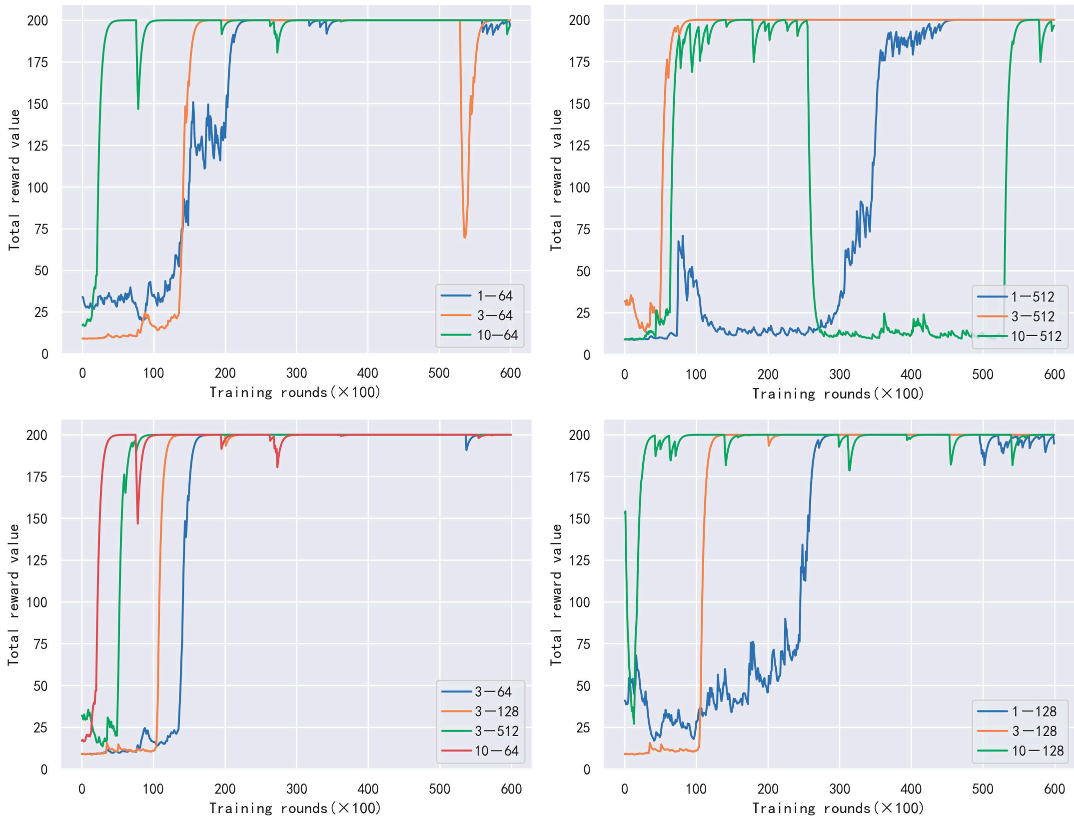

Figure 10. Performance of different hiding layer parameters.

4.3.4. Update Frequency

The update frequency of WOA was set as 5 to balance the influence of parameters in estimated network. After the objective network was updated 5 times by DQN, WOA was used to update the objective network once. The update frequency of DQN was set as 5, 20, 50, 200 and 500, therefore the update frequency of the objective network is 25, 100, 250, 1000 and 2500. By comparing the performance of all cases in Figure 11, it draws the conclusion that the agent has the fastest convergence speed and the best stability when the update frequency is 20.

In summary, the WOA-DQN proposed in this study has good applicability. The experimental results have shown that a hidden layer configuration with 3 layers and 512 neurons per layer can enhance the learning ability of DQN in complex nonlinear relationships, enabling it to better capture complex features in sensor placement problems. This configuration enhances the expressive power and exploration efficiency of the model, helping to more accurately determine the optimal sensor position during the optimization process, while controlling computational complexity to ensure the stability and generalization ability of the learning process. This enables DQN to achieve higher deployment accuracy in complex tunnel monitoring environments. And because a smaller learning rate can ensure smooth parameter updates, avoid oscillations during the convergence process due to large step sizes, and ensure that the model can gradually approach the optimal solution, the learning rate λ is set to 0.001. In order to ensure sufficient exploration intensity in the initial stage and prevent excessive noise from causing the model to converge, the noise parameter is set to 0.4.

