

Article

Pose Error Compensation of Drilling and Anchoring Arm Based on Improved DDPG Algorithm

Xuan Dong ^{1,2,3,*} , Jianjian Yang ^{1,2}, Zhaowei Li ^{1,2} , Guoyong Wang ¹ and Haifeng Han ^{1,2} ¹ Inner Mongolia Research Institute, China University of Mining and Technology (Beijing), Ordos 017004, China; yangjiannedved@163.com (J.Y.)² School of Mechanical and Electrical Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China³ Shaanxi Blower (Group) Co., Ltd., Xi'an 710082, China

* Correspondence: kzlcvv02603879@163.com

Abstract

Aiming at the engineering problems of composite roll, pitch, and yaw pose errors of the roadheader body induced by floor undulation and geological variation, as well as insufficient anchoring accuracy caused by the incapability of traditional 1–3-degree-of-freedom (DOF) drilling–anchoring arms in dynamic error compensation during coal mine roadway excavation and bolting, this paper proposes an inverse kinematics solving method for a 5-DOF drilling–anchoring arm based on an improved Deep Deterministic Policy Gradient (DDPG) algorithm. Firstly, the modified Denavit–Hartenberg (MDH) approach is adopted to establish a full-link kinematic model incorporating body pose errors, where the drill rod length, mounting offset, and world coordinate transformation are fully considered. Secondly, an Actor–Critic dual-network architecture tailored for drilling and anchoring tasks is constructed with an 11-dimensional state space and a 5-dimensional action space. The coupling optimization between body pose errors and joint adjustments is realized by designing a hierarchical gradient reward function, a dynamic noise decay exploration strategy, and an optimal state restart mechanism. Finally, 1000 episodes of training and verification are carried out on a Python 3.10 simulation platform. The simulation results reveal that the average end-effector position error of the improved algorithm reaches 0.71 mm, and the deflection angle toward the roof is less than 1°, which outperforms the specified industrial standard. The proposed method realizes real-time compensation for dynamic body pose errors and provides crucial technical support for intelligent excavation and anchoring in underground coal mines.

Keywords: drilling and anchoring arm; improved DDPG algorithm; inverse kinematics; pose error compensation; coal mine roadway support



Academic Editor: Arkadiusz Gola

Received: 16 May 2026

Revised: 19 June 2026

Accepted: 7 July 2026

Published: 27 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

During the coal mine roadway excavation process, roof safety is the core prerequisite for ensuring smooth construction progress. As a key part of roof support, the positional and directional accuracy of roof bolt drilling directly determines the support quality and the long-term stability of the roadway. The roadheader, as the core equipment for roadway excavation, simultaneously undertakes the roof support task with the roof drill mounted on its head while performing excavation. However, due to factors such as floor undulations, geological variations, and machine body vibration, the roadheader body is prone to composite pose changes in pitch, roll, and yaw angles. Traditional drilling and anchoring

manipulators are limited by roadway space layout and development costs, often featuring 1–3-degrees-of-freedom (DOF) structures [1]. These structures are only capable of fixed trajectory operations with minor adjustment capabilities parallel to the working face, lacking the ability to compensate for real-time complex pose errors of the machine body. This leads to excessive deviations in bolt hole perpendicularity and positional accuracy, reducing the load-bearing capacity of the support structure. A schematic diagram of its operation is shown in Figure 1. When deviations exceed standards, rework and hole patching are required, occupying excavation time and affecting production efficiency. Therefore, constructing a multi-DOF drilling and anchoring arm structure capable of compensating for body pose errors and designing an efficient pose compensation algorithm is of significant engineering importance for improving the safety and intelligence level of underground excavation and anchoring operations.

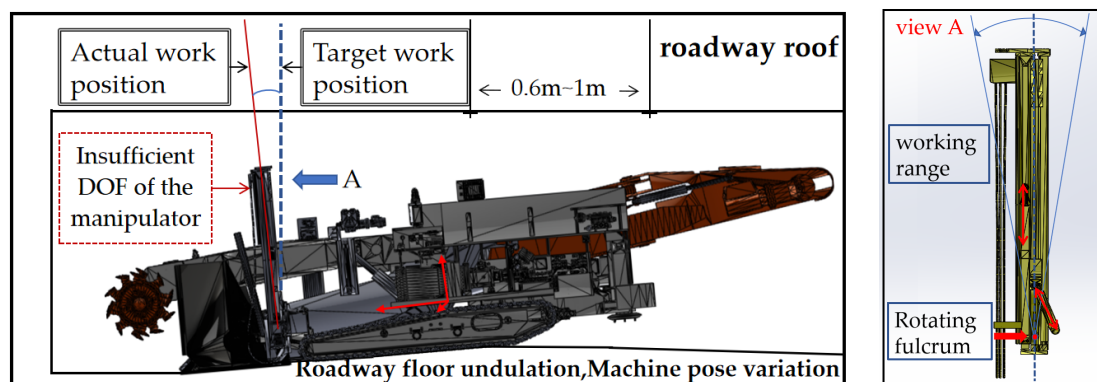


Figure 1. Drilling and anchoring arm structure and operation error schematic diagram.

To achieve dynamic compensation for machine body pose errors, the core objective is to adjust the motion of each joint of the drilling and anchoring arm to counteract any end-effector pose deviations caused by machine body pose variations. This essentially boils down to the inverse kinematics problem of the drilling and anchoring arm—namely, solving the corresponding joint adjustment amounts based on the target end-effector pose and current machine body pose errors. Therefore, the accuracy, convergence speed, and anti-interference capability of the inverse kinematics algorithm directly determine the effectiveness of pose error compensation, serving as a core technology for high-precision anchoring operations. Traditional classical methods, such as analytical geometry [2], iterative numerical methods [3], and Jacobian pseudoinverse [4], rely heavily on precise kinematic models, suffering from complex solution logic, limited applicable working conditions, susceptibility to local optima, and sensitivity to initial values. These methods struggle to effectively address the strongly coupled nonlinear problem between dynamic machine body pose errors and joint adjustments in underground scenarios. By contrast, agent-based algorithms enable adaptive error optimization through real-time interaction with the environment, exhibiting superior dynamic adaptability, real-time compensation capability, and engineering robustness.

In recent years, many researchers have explored inverse kinematics solutions and error optimization for robotic arms in various scenarios, proposing multiple intelligent optimization methods. Zhang et al. [5] developed a general inverse kinematics solution for assembly robots using an improved Particle Swarm Optimization algorithm, enhancing solution accuracy and convergence speed through a nonlinear weight strategy and constraint integration via penalty terms. Ning et al. [6] presented a multi-objective optimization framework for the inverse kinematics of upper-limb exoskeleton rehabilitation robots, aiming to achieve more human-like motion trajectories by simultaneously optimizing

end-effector position, joint comfort, and energy consumption. Zhao et al. [7] introduced a deep reinforcement learning-based inverse kinematics and control method (MAPPO-IK) for 6-DOF manipulators, facilitating accurate pose tracking in complex dynamic scenarios through a dual-distance (Euclidean and cosine) reward design. Luo et al. [8] proposed a multi-objective full-parameter optimization particle swarm optimization algorithm, which designs a comprehensive objective function considering end-effector position, orientation, and joint changes, and optimizes inertia weight, asynchronous learning factors, and time factors to address the problems of poor accuracy, low efficiency, and instability in manipulator inverse kinematics solutions. Naranjo-Campos et al. [9] proposed an apprenticeship learning method via inverse reinforcement learning based on expert-trajectory features combined with a reverse discount mechanism and integrated with the Proximal Policy Optimization algorithm, aiming to infer reward functions more efficiently from expert demonstrations for policy learning in complex robotic manipulation tasks within continuous state spaces. Bai et al. [10] proposed a fruit fly optimization algorithm optimized BP neural network algorithm, which effectively improves the shortcomings of slow convergence, tendency to fall into local minima, and large output errors of traditional BP neural networks in solving robot inverse kinematics by using the FOA (Fruit Fly Optimization Algorithm) to optimize the network's initial weights and thresholds. Sekkat et al. [11] proposed a vision-based robotic arm control algorithm combining the Deep Deterministic Policy Gradient (DDPG) algorithm with inverse kinematics, providing a solution for multi-DOF manipulators that does not require precise kinematic models and outperforms traditional methods in autonomous object grasping. Yiyang et al. [12] proposed an improved Particle Swarm Optimization algorithm suitable for general robots that do not meet the Pieper criterion, which introduces a nonlinear dynamic inertia weight based on the concept of similarity and a multi-particle swarm strategy with immigration operators to enhance convergence accuracy, speed, and stability when solving high-dimensional inverse kinematics problems. Danaci et al. [13] proposed a general inverse kinematics solution method for serial robot manipulators based on an improved Particle Swarm Optimization algorithm and POSIX threads parallel implementation, which converges to the complete end-effector pose and significantly improves computational efficiency. Ostermeier et al. [14] proposed an automatic geometric decomposition method for the analytical inverse kinematics of manipulators, which enables, for the first time, the derivation and computation of analytical inverse kinematics in less than a millisecond by decomposing a given kinematic chain online into a series of pre-solved geometric subproblems. Ning et al. [15] developed an inverse kinematics solution for a 6-DOF variable stiffness manipulator based on the product of exponential formula and Paden–Kahan subproblems, which simplifies the analysis for an (RRR)-R-(RR) configuration with coupled position and orientation by avoiding the establishment of local coordinate frames for each link and complex spatial geometric analysis. Seyyedabbasi [16] investigated the use of metaheuristic algorithms, including the Sand Cat Swarm Optimization algorithm, to solve the inverse kinematics of a PUMA 560 robot arm by transforming the problem of finding joint angles into a cost function optimization, demonstrating its competitiveness in terms of convergence accuracy and speed compared to algorithms like PSO (Particle Swarm Optimization) and GWO (Grey Wolf Optimizer). Diaz Posada et al. [17] proposed external sensor compensation and compliance model-based compensation methods for robotic drilling tasks in the aviation industry, improving positioning accuracy through systematic error source analysis. Lei et al. [18] proposed a hybrid visual servo control strategy combining PBVS (position-based visual servoing) and IBVS (image-based visual servoing) for coal mine drilling and anchoring robots, effectively improving drilling alignment accuracy and operational efficiency.

However, from the perspective of pose error compensation for drilling and anchoring arms in coal mine roadways, several limitations still exist in current studies. First, most inverse kinematics methods are mainly developed for fixed-base industrial manipulators, assembly robots, rehabilitation robots, or robotic drilling systems in aviation, where the manipulator base is usually assumed to be fixed or known. For a drilling and anchoring arm mounted on a roadheader, floor undulation, geological variation, and cutting-induced vibration may cause composite roll, pitch, and yaw errors of the machine body. Such machine-body pose disturbances directly change the spatial relationship between the base coordinate system of the drilling and anchoring arm and the world coordinate system. However, few existing studies have incorporated this type of mobile-equipment base disturbance into the full-link kinematic model, making it difficult to accurately describe the coupling relationship between machine-body pose variation and drill-end pose deviation. Second, roof bolting in coal mine roadways requires not only that the drill end reaches the designed bolt-hole position, but that the drilling direction remains as perpendicular as possible to the roof. Existing intelligent optimization or deep-reinforcement-learning-based inverse kinematics studies have mainly taken end-effector position error minimization as the primary objective. Even when orientation error is considered, it is usually designed for general robotic grasping, assembly, or trajectory tracking tasks, rather than for the dual constraints of bolt-hole positioning accuracy and drilling perpendicularity. Third, the existing DDPG-related studies have mainly focused on convergence performance in general continuous-control tasks and have not been specifically adapted to underground drilling and anchoring operations characterized by limited roadway space, dynamic machine-body disturbance, millimeter-level positioning accuracy, and angular deviation constraints. Therefore, establishing a drilling and anchoring arm kinematic model considering dynamic machine-body pose errors and designing a reinforcement-learning-based inverse kinematics algorithm that can simultaneously optimize end position and drilling direction constitute the core research gap to be addressed for intelligent drilling and anchoring equipment in underground coal mines.

Among various intelligent optimization algorithms, deep reinforcement learning enables adaptive learning through real-time interaction with the environment, making it particularly suitable for decision-making problems in continuous action spaces under dynamic working conditions. As a classic deep reinforcement learning algorithm for continuous action spaces, the Deep Deterministic Policy Gradient (DDPG) algorithm integrates the feature extraction capability of deep learning and the decision-making capability of reinforcement learning. Compared with swarm intelligence algorithms such as PSO and FOA, it exhibits superior real-time performance and adaptability, making it more suitable for the real-time compensation scenario of dynamically changing machine body poses in underground coal mines. Additionally, compared with algorithms like PPO (Proximal Policy Optimization) [19] and MAPPO (Multi-Agent Proximal Policy Optimization) [20], the DDPG algorithm features a simpler structure, facilitating engineering implementation on underground embedded controllers. Nevertheless, the existing DDPG-based inverse kinematics research for manipulators has failed to design optimized strategies tailored to the dual high-precision requirements for position and attitude in drilling and anchoring operations, resulting in insufficient convergence accuracy and poor engineering adaptability.

However, this study does not simply apply the standard DDPG algorithm to the inverse kinematics solution of the drilling and anchoring arm. Instead, the algorithm is adapted according to the specific characteristics of underground roadway drilling and anchoring operations. Specifically, a 5-DOF drilling and anchoring arm with an $R \perp R \perp P \perp R \perp R$ joint configuration is first adopted according to the limited roadway space and the requirement for end-effector pose adjustment, thereby providing the mechanical ba-

sis for position and direction compensation. Then, a full-link kinematic model is established based on the modified Denavit–Hartenberg method. In this model, the roll, pitch, and yaw pose errors of the roadheader body are introduced, and the drill-rod length, mounting offset of the drilling arm, and coordinate transformation from the body coordinate system to the world coordinate system are also considered. In this way, the algorithm training environment can reflect the influence of dynamic machine-body disturbances on the end-effector pose under underground working conditions. On this basis, the inverse kinematics problem of the drilling and anchoring arm is reformulated as a reinforcement learning decision-making problem in a continuous action space. An eleven-dimensional state space is constructed using five joint states, three end-effector position error components, and three end-effector direction error components, while the adjustment values of the five joints are defined as a five-dimensional action space.

To meet the dual requirements of bolt-hole positioning accuracy and drilling perpendicularity in roof bolting operations, a hierarchical gradient reward function consisting of position error reward, direction error reward, and high-precision reward is further designed to guide the agent to optimize both the drill-end position and the drilling direction. To improve exploration efficiency and convergence stability during training, a dynamic noise decay strategy is introduced, allowing the agent to maintain sufficient exploration ability in the early training stages and to gradually shift to high-precision convergence in the later stages. Meanwhile, an optimal-state restart mechanism is adopted, enabling the agent to restart exploration from historically better states with a certain probability, thereby reducing the risk of falling into local optima. Finally, random roll, pitch, and yaw disturbances of the machine body are set in a Python simulation platform, and the proposed method is compared with the standard DDPG algorithm in terms of reward convergence, end-effector position error, and direction deviation. Through the above modeling and algorithmic design, this study realizes coupled optimization between dynamic machine-body pose errors and joint adjustment values of the drilling and anchoring arm, providing a feasible inverse kinematics solution for real-time pose error compensation in underground coal mine drilling and anchoring operations.