Finally, the sensor placement is set at [0°, 27°, 30°, 47°, 51°, 111°, 126°, 219°, 221°, 289°] to get the best accuracy with the learning rate 0.001, the initial noise 0.4, the hidden layer 3–512 and the updated frequency 20. These hyperparameters are used to compare our new algorithm with other algorithms in this study.

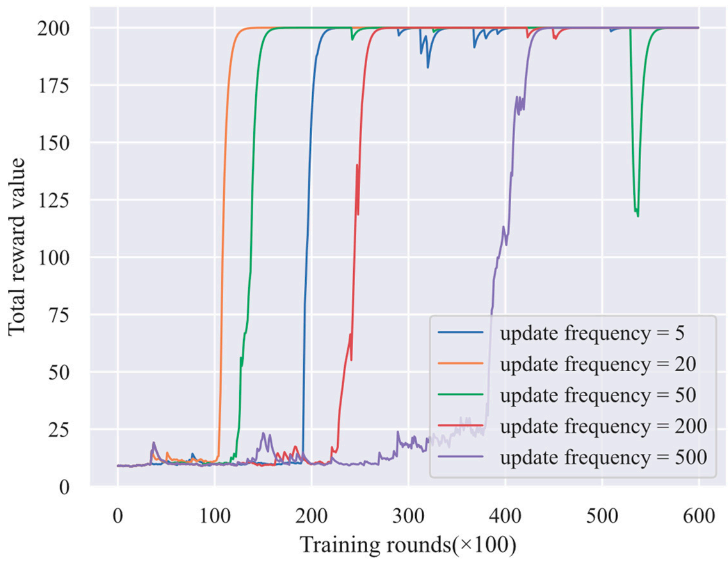

Figure 11. Performance of different updating frequencies.

4.4. Comparison of Different Algorithms for Placement Optimization

The study compared the performance of different algorithms based on the datasets of Case 1, including Noisy DQN, Prioritized DQN, DQN, WOA, and proposed WOA-DQN. The total reward of different algorithms for the FBG placement optimization is shown in Figures 12 and 13. Obviously, it could be seen that the WOA possesses the worst convergence rate and the final result even is unstable. The proposed WOA-DQN performed best. The results verified that it is superior to other methods and suitable for optimizing the placement of FBG sensors when the WOA-DQN was tested further in other cases. When the number of training rounds is 340, the total reward value of WOA-DQN under case 1, 3, 4, 5, 6, 7, 8, 9 is best shown in Table 7. As can be seen from the Figure 13e, WOA-DQN quickly outperforms other methods such as Noisy DQN and Prioritized DQN during initial training (first 100–200 rounds), and eventually shows a more consistent and higher total return after long-term training. In contrast, the traditional DQN converges slowly and behaves erringly over some training rounds. Noisy DQN and Prioritized DQN, although close to WOA-DQN at some stages, are still slightly inferior overall.

Table 7. The total reward value.

Case	WOA-DQN	Noisy DQN (Fortunato, M et al. [34])	Prioritized DQN (Schaul, T et al. [35])	DQN (Luo Lei et al. [23])	WOA (Mirjalili et al. [28])
	Total Reward Value				
Case1	**158**	152	136	138	116
Case2	143	150	**158**	133	108
Case3	**158**	**158**	144	148	103
Case4	**167**	156	158	135	119
Case5	**155**	150	154	135	89
Case6	**153**	148	136	139	103
Case7	**159**	156	148	148	102
Case8	**148**	148	148	146	128
Case9	**156**	145	146	140	58