To address the above research gap, this paper proposes a pose error compensation method for a 5-DOF drilling and anchoring arm based on an improved DDPG algorithm. The main scientific contributions of this study are summarized as follows:

- (1) A full-link kinematic model of a 5-DOF drilling and anchoring arm considering roadheader body pose errors is established. Different from conventional fixed-base manipulator models, the proposed model incorporates the roll, pitch, and yaw errors of the machine body, drill-rod length, mounting offset of the drilling arm, and world-coordinate transformation into the MDH-based kinematic modeling process, thereby revealing the mapping relationship between dynamic machine-body pose disturbances and drill-end pose deviations.
- (2) An improved DDPG-based inverse kinematics framework is proposed for coal mine drilling and anchoring operations. By constructing an eleven-dimensional state space consisting of joint states, end-effector position errors, and end-effector direction errors, together with a continuous five-dimensional action space corresponding to the five joint adjustment values, the proposed framework realizes coupled optimization between the machine-body pose errors and joint compensation values of the drilling and anchoring arm.
- (3) A position–direction coordinated reward mechanism is designed according to the accuracy requirements of roof bolting. The reward mechanism simultaneously considers end-effector position error, drilling direction error, and high-precision convergence reward, enabling the algorithm not only to reduce the drill-end position deviation but

to constrain the drilling direction to approach the roof normal direction, which better satisfies the engineering requirements of coal mine roof support.

- (4) A dynamic noise decay strategy and an optimal-state restart mechanism are introduced to improve the exploration efficiency and convergence stability of the improved DDPG algorithm when solving the strongly coupled nonlinear inverse kinematics problem. The simulation results show that the proposed method can achieve millimeter-level end-effector position compensation and small-angle direction deviation control under random machine-body pose disturbances, providing a reproducible algorithmic framework for the real-time pose error compensation of underground drilling and anchoring arms.

The remainder of this paper is organized as follows. First, the mechanical structure design of the 5-DOF drilling and anchoring arm is elaborated in detail, and a forward kinematic model incorporating machine body pose errors is established using the improved MDH method. Second, an improved DDPG-based inverse kinematics solution algorithm is constructed, including a definition of the state–action space, optimization of network architecture, and design of core reward function and exploration strategy. Third, a simulation platform is built to verify the convergence performance and pose compensation accuracy of the improved algorithm by comparing it with the standard DDPG algorithm. Finally, the research conclusions are summarized, the limitations of the current study are analyzed, and future prospects for engineering implementation and optimization are discussed.

2. Materials and Methods

2.1. Structure of the 5-DOF Drilling and Anchoring Arm

The width of underground coal mine roadways is typically 4.5–5.5 m, with a height of 3–4 m, resulting in limited available space. Meanwhile, the installation space on the body of the drilling and anchoring integrated machine is constrained, requiring error compensation to be implemented within a compact mechanical structure. The drilling and anchoring manipulator mounted on the roadheader body is mostly a 1–3-DOF mechanism. Facing body pose changes, it often results in bolts not being drilled vertically into the roof. To adapt to multi-angle changes of the floor and the body, a 5-DOF drilling and anchoring mechanism is adopted at the front of the roadheader body. Considering the functional requirement of domestic drilling–bolting arms, which primarily focus on the end-effector pose adjustment, a spherical joint configuration scheme of $R\perp R\perp R$ is initially adopted. To meet the operational demands of medium and high cross-section roadways, a prismatic joint (P) is incorporated to enhance the working height and coverage range. This leads to the determination of the configuration for the first four joints as $R\perp R\perp P\perp R$. After a comprehensive multi-dimensional consideration of factors, including roof support posture constraints, underground spatial layout, pose error compensation, engineering reliability, and control system compatibility, a five-degree-of-freedom (5-DOF) drilling–bolting arm is finally selected. Its joint configuration is $R\perp R\perp P\perp R\perp R$, as shown in Figure 2. Joint 1 enables horizontal rotation of the drill arm, adjusting the working horizontal direction. Joint 2 enables pitch motion of the drill arm in the vertical plane, controlling the overall lifting angle. Joint 3 is responsible for adjusting the length of the main connecting rod of the drill arm, extending the working radius, with an adjustable length of $d_3 = 1000$ mm (see Figure 3). Joint 4 adjusts the horizontal posture of the end drill to compensate for any body roll error. Joint 5 controls the pitch angle of the drill end, ensuring the drilling operation does not deviate from the vertical direction. The end link is connected to the roof bolter, and the rotation of the drill around its axis is driven independently by the power head, which is not included in the manipulator control. The drill arm adopts a serial “base-link-end effector” structure. The modified Denavit–Hartenberg method (MDH)

method is used to establish the kinematic model of the drill arm to accurately describe the geometric relationships between joints. The specific parameters are shown in Table 1, where α_{i-1} is the link twist angle, a_{i-1} is the link length, θ_i is the joint angle, and d_i is the joint offset.

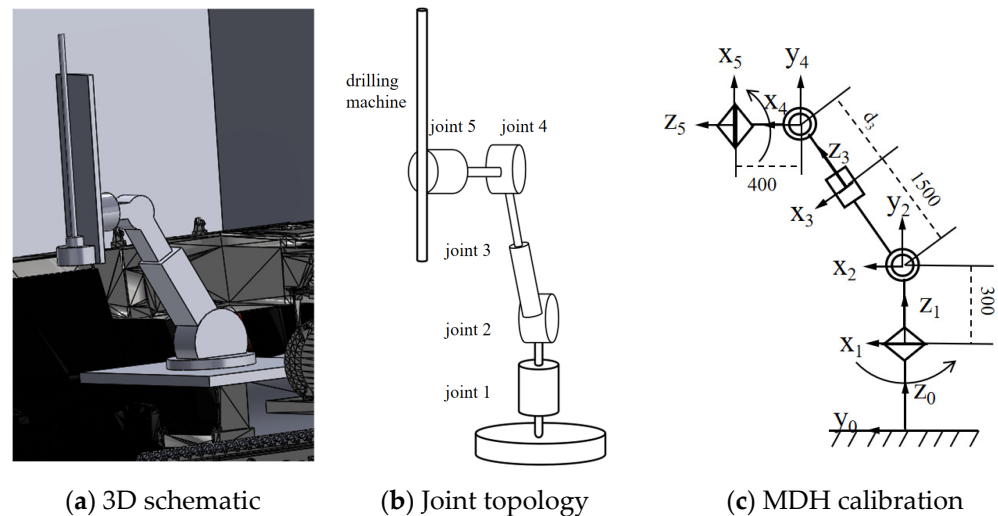


Figure 2. Schematic diagram of drill arm mechanism and the MDH coordinate system.

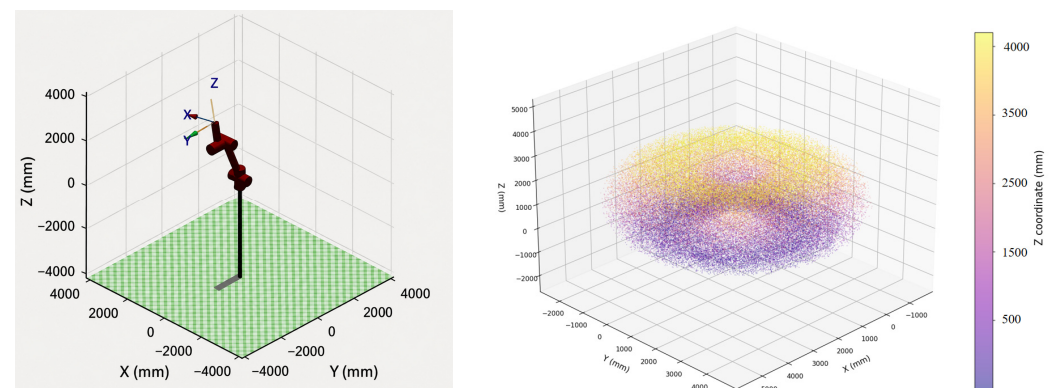


Figure 3. Visualization model and workspace scatter plot.