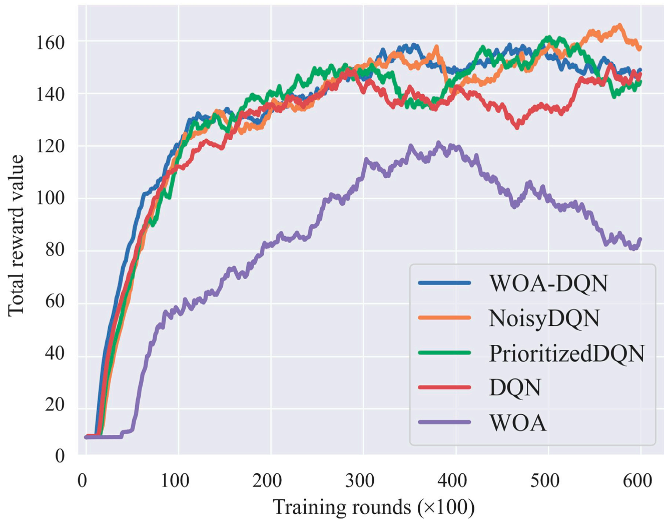

Figure 12. Performance of 5 algorithms in Case 1.

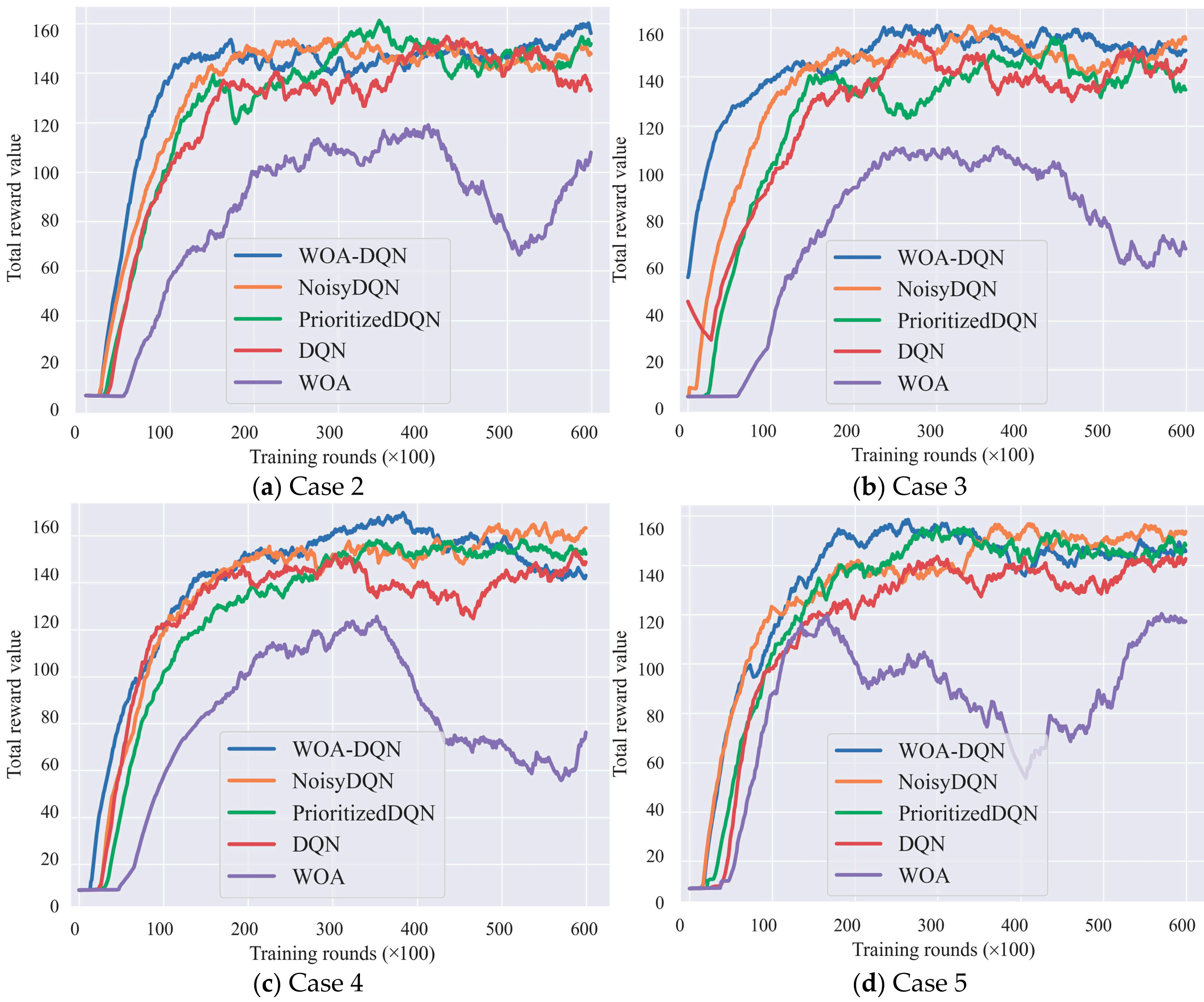

Figure 13. *Cont.*

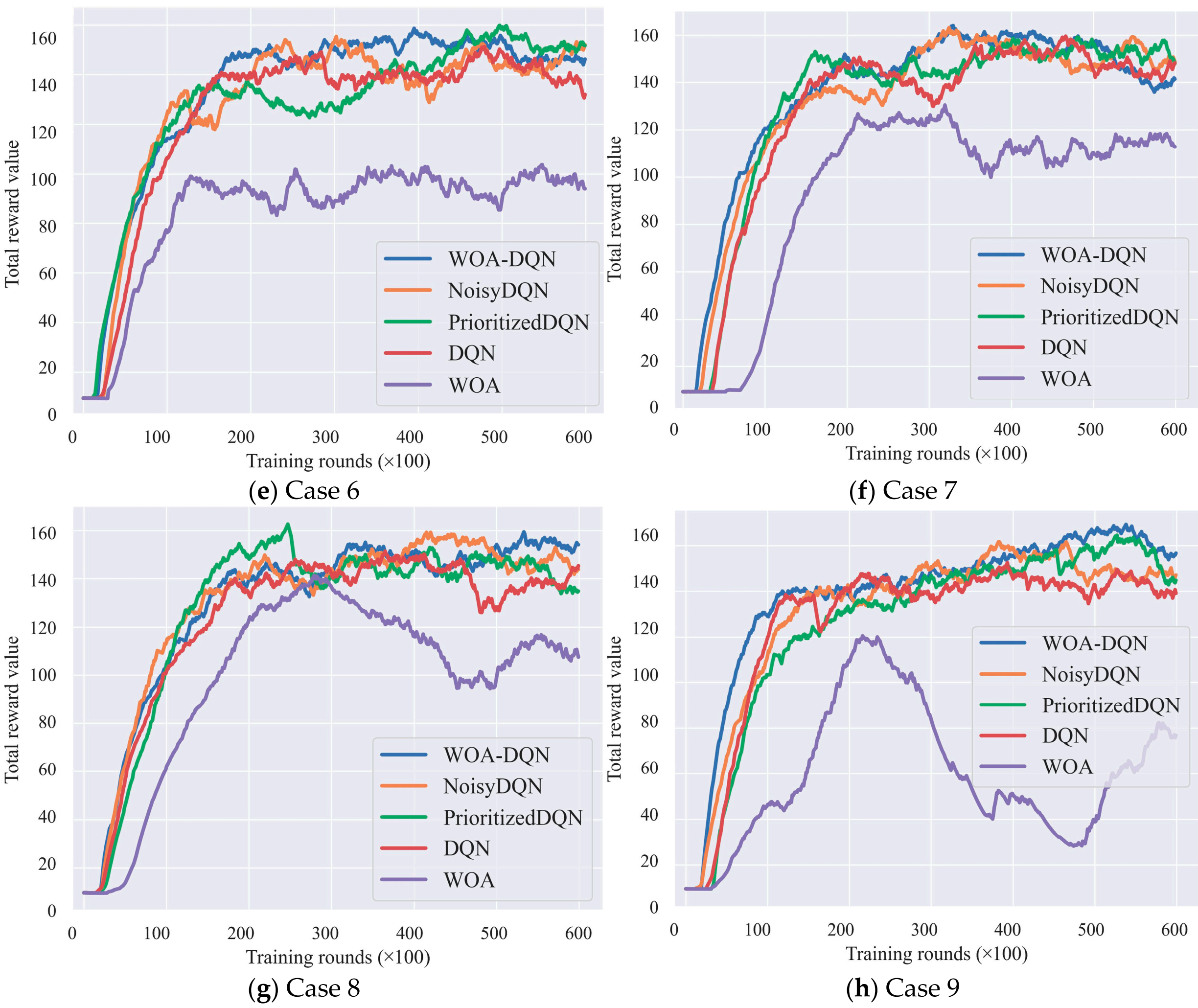

Figure 13. Performance of 5 algorithms in all Cases.