Table 1. MDH Parameters Table.

i	α_{i-1}	a_{i-1}	θ_i (°)	d_i (mm)
1	0	0	q_1	300
2	$\pi/2$	0	q_2	0
3	$-\pi/2$	0	0	$d_3 + 1500$
4	$\pi/2$	0	q_4	0
5	$\pi/2$	0	q_5	400

The parameters are as follows: q_1, q_2, q_4, q_5 are rotational joint angle variations (unit: °); d_3 is the prismatic joint displacement variation (unit: mm); the joint's reference offset is 1500 mm; and the actual offset is the sum of the reference value and the input displacement. Because the rotation axis (Z-axis) of Joint 1 is collinear with the rotation center axis of Joint 2, and their rotation centers coincide, the joint offsets of Joint 1 and Joint 2 are combined into $d_1 = 300$ mm. The model is visualized and modeled in MATLAB (R2023b) using the Robotics Toolbox, as shown in Figure 3. The established forward kinematics model can accurately reflect the mapping relationship between the joint motion and the end-effector pose. To further verify the kinematic mapping accuracy and workspace coverage validity

of the model and to intuitively demonstrate the reachable workspace of the drilling and anchoring arm, the end-effector positions in the world coordinate system were calculated by uniformly sampling the joint space based on the above kinematic model, and a three-dimensional scatter plot of the manipulator's workspace was generated (see Figure 3). This visualization clearly presents the set of all reachable end-effector positions of the drilling and anchoring arm within the specified joint ranges: Joint 1: $\pm 180^\circ$; Joint 2: -60° to 0° ; Joint 3: 0 to 1000 mm; Joint 4: $\pm 160^\circ$; Joint 5: $\pm 180^\circ$. Its spatial morphology is consistent with the structural design expectations, i.e., the manipulator can cover the roof working area of roadways with a width of 4.5–5.5 m and a height of 3–4 m, while ensuring structural stiffness and a compact installation space.

2.2. Forward Kinematic Modeling Considering Machine-Body Pose Errors

The coordinate transformation relationship between each link of the drill arm is described by a 4×4 homogeneous transformation matrix, which encapsulates the translational components and rotational components between adjacent link coordinate frames. The expression of the transformation matrix for the i -th link relative to the $(i - 1)$ -th link is as follows [21]:

$$T_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i & 0 & a_{i-1} \\ \sin \theta_i \cos \alpha_{i-1} & \cos \theta_i \cos \alpha_{i-1} & -\sin \alpha_{i-1} & -d_i \sin \alpha_{i-1} \\ \sin \theta_i \sin \alpha_{i-1} & \cos \theta_i \sin \alpha_{i-1} & \cos \alpha_{i-1} & d_i \cos \alpha_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The pose matrix of the drill arm end-effector relative to the world coordinates is obtained by sequentially multiplying the homogeneous transformation matrices of each link. As shown in Figure 4, due to body pose errors causing composite angle changes in the roll angle ϕ , pitch angle θ , and yaw angle ψ , the body rotation matrix R_{body} needs to be introduced to correct the coordinates, transforming the drill arm coordinate system (X_V, Y_V, Z_V) into the corrected coordinate system (X_C, Y_C, Z_C). The body pose rotation matrix can be expressed as follows:

$$R_{body} = R_z(\psi) \cdot R_y(\theta) \cdot R_x(\phi) \quad (2)$$

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{bmatrix}, \quad R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix},$$

$$R_z(\psi) = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Simultaneously, the length of the drill end in the fifth link coordinate system $L_{drill} = 1300$ mm, the offset of the drill arm base in the body coordinate system $O_{base} = [2000, 1150, 0]^T$, and the offset of the body coordinate system in the world coordinate system $O_{body} = [0, 0, 1000]^T$ are also incorporated. The position vector of the drill arm end P_{world} can be expressed as follows:

$$P_{world} = R_{body} \cdot ((T_{total} \cdot P_{local})_{1:3} + O_{base}) + O_{body} \quad (3)$$

where $T_{total} = T_1 \cdot T_2 \cdot T_3 \cdot T_4 \cdot T_5$ is the MDH transformation matrix for Joint 5, $P_{local} = [0, 0, L_{drill}, 1]^T$ is the homogeneous coordinate of the drill in the Joint 5 coordinate system, and $(T_{total} \cdot P_{local})_{1:3}$ is used to extract the first three components of $T_{total} \cdot P_{local}$, obtaining

the three-dimensional position vector of the drill end-effector in the manipulator base coordinate system.

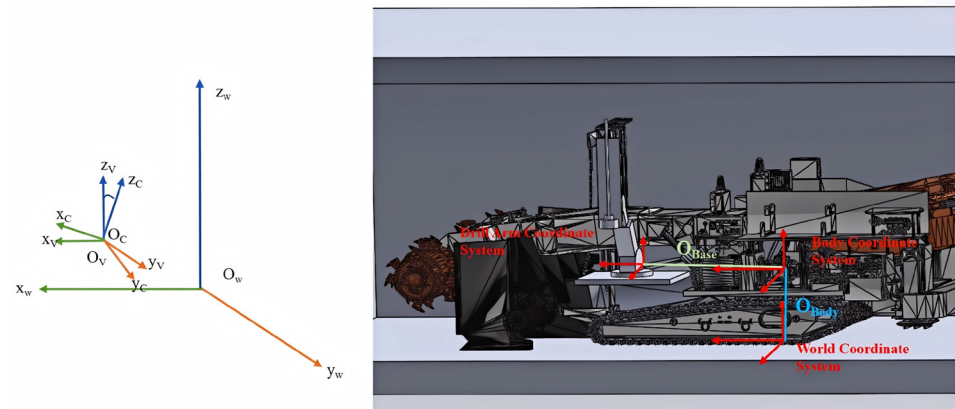


Figure 4. Coordinate system transformation schematic diagram.

Substituting the link transformation matrices into the drill end pose matrix operation yields the drill end pose matrix, which can be expressed as follows:

$$P_{world_h} = \begin{bmatrix} n_x & o_x & a_x & P_{dw,x} \\ n_y & o_y & a_y & P_{dw,y} \\ n_z & o_z & a_z & P_{dw,z} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

where $[n_x, n_y, n_z]^T$, $[o_x, o_y, o_z]^T$, $[a_x, a_y, a_z]^T$ are the normal vector, sliding vector, and approach vector of the drill end, respectively, and $P_{dw,x}$, $P_{dw,y}$, $P_{dw,z}$ are the position of the drill end in the world coordinate system.

The forward kinematics model of the drill arm established above accurately describes the mapping relationship between the joint motion and the end-effector pose, providing a reliable environment interaction model for the improved DDPG algorithm. The algorithm iteratively optimizes the joint adjustment strategy through real-time interaction with this model to achieve inverse kinematics solution and to pose error compensation.

2.3. Inverse Kinematics Problem Formulation

The inverse kinematics solution for the drilling and anchoring arm is a core technology for intelligent roadway equipment. Traditional geometric and numerical methods suffer from poor adaptability and the curse of dimensionality [22]. Reinforcement learning, as a model-free adaptive optimization scheme, does not require an accurate mathematical analytical model. It can learn autonomously through trial-and-error and rewards during the interaction between the agent and the environment, adapting to the complex structural characteristics of the multi-joint, strongly coupled drill arm. Traditional reinforcement learning algorithms have issues such as insufficient representational capacity when handling high-dimensional state spaces and continuous action spaces. Integrating the feature extraction capability of deep learning and the decision-making capability of reinforcement learning, the DDPG algorithm employs methods like the Actor–Critic dual-network structure and experience replay mechanism, effectively solving the instability problems of traditional methods in continuous action spaces. The core of the algorithm is to iteratively update the state–action cumulative reward Q-value. Its iteration formula can be expressed as follows:

$$Q(s, a) \leftarrow r + \gamma Q'(s', a'; \theta^-) \quad (5)$$

where $Q(s, a)$ is the value estimate of the current state–action pair, r is the immediate reward after executing action a , γ is the discount factor, s' is the new state after executing action a , a' is the optimal action output by the Actor network in the new state, $Q'(s', a'; \theta^-)$ is the new state–action value output by the target Critic network, and θ^- is the parameter of the target network. Based on this, this study proposes an improved DDPG algorithm optimized for the drilling and anchoring operation scenario to achieve a high-precision inverse kinematics solution. The optimization for drilling and anchoring operations is reflected in the task-specific design of the state–action space, reward function, exploration strategy, and network mechanism, which are modeled and analyzed in detail in the following subsections. Firstly, by defining state–action spaces and an error objective function adapted to the characteristics of the drill arm, a sequential decision-making model for reinforcement learning is constructed. Then, the Actor–Critic dual-network architecture and core functions are designed to achieve efficient iteration of the state perception–action output–reward feedback–parameter update.