5. Conclusions

Based the swarm intelligence algorithm, the proposed WOA-DQN algorithm has better ability of exploration in the deep reinforcement learning algorithm. It improves the capacity of agent to explore the optimal solution equipped with a loop structure of SIA-DRL. The datasets of the tunnel were used to train the WOA-DQN model. By discussion above, the following conclusions could be drawn:

(1) After numerous experiments, the optimal sensor placement in the tunnel for various cases are obtained. The FBG sensors was finally set as 10 and at [0°, 27°, 30°, 47°, 51°, 111°, 126°, 219°, 221°, 289°] to detect the precise deformation of the tunnel. Compared with the traditional layout, the optimized placement obtained has better performance in all cases.

(2) The results demonstrate the efficacy of the WOA-DQN in resolving the optimal placement problem of FBG sensors in the tunnel. It provides theoretical basis to the placement of sensors for structural health monitoring and increases the average reconstruction accuracy with the max error 8.66 mm.

(3) It was found that the proposed WOA-DQN algorithm has the fastest convergence speed and the best stability to obtain the optimal sensor placement with the learning rate 0.001, the initial noise 0.4, the hidden layer 3–512 and the updated frequency 20.

(4) Additionally, the experiments show that the improved algorithm is superior to other methods including Noisy DQN, Prioritized DQN, DQN and WOA, and more suitable for optimizing the placement of FBG sensors.

To sum up, the method in this study is mainly applicable to shield tunnels, and other types of tunnels are not taken into account. The placement of sensors in other types of tunnels can be explored in the future. In large-scale tunnel deformation monitoring, this study relies on a large number of numerical model data for training and optimization, so it faces the problems of high computational complexity and insufficient real-time processing ability, which makes it difficult to adjust the sensor position in real time during the monitoring process. Future research can solve these problems by improving computing speed, optimizing sensor layout algorithms, and developing real-time monitoring systems, thereby improving the accuracy and efficiency of monitoring.

Author Contributions: Conceptualization, J.L.; Methodology, M.S.; Software, H.S.; Validation, L.W.; Formal Analysis, M.S.; Investigation, M.S.; Resources, J.L.; Data Curation, K.W.; Writing—Original Draft Preparation, J.L. and M.S.; Writing—Review & Editing, J.L. and M.S.; Visualization, H.S.; Supervision, W.P.; Project Administration, J.L.; Funding Acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Major Science and Technology Special Project of Xinjiang Uygur Autonomous Region [No.2020A03003-4].

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the restriction of privacy.

Conflicts of Interest: All authors are employed by Research and Development Center on Tunnel and Underground Space Technology and CCCC Second Highway Consultant Co., Ltd. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Tan, X.; Chen, W.; Wu, G.; Wang, L.; Yang, J. A structural health monitoring system for data analysis of segment joint opening in an underwater shield tunnel. *Struct. Health Monit.* **2020**, *19*, 1032–1050. [CrossRef]
2. Liu, W.; Wu, X.; Zhang, L.; Wang, Y.; Teng, J. Sensitivity analysis of structural health risk in operational tunnels. *Autom. Constr.* **2018**, *94*, 135–153. [CrossRef]
3. Yang, J.P.; Chen, W.Z.; Li, M.; Tan, J.X.; Yu, J.X. Structural health monitoring and analysis of an underwater TBM tunnel. *Tunn. Undergr. Space Technol.* **2018**, *82*, 235–247. [CrossRef]
4. Pan, J.B.; Ye, J.; Ai, H.J.; Wang, J.W.; Wan, Y. Parameter optimization of a pure electric sweeper dust port by a backpropagation neural network combined with a whale algorithm. *Mech. Sci.* **2023**, *14*, 47–60. [CrossRef]
5. Huang, Z.; Fu, H.; Chen, W.; Zhang, J.B.; Li, J.; Huang, H.W. Damage detection and quantitative analysis of shield tunnel structure. *Autom. Constr.* **2018**, *94*, 303–316. [CrossRef]
6. Xing, R.; Jiang, S.; Xu, P. Long-term temperature monitoring of tunnel in high-cold and high-altitude area using distributed temperature monitoring system. *Measurement* **2017**, *95*, 456–464. [CrossRef]
7. Duan, C.; Hu, M.; Zhang, H. Comparison of ARIMA and LSTM in Predicting Structural Deformation of Tunnels during Operation Period. *Electronics* **2023**, *8*, 104. [CrossRef]
8. Simeoni, L.; Zanei, L. A method for estimating the accuracy of tunnel convergence measurements using tape distometers. *Int. J. Rock Mech. Min. Sci.* **2009**, *46*, 796–802. [CrossRef]
9. Puente, I.; González-Jorge, H.; Martínez-Sánchez, J.; Arias, P. Automatic detection of road tunnel luminaires using a mobile LiDAR system. *Measurement* **2014**, *47*, 569–575. [CrossRef]
10. Wang, T.T.; Jaw, J.J.; Chang, Y.H.; Jeng, F.S. Application and validation of profile-image method for measuring deformation of tunnel wall. *Tunn. Undergr. Space Technol.* **2009**, *24*, 136–147. [CrossRef]
11. Luo, Y.; Chen, J.; Xi, W.; Zhao, P.; Li, J.; Qiao, X.; Liu, Q. Application of a total station with RDM to monitor tunnel displacement. *J. Perform. Constr. Facil.* **2017**, *31*, 04017030. [CrossRef]
12. Yüksel, K.; Kinet, D.; Moeyaert, V.; Kouroussis, G.; Caucheteur, C. Railway monitoring system using optical fiber grating accelerometers. *Smart Mater. Struct.* **2018**, *27*, 105033. [CrossRef]
13. Minardo, A.; Catalano, E.; Coscetta, A.; Porfido, C.; Maio, C.; Bernini, R.; Zeni, L.; Aiello, G.; Giordano, G.; Lapenna, V.; et al. Distributed fiber optic sensors for the monitoring of a tunnel crossing a land-slide. *Remote Sens.* **2018**, *10*, 1291. [CrossRef]
14. Feng, X.; Han, Y.; Wang, Z.; Li, H. Structural performance monitoring of buried pipelines using distributed fiber optic sensors. *J. Civ. Struct. Health Monit.* **2018**, *8*, 509–516. [CrossRef]
15. Kammer, D.C.; Tinker, M.L. Optimal placement of triaxial accelerometers for modal vibration tests. *Mech. Syst. Signal Process.* **2004**, *18*, 29–41. [CrossRef]