2.3.1. Definition of State and Action Spaces

This study transforms the inverse kinematics problem of the drill arm into a sequential decision-making problem in reinforcement learning, requiring clear definition of the state space and action space. The state space S integrates the current joint states, positional deviation, and directional deviation, constructing an eleven-dimensional state vector. The specific expression is as follows:

$$s = \left[\frac{q_1}{180}, \frac{q_2}{90}, \frac{q_3}{1000}, \frac{q_4}{180}, \frac{q_5}{180}, \frac{e_{px}}{1000}, \frac{e_{py}}{1000}, \frac{e_{pz}}{1000}, \frac{e_{dx}}{2}, \frac{e_{dy}}{2}, \frac{e_{dz}}{2} \right] \quad (6)$$

where $q_1 \sim q_5$ is the state parameter of each joint, $e_{px} = P_{dw,x} - P_{t,x}$, $e_{py} = P_{dw,y} - P_{t,y}$, $e_{pz} = P_{dw,z} - P_{t,z}$ are the error components between the end position and the target position, $e_{dx} = d_{curr,x} - d_{t,x}$, $e_{dy} = d_{curr,y} - d_{t,y}$, $e_{dz} = d_{curr,z} - d_{t,z}$ are the error components between the end direction and the target direction; and P_{dw} = drill position, P_t = target position, d_{curr} = drill direction, $d_t = [0, 0, 1]$. Each dimensional parameter is normalized to effectively ensure the stability of network training. To balance input scales and to stabilize training, each dimension is normalized by its maximum variation range: rotational joints $180^\circ (\div 180)$, prismatic joint $0 - 1000$ mm ($\div 1000$), position error ± 1000 mm ($\div 1000$), and direction cosine deviation $\pm 2 (\div 2)$. Normalization restricts all inputs to $[-1, 1]$, mitigating gradient explosion/vanishing and improving training stability.

The action space A contains the adjustment amounts of the five joints, which can be expressed as follows:

$$a = [\Delta q_1, \Delta q_2, \Delta q_3, \Delta q_4, \Delta q_5] \quad (7)$$

where $\Delta q_1, \Delta q_2, \Delta q_3, \Delta q_4, \Delta q_5$ are the angle adjustment amounts of the rotational joints (unit: $^\circ$), Δq_3 is the displacement adjustment amount of the prismatic joint (unit: mm), and the action output is mapped to the $[-0.1, 0.1]$ interval via the tanh function, then converted to actual joint adjustment amounts through a scaling factor.

2.3.2. Error Objective Function

The inverse solution search problem can be defined as an error minimization optimization problem, thereby defining the error objective function for inverse kinematics solution. Its functional expression is as follows:

$$f(s) = \lambda_p \cdot \sqrt{e_{px}^2 + e_{py}^2 + e_{pz}^2} + \lambda_d \cdot \arccos \left(\frac{\vec{d}_{curr} \cdot \vec{d}_{target}}{|\vec{d}_{curr}| \cdot |\vec{d}_{target}|} \right) \quad (8)$$

where $\lambda_p = 0.6$, $\lambda_d = 0.4$ are error weighting coefficients used to balance the influence of position error and end-effector orientation angle error, according to the actual working conditions at coal mine sites, and the requirement for position accuracy is slightly higher than that for verticality, $\vec{d}_{curr}$ is the actual direction vector of the end-effector, and $\vec{d}_{target}$ is the target direction vector $[0, 0, 1]^T$ representing vertically upward. A smaller value of this expression indicates that the drill end pose of the drilling and anchoring arm is closer to the preset target pose.

2.4. Improved DDPG Algorithm Framework

2.4.1. Network Architecture Design

To translate the mathematical model into a feasible solution algorithm based on the classic DDPG framework, the network structure and training mechanism are optimized for the engineering requirements of the drilling and anchoring arm inverse solution. The specific framework design is as follows.

The improved DDPG algorithm adopts an Actor–Critic dual-network architecture, incorporating target networks and an experience replay mechanism to enhance algorithm stability during the solution process and to improve the convergence rate. A schematic diagram of the network iteration is shown in Figure 5.

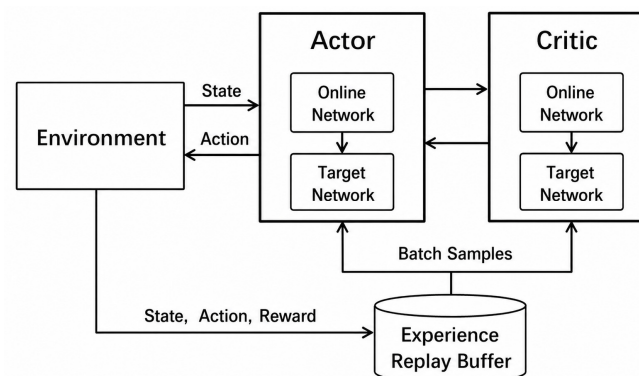


Figure 5. Network iteration schematic diagram.

The Actor network takes the state quantity s as the input and outputs the action vector a . The network uses a three-layer fully connected structure design. The hidden layers use the ReLU activation function, and the output layer uses the tanh function for normalization. Its expression is as follows:

$$a = \max_{action} \cdot \tanh(W_3 \cdot \text{ReLU}(W_2 \cdot \text{ReLU}(W_1 s + b_1) + b_2) + b_3) \quad (9)$$

where $W_1 \in R^{256 \times 11}$, $W_2 \in R^{128 \times 256}$, $W_3 \in R^{5 \times 128}$ are weight matrices and b_1, b_2, b_3 are bias terms. Hidden sizes 256/128 are determined via grid search, balancing performance, and computational efficiency.

The Critic network takes the state vector s and the action vector a as the inputs, and outputs the state–action value $Q(s, a)$. It also employs a three-layer fully connected

structure, concatenating the state and action vectors before inputting them into the network. Its expression is as follows:

$$Q(s, a) = W'_3 \cdot \text{ReLU}\left(W'_2 \cdot \text{ReLU}(W'_1 \cdot [s; a] + b'_1) + b'_2\right) + b'_3 \quad (10)$$

where $W'_1 \in R^{256 \times 11}$, $W'_2 \in R^{128 \times 256}$, $W'_3 \in R^{5 \times 128}$ are weight matrices and $[s; a]$ denotes the concatenation of state and action.

To avoid oscillation phenomena during parameter updates in training, the Actor target network and the Critic target network are set up separately. Their parameters are obtained from the main network through a soft update method. The update formula can be expressed as follows:

$$\theta_{\text{target}} = \tau \cdot \theta_{\text{main}} + (1 - \tau) \cdot \theta_{\text{target}} \quad (11)$$

where θ_{target} is the target network parameter; θ_{main} is the main network parameters; and τ is the soft update coefficient. This soft update is executed at every training iteration to update the target Actor/Critic networks, thereby stabilizing training and avoiding oscillation.