16. Lai, J.; Qiu, J.; Fan, H.; Zhang, Q.; Hu, Z.; Wang, J.; Chen, J. Fiber Bragg grating sensors-based in situ monitoring and safety assessment of loess tunnel. *J. Sens.* **2016**, *2016*, 8658290. [CrossRef]
17. Xia, N.; Wen, X.; Du, H.; Xu, C. Study on hunting deployment of underwater sensors based on level set method. *Chin. J. Comput.* **2021**, *44*, 1024–1036. [CrossRef]
18. Li, M.; Shao, G.; Han, G.; Yu, B.; Yang, T.; Hu, Z. Design and placement path of optical fiber monitoring system for shield tunnel. *J. Wuhan Univ. Technol.* **2022**, *44*, 78–83. [CrossRef]
19. Chakraborty, S.; Chakraborty, S. A systematic review of the whale optimization algorithm: Theoretical foundation, improvements, and hybridizations. *Arch. Comput. Methods Eng.* **2022**, *29*, 89–112. [CrossRef]
20. Sun, J.; Wu, Z.; Li, X.; Wang, Y.; Zhang, P.; Liu, F.; Chen, H.; Xu, J.; Li, G.; Zhao, R.; et al. An Improved Whale Optimization Algorithm Based on Nonlinear Parameters and Feedback Mechanism. *Int. J. Comput. Intell. Syst.* **2022**, *15*, 38–52. Available online: https://link.springer.com/article/10.1007/s44196-022-00092-7 (accessed on 14 September 2024). [CrossRef]
21. Habib, M.; Abbas, G.; Ghaffari, A.; Jumani, T.; Bhutto, A.; Mirsaeidi, S.M.; Ahmed, E. Improved whale optimization algorithm for transient response, robustness, and stability enhancement of an automatic voltage regulator system. *Energies* **2022**, *15*, 5037. [CrossRef]
22. Li, Y.; Pei, W.; Zhang, Q. Improved whale optimization algorithm based on hybrid strategy and its application in location selection for electric vehicle charging stations. *Energies* **2022**, *15*, 7035. [CrossRef]
23. Luo, L.; Zhao, N.; Zhu, Y.; Sun, Y. A* guiding DQN algorithm for automated guided vehicle pathfinding problem of robotic mobile fulfillment systems. *Comput. Ind. Eng.* **2023**, *178*, 109112. [CrossRef]
24. Sun, W.; Zhang, W.; Ma, N.; Min, J. A Multi-Branch DQN-Based transponder resource allocation approach for satellite communications. *Electronics* **2023**, *12*, 916. [CrossRef]
25. Meng, W.; Zheng, Q.; Shi, Y.; Pan, G. An Off-Policy trust region policy optimization method with monotonic improvement guarantee for deep reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 2223–2235. [CrossRef]
26. Liu, H.; Wu, Y.; Sun, F.; Li, J.; Zhang, M. Extreme trust region policy optimization for active object recognition. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2253–2258. [CrossRef]
27. Huang, R.; He, H. Naturalistic data-driven and emission reduction-conscious energy management for hybrid electric vehicle based on improved soft actor-critic algorithm. *J. Power Sources* **2023**, *559*, 232648. [CrossRef]
28. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]
29. He, W.; Wang, J.; Liu, Y.; Qin, Z.; Sun, C.; You, H.; Wei, X.; Liu, Y. A novel symmetrical peak fitting method based on improved WOA algorithm for the analysis of microchip electrophoresis signals. *Symmetry* **2022**, *14*, 2603. [CrossRef]
30. Lou, J.; Li, L. Dynamic economic environmental dispatch of power system using an improved whale algorithm. *J. Phys. Conf. Ser.* **2022**, *2409*, 012017. [CrossRef]
31. Parent, F.; Loranger, S.; Mandal, K.K.; Iezzi, V.; Lapointe, J.; Boisvert, J.; Baiad, M.; Kadoury, S.; Kashyap, S. Enhancement of accuracy in shape sensing of surgical needles using optical frequency domain reflectometry in optical fibers. *Biomed. Opt. Express* **2017**, *8*, 2210–2221. [CrossRef] [PubMed]
32. Zhang, A.P.; Gao, S.; Yan, G.; Bai, Y.B. Advances in optical fiber Bragg grating sensor technologies. *Photonic Sens.* **2012**, *2*, 1–13. [CrossRef]
33. Zheng, K.; Liu, J.; Zeng, W. Neural Architecture Search with Representation Mutual Information. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022. Available online: https://openaccess.thecvf.com/content/CVPR2022/html/Zheng_Neural_Architecture_Search_With_Representation_Mutual_Information_CVPR_2022_paper.html (accessed on 14 September 2024).
34. Fortunato, M.; Azar, M.G.; Piot, B.; Menick, J.; Osband, I.; Graves, A.; Mnih, V.; Munos, R.; Hassabis, D.; Pietquin, O.; et al. Noisy networks for exploration. *arXiv* **2017**. [CrossRef]
35. Schaul, T.; Quan, J.; Antonoglou, I. Prioritized Experience Replay. *arXiv* **2016**, arXiv:arXiv:1511.05952. Available online: https://arxiv.org/abs/1511.05952 (accessed on 14 September 2024).