An experience replay buffer with a capacity of 10,000 is set up to store experience samples from the agent–environment interaction. During the training phase, batch samples are randomly drawn to update the network parameters, thereby breaking the correlation between samples and further enhancing training stability.

2.4.2. Core Function Design

(1) Reward Function

The overall reward function adopts a hierarchical gradient reward setting to balance convergence efficiency and solution accuracy. The overall expression is as follows:

$$R = R_{\text{pos}} + R_{\text{dir}} + R_{\text{bonus}} \quad (12)$$

where R_{pos} is the position reward, designed in segments based on the magnitude of the position error. When the error is less than 50 mm, an exponential reward form is used to encourage the agent to achieve high-precision convergence, which can be expressed as follows:

$$R_{\text{pos}} = \begin{cases} 100 \cdot e^{-e_p/25} & e_p < 50 \\ -e_p/50 & e_p \geq 50 \end{cases} \quad (13)$$

where R_{dir} is the direction reward and $e_p = \sqrt{e_{px}^2 + e_{py}^2 + e_{pz}^2}$ is the position error (unit: mm). A linear penalty is designed based on the end direction angle error of the drilling and anchoring arm to guide the agent's end-effector to adjust to a direction perpendicular to the roof, which can be expressed as follows:

$$R_{\text{dir}} = -5 \cdot \arccos\left(\frac{\vec{d}_{\text{curr}} \cdot \vec{d}_{\text{target}}}{|\vec{d}_{\text{curr}}| \cdot |\vec{d}_{\text{target}}|}\right) \times \frac{180}{\pi} \quad (14)$$

where R_{bonus} is the high-precision reward. A stepwise reward mechanism is set up to encourage the drilling arm agent to achieve millimeter-level operation accuracy, which can be expressed as follows:

$$R_{bonus} = \begin{cases} 850 & e_p < 1 \text{ and } e_d < 1^\circ \\ 350 & e_p < 1 \text{ and } e_d \geq 1^\circ \\ 150 & e_p < 2 \\ 50 & e_p < 5 \\ 0 & \text{other} \end{cases} \quad (15)$$

where $e_d = \arccos(d_{curr} \cdot d_{target}) \times 180/\pi$ is the direction error (unit: $^\circ$) and the coefficients are calibrated for convergence.

(2) Loss Function

The Critic network loss function uses mean squared error loss to measure the deviation between the current Q-value and the target Q-value. The expression is as follows:

$$L_{critic} = \frac{1}{B} \sum_{i=1}^B (Q(s_i, a_i) - y_i)^2 \quad (16)$$

where $B = 64$ for the batch size, $y_i = r_i + (1 - d_i) \cdot \gamma \cdot Q'(s_{i+1}, a'_{i+1})$ is the target Q-value, γ is the discount factor, r_i is the immediate reward of the i -th sample, d_i is the termination variable, where $d_i = 1$ indicates the required accuracy is reached, Q' is the output of the Critic target network, and a' is the output of the Actor target network.

The Actor network loss function uses policy gradient loss. The core objective is to maximize the state–action value. The expression is as follows:

$$L_{actor} = -\frac{1}{B} \sum_{i=1}^B Q(s_i, a_{actor}(s_i)) \quad (17)$$

where $a_{actor}(s_i)$ is the action vector output by the Actor network.

(3) Exploration Strategy

This study adopts a dynamic noise attenuation strategy to reasonably balance the agent's exploration and exploitation capabilities. For the first 10% of episodes, a noise intensity of 0.3 is used, gradually decreasing until episode 600. For the last 400 episodes, the noise is 0. The noise intensity σ expression is as follows:

$$\sigma = \begin{cases} 0.3 & \text{if } k < 0.1M \\ \max\left(0.01, 0.3 - \frac{0.3(k-100)}{500}\right) & \text{if } 0.1M \leq k < 600 \\ 0.0 & \text{if } k \geq 600 \end{cases} \quad (18)$$

where k is the current training episode number and M is the total number of training episodes.

Additionally, an optimal state restart mechanism is introduced, with a 30% probability of starting exploration from the historical best state, which can help the agent avoid becoming trapped in local optima. The hierarchical gradient reward function guides the agent to prioritize convergence to high-precision regions. The dynamic noise attenuation strategy ensures exploration breadth in the early training phase and improves convergence precision in the later phase. The optimal state restart mechanism and the experience replay mechanism work together to break sample correlation. These three aspects collectively

enhance the stability and efficiency of the algorithm in solving the inverse kinematics problem of the drilling and anchoring arm.

2.5. Simulation Configuration and Workflow

To verify the effectiveness of the improved DDPG algorithm, experiments were conducted for simulation verification based on the VS Code 1.120 development environment, with software dependencies including Anaconda3 3.11.7 and Matplotlib 3.7.1. The initial experimental parameters were set with reference to the actual working conditions of the drilling and anchoring arm as follows: total training episodes 1000, experience replay buffer capacity 10,000, and batch sampling size 64. The experiment used the standard DDPG algorithm as a comparative baseline. Both groups were tested under the same initial joint states, with initial joint values set to $[0^\circ, 60^\circ, 0 \text{ mm}, 0^\circ, 0^\circ]$. The initial body pose errors for pitch, roll, and yaw angles were set to vary randomly within the interval $[-3^\circ, 3^\circ]$. The target position was set to $[2134, 1070, 4430] \text{ mm}$, and the target pose direction vector was $[0, 0, 1]$.

The MATLAB-based workspace visualization procedure is illustrated in Figure 6. First, the feasible intervals of the MDH parameters and joint variables are defined, and the Monte Carlo sample size is set. Then, random joint-variable samples are generated, and the corresponding end-effector positions are calculated through forward kinematics. Finally, all sampled positions are aggregated to generate the end-effector workspace point-cloud map.

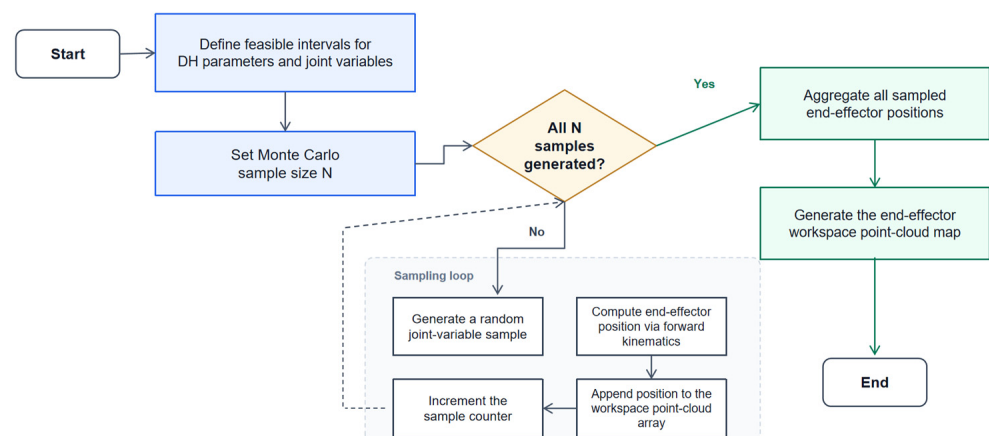


Figure 6. Flowchart of the MATLAB-based end-effector workspace point-cloud generation process.

3. Results and Discussion

The generated reward curve and position error curve comparison charts are shown in Figures 7 and 8, respectively.

The reward curve and position error curve collectively reveal the significant advantages of the improved DDPG algorithm in training convergence and solution accuracy. As shown in Figure 6, the reward value of the improved DDPG algorithm rises rapidly and stabilizes within the first 500 episodes, with negligible fluctuations in the later stages, demonstrating excellent convergence stability. By contrast, the standard DDPG algorithm exhibits slow reward growth and severe oscillations throughout training, failing to achieve effective convergence. This fully indicates that the hierarchical gradient reward function, dynamic noise attenuation strategy, and optimal state restart mechanism can effectively balance exploration and exploitation, accelerate algorithm convergence, and suppress training oscillations. Correspondingly, the position error curve in Figure 6 shows a completely opposite trend to the reward curve: the end-effector position error of the improved DDPG algorithm decreases rapidly as training progresses, converging stably to within 1 mm after 500 episodes. By comparison, the standard DDPG algorithm maintains a relatively high

error of 2–3 mm after convergence, resulting in low accuracy. The consistent variation patterns of the two curves mutually verify that the improved algorithm can effectively couple dynamic machine body pose errors with joint adjustments to achieve a high-precision end-effector pose solution. It significantly outperforms the conventional DDPG algorithm in both convergence efficiency and solution accuracy, providing reliable algorithmic support for the real-time pose compensation of drilling and anchoring arms in underground coal mines. After training, the model was saved and tested for three rounds. The test error results are shown in Table 2.

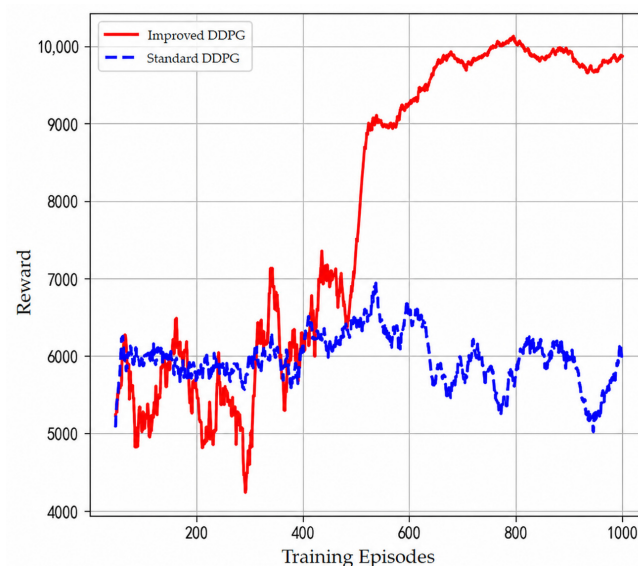


Figure 7. Reward value curve graph.

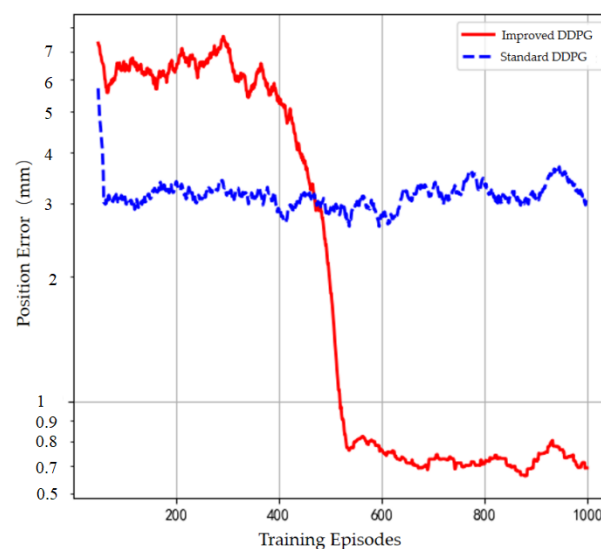


Figure 8. Position error curve graph.

Table 2. Training Test Error Table.

Episode	End Coordinate (mm)	Position Error (mm)	Direction Vector	Deviation Angle (°)
1	[2134.457,1070.189,4429.515]	0.6927	[0.002,−0.012,0.998]	0.697
2	[2134.475,1070.293,4429.517]	0.7381	[0.003,−0.015,0.998]	0.876
3	[2134.361,1070.203,4429.813]	0.4544	[0.003,−0.012,0.999]	0.711

After 1000 complete training episodes, compared with the standard DDPG algorithm, the improved DDPG algorithm converges after 500 episodes, and the solution accuracy of the model meets the engineering application criteria. In the final 50 stable training episodes, the average position error at the end of the drill bit of the drilling and anchoring arm is 0.71 mm, and the maximum position error is no more than 0.95 mm. The directional angular deviation is controlled within 1° . Such accuracy is far superior to the specifications stipulated in the Chinese industry standard MT/T 1171-2019 Technical Specification [23] for Bolt Support in Coal Mine Roadways, which limits the position deviation to no more than 5 mm and the allowable angular deviation between the actual drilling angle and the designed angle of bolt holes to within 5° . The proposed algorithm satisfies the stringent requirements on positioning and attitude accuracy for drilling and anchoring operations in roadway support, which verifies its high-precision solving capability under complex working conditions.

In underground roadway support engineering, the perpendicularity of bolt holes directly determines the bearing capacity of the support structure. Excessive angular deviation may lead to support failure and even potential roadway instability risks. By introducing a weight term of directional error into the reward function, the improved algorithm incorporates end-effector attitude optimization into the core optimization objective, achieves the collaborative optimization of position and attitude, and guarantees the safety and reliability of anchoring operations from the algorithm level.

4. Conclusions

To tackle practical issues including composite roll, pitch, and yaw errors caused by uneven roadway floors and the insufficient dynamic compensation performance of traditional 1–3-DOF drilling and anchoring arms, this paper develops an improved DDPG algorithm for inverse kinematics and pose compensation of a 5-DOF manipulator. A full-link MDH kinematic model is established with body pose deviation, drill rod length, and assembly offset incorporated into coordinate transformation. Targeting the dual precision requirements of actual bolting tasks, we construct an eleven-dimensional state space and a five-dimensional action space, and optimize the standard DDPG algorithm using the hierarchical gradient reward function, dynamic noise decay strategy, and optimal state restart rule. Under simulated $\pm 3^\circ$ random body pose disturbances, the algorithm converges steadily after 500 training episodes, achieving an average position error of 0.71 mm and an angular error below 1° , which distinctly outperforms the permissible tolerance stated in the MT/T 1171-2019 industrial specification. Different from previous studies ignoring fuselage disturbance, this method realizes coupled optimization between body errors and joint adjustments, and its compact calculation structure is suitable for embedded control in underground equipment.

Nevertheless, all validations are completed via numerical simulation based on an ideal kinematic model, ignoring various practical interferences, such as mechanical clearance, hydraulic response delay, structural deformation, and sensor noise, existing in real equipment. In follow-up work, we plan to deploy the algorithm on a physical integrated roadheader-bolter test bench to verify its practical robustness. In addition, relevant online calibration technologies will be studied to improve environmental adaptability, and further research will focus on collaborative error compensation for multiple manipulators to promote the intelligent development of coal mine excavation and anchoring operations.

Author Contributions: Investigation, Z.L.; Resources, H.H.; Writing—original draft preparation, X.D.; Writing—review and editing, J.Y.; Supervision, G.W. All authors have read and agreed to the published version of the manuscript.

Funding: (1) Research and application of intelligent mining and transportation equipment with robotization for mines (MRI23002); (2) Theory and Method of Excavation-Support-Anchor Parallel Control for Intelligent Excavation complex System (Grant No. 52104169).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are unavailable due to privacy restrictions. No public involvement in any aspect of this research, No specific reporting guidelines were used in drafting this manuscript; AI or AI-assisted tools were not used in drafting any aspect of this manuscript.

Conflicts of Interest: Author Xuan Dong was employed by the company Shaanxi Blower (Group) Co., Ltd. The remaining 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.

Nomenclature

Symbol	Physical Meaning	Unit
a_{i-1}	Length of the (i − 1)-th connecting link	mm
α_{i-1}	Torsion angle of the (i − 1)-th connecting link	rad
d_i	Joint offset of the i-th joint	mm
θ_i	Rotation angle of revolute joint i	°
Δq_3	Displacement variation of prismatic Joint 3	mm
R_{body}	Rotation transformation matrix of whole drilling–anchoring arm base	Dimensionless
T_{total}	Total homogeneous transformation matrix from joint to manipulator base	Dimensionless
P_{local}	Local coordinate vector of drill end at terminal link	mm
$(T_{total} \cdot P_{local})_{1:3}$	First three components of position vector extracted from homogeneous coordinates	mm
O_{base}	Mounting offset vector of manipulator base in fuselage coordinate system	mm
O_{body}	Offset vector of fuselage coordinate relative to world coordinate	mm
S	State space of reinforcement learning	Dimensionless
A	Action space of reinforcement learning	Dimensionless
R	Overall hierarchical gradient reward function	Dimensionless
e_p	Euclidean position error between actual and target drill position	mm
e_d	Angular deviation of drilling direction from ideal vertical direction	°
γ	Discount factor of reinforcement learning	Dimensionless
σ	Attenuation noise intensity for exploration strategy	Dimensionless
τ	Soft-update coefficient of target network	Dimensionless
B	Batch sampling size of experience replay	Dimensionless

References

1. Sun, L.L.; Liang, F.Y.; Fang, L.J. Design and performance analysis of an industrial robot arm for robotic drilling process. *Ind. Robot-Int. J. Robot. Res. Appl.* **2019**, *46*, 7–16. [\[CrossRef\]](#)
2. Zhiwei, W.; Lei, Y.; Tianhong, C.; Wenfu, X.; Bin, L. An Efficient Geometry-Informed Inverse Kinematics of a 7 DOF Cable-Driven Manipulator with Non-Sphere Shoulder and Wrist. *Robot. Auton. Syst.* **2026**, *198*, 105335. [\[CrossRef\]](#)
3. Quiñonez, Y.; Zatarain, O.; Lizarraga, C.; Aguayo, R.; Mejía, J. Numerical Method Using Homotopic Iterative Functions Based on the via Point for the Joint-Space Trajectory Generation. *Appl. Sci.* **2023**, *13*, 1142. [\[CrossRef\]](#)
4. Katiyar, S.; Dutta, A. End-effector path tracking of a 14 DOF Rover Manipulator System in CG-Space framework. *Robotica* **2024**, *42*, 265–301.

5. Zhang, S.; Li, A.; Ren, J.; Ren, R. Kinematics inverse solution of assembly robot based on improved particle swarm optimization. *Robotica* **2024**, *42*, 833–845. [[CrossRef](#)]
6. Ning, Y.; Sang, L.; Wang, H.; Wang, Q.; Vladareanu, L.; Niu, J. Upper limb exoskeleton rehabilitation robot inverse kinematics modeling and solution method based on multi-objective optimization. *Sci. Rep.* **2024**, *14*, 25476. [[CrossRef](#)] [[PubMed](#)]
7. Zhao, C.; Wei, Y.; Xiao, J.; Sun, Y.; Zhang, D.; Guo, Q.; Yang, J. Inverse kinematics solution and control method of 6-degree-of-freedom manipulator based on deep reinforcement learning. *Sci. Rep.* **2024**, *14*, 12467. [[CrossRef](#)] [[PubMed](#)]
8. Luo, S.; Chu, D.; Li, Q.; He, Y. Inverse Kinematics Solution of 6-DOF Manipulator Based on Multi-Objective Full-Parameter Optimization PSO Algorithm. *Front. Neurobot.* **2022**, *16*, 791796. [[CrossRef](#)] [[PubMed](#)]
9. Naranjo-Campos, F.J.; Victores, J.G.; Balaguer, C. Expert-Trajectory-Based Features for Apprenticeship Learning via Inverse Reinforcement Learning for Robotic Manipulation. *Appl. Sci.* **2024**, *14*, 11131. [[CrossRef](#)]
10. Bai, Y.; Luo, M.; Pang, F. An Algorithm for Solving Robot Inverse Kinematics Based on FOA Optimized BP Neural Network. *Appl. Sci.* **2021**, *11*, 7129. [[CrossRef](#)]
11. Sekkat, H.; Tigani, S.; Saadane, R.; Chehri, A. Vision-Based Robotic Arm Control Algorithm Using Deep Reinforcement Learning for Autonomous Objects Grasping. *Appl. Sci.* **2021**, *11*, 7917. [[CrossRef](#)]
12. Liu, Y.; Xi, J.; Bai, H.; Wang, Z. A General Robot Inverse Kinematics Solution Method Based on Improved PSO Algorithm. *IEEE Access* **2021**, *9*, 32341–32351. [[CrossRef](#)]
13. Danaci, H.; Nguyen, L.A.; Harman, T.L.; Pagan, M. Inverse Kinematics for Serial Robot Manipulators by Particle Swarm Optimization and POSIX Threads Implementation. *Appl. Sci.* **2023**, *13*, 4515. [[CrossRef](#)]
14. Ostermeier, D.; Külz, J.; Althoff, M. Automatic Geometric Decomposition for Analytical Inverse Kinematics. *IEEE Robot. Autom. Lett.* **2025**, *10*, 9964–9971. [[CrossRef](#)]
15. Ning, Y.; Hu, R.; Chen, X.; Ren, S.; Chen, Y.; Li, B. POE-Based Inverse Kinematics Analysis for a Six-DOF Robotic Manipulator. In *Proceedings of the 2022 IEEE 5th International Conference on Electronics Technology (ICET)*, Chengdu, China, 13–16 May 2022; IEEE: New York, NY, USA, 2022.
16. Seyyedabbasi, A. Solve the Inverse Kinematics of Robot Arms using Sand Cat Swarm Optimization (SCSO) Algorithm. In *Proceedings of the 2022 International Conference on Theoretical and Applied Computer Science and Engineering (ICTACSE)*, Istanbul, Turkey, 29 September–1 October 2022; IEEE: New York, NY, USA, 2022.
17. Diaz Posada, J.R.; Schneider, U.; Pidan, S.; Geravand, M.; Stelzer, P.; Verl, A. High Accurate Robotic Drilling with External Sensor and Compliance Model-Based Compensation. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*, Stockholm, Sweden, 16–21 May 2016; IEEE: New York, NY, USA, 2016.
18. Lei, M.Y.; Zhang, X.H.; Yang, W.J.; Wan, J.C.; Dong, Z.; Zhang, C.; Zhang, G.M. High-Precision Drilling by Anchor-Drilling Robot Based on Hybrid Visual Servo Control in Coal Mine. *Mathematics* **2024**, *12*, 2059. [[CrossRef](#)]
19. Li, S.; Chen, Z.M. Autonomous Capture Control of Non-Cooperative Targets by Space Manipulator Based on PPO Algorithm. In *Proceedings of the 23rd IFAC Symposium on Automatic Control in Aerospace (ACA)*, Harbin, China, 2–6 August 2025.
20. Niu, J.X.; Yu, Q.Y.; Bi, M.B.; Zhao, J.L.; Zhang, T. Deep Reinforcement Learning-Based Cooperative Harvesting Strategy for Dual-Arm Robots in Apple Picking. *Agronomy* **2025**, *15*, 2565. [[CrossRef](#)]
21. Wang, Z.; Zhang, J.; Chu, Y.; Wu, Y.; Mou, Y.; Wang, X.; Yang, H. Layered and Decoupled Calibration: A High-Precision Kinematic Identification for a 5-DOF Serial-Parallel Manipulator with Remote Drive. *Actuators* **2025**, *14*, 577. [[CrossRef](#)]
22. Lin, G.; Huang, P.; Wang, M.; Xu, Y.; Zhang, R.; Zhu, L. An Inverse Kinematics Solution for a Series-Parallel Hybrid Banana-Harvesting Robot Based on Deep Reinforcement Learning. *Agronomy* **2022**, *12*, 2157. [[CrossRef](#)]
23. MT/T 1171-2019; Technical Specifications of Floating Dust Detection for Underground Coal Mine. National Mine Safety Administration: Beijing, China, 2019.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.