

## **Research Article**

# Deep Reinforcement Learning and Auto-Differential Evolution Co-Guided Coal Washing

## Mingcheng Zuo 🕞

Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou, China

Correspondence should be addressed to Mingcheng Zuo; mingcheng.zuo@cumt.edu.cn

Received 8 January 2024; Revised 23 February 2024; Accepted 4 March 2024; Published 12 March 2024

Academic Editor: Ewa Pawluszewicz

Copyright © 2024 Mingcheng Zuo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Background.* Coal washing is a complicated process and difficult to control, which has many controlling parameters with strong coupling relationship. It is still a challenge to realize the self-perception, self-adjustment, and self-evaluation of coal washing machine, improve the quality of coal washing, ensure production safety, and reduce labor cost. *Methods.* Through the intelligent transformation of jig, this paper proposes an intelligent washing method with cooperated deep reinforcement learning and evolutionary computation. First, it designs a fault warning method based on statistical analysis, helping to recover the normal running state of jig with manual maintenance. Then, it constructs a regulation strategy generation method with deep reinforcement learning supported by the fusion of artificial experience and historical data. Last, for the lack of monitoring data caused by poor communication quality and environment, the regulation strategy prediction method with evolutionary computation and surrogate model is proposed. *Results.* In practice, this method shows accurate fault warning accuracy and rapid cleaned coal ash adjustment response ability. *Conclusions.* This shows that the method proposed in this paper is of great significance for intelligent washing and can better cope with the special situation when the washing equipment sensing data are missing.

## 1. Introduction

Coal preparation plants use various gravity processes to process most of the raw coal, which are known for their low cost and high process efficiency. Jig coal washing is the process of physically separating different substances from raw coal to form coal products of various quality specifications [1]. Intelligent coal washing can reduce the impact of manual activities on the separation process, shorten the delay time of process control, and improve efficiency. With the construction of intelligent mines, especially the demand for safe and high-quality production, it has become particularly important to quickly identify washing faults and independently adjust the washing process [2]. In addition to the basic problems of severe aging of raw coal and low sorting rate, the process of washing equipment is also affected by many factors, such as frequency, air pressure, air valve adjustment, hydraulic cylinder, coal gangue valve opening, and float weight, all of which have an important

impact on washing quality [3]. How to monitor the working status of fixtures in real time during the cleaning process, accurately perceive harsh working conditions, and effectively ensure production quality remains a challenge [4, 5]. The core of intelligent cleaning is to collect multidimensional operation data of jigs, quickly mine real-time status, and timely return accurate control strategies [6]. In previous studies, the coal separation process was regarded as a physical and mathematical modeling problem, and the control strategy was usually generated by referring to the established model [7-10]. With the ability to obtain massive sensor data and deeply mine available information, Internet of Things [11] and big data [12] techniques are potential to achieve intelligent coal washing. In terms of the jig running state, diagnosing the faults in time is the key to ensuring the safe operation of the system. Wang et al. [13] proposed a fault diagnosis method based on AlexNet convolutional neural network (CNN) from a data-driven perspective. When the washing equipment is running normally, monitoring the control parameters of the equipment and adjusting them in a timely manner can help effectively adjust the washing quality. Wang et al. [1] discovered the ability to intelligently control the density of heavy medium separation. The control process is implemented through the following program. Based on online scanning of raw coal ash content, the properties of excess coal are first calculated to predict and optimize process parameters including heavy medium separation density. The required density for optimizing the circulating medium will be transmitted to the control system. Wang et al. [14] developed an intelligent analysis system for raw coal float and sink test data, high-precision intelligent monitoring instruments and equipment for position, liquid level, and ash content, and highly reliable sensor equipment.

However, in the process of intelligent extraction and analysis of real-time data, the current methods mostly use statistical analysis, which has large errors and is difficult to integrate into the existing artificial experience. In addition, the rough environment of the coal mine washing site has seriously affected the stability of network communication, resulting in a large amount of data missing, which is difficult to meet the needs of real-time analysis and return the control strategy. In view of this, this paper embeds the existing artificial experience into the machine real-time control process by adopting deep reinforcement learning. When the collected data cannot meet the requirements of the analysis, we build a surrogate model with early accumulated historical data and use our proposed auto-differential evolution algorithm to quickly solve the potential operation scheme. At the time of automatically controlling the washing process of the jig, the detection of jig fault is also monitored to ensure the washing efficiency.

#### 2. Proposed Methods

2.1. Hardware Infrastructure. The overall architecture of the hardware foundation relied on by the proposed method is shown in Figure 1, which includes three parts: intelligent sensing of state data, intelligent analysis of state data, and generation and transmission of regulatory strategies. First, intelligent perception of data requires installing different types of sensors in the critical control of the jig, realizing the real-time collection of cleaned coal ash content, wind pressure, water pressure, hydraulic value, medium coal and gangue gate opening, buoy counterweight, coal gangue bucket lifting amount, and buoy value. These data are collected by maintaining a synchronized or similar data acquisition frequency, which is convenient for postprocessing. Then, the collected data are gathered in the data server via the OPC protocol, the current jig operational fault is analyzed, and the on-site personnel are informed to deal with it in time. Last, regulatory models based on deep reinforcement learning and evolutionary computation are constructed to generate regulatory strategies for the current jig operation, and the generated regulation strategy is sent back to the control end via the OPC protocol, realizing the automatic operation of jig. It is worth noting that the regulation strategy generated by this method does not consider the parameter adjustment range of jig in the actual production, so it is necessary to set the strategy filter at the PLC end to fine-tune the received regulation strategy, ensuring the safe operation of the jig. The technical core of this framework is the intelligent analysis method of real-time data for the regulatory model. The regulatory model of deep reinforcement learning and evolutionary computation will be described in detail below.

2.2. Regulation Model. In industrial processes, the safety and reliability of mechanical systems determine the quality of products. Timely diagnosis of small faults is the key to ensuring the safe operation of the system and suppressing the deterioration of faults. Given that the important prerequisite for regulating the jig is to maintain it under normal operating conditions, it is necessary to monitor the working status of the jigger in real time. When operating faults occur, warn the driver to handle them in a timely manner to improve regulation performance. The regulation model of the jigging machine adopts collaborative deep reinforcement learning and evolutionary computing (hereinafter referred to as DEIS), as shown in Figure 2. In this model, assuming that the sampling frequency of all jig sensors is set to f, a total of f pieces of data were collected within one second, each containing 32 values, including clean coal ash, air pressure, water pressure, hydraulic cylinder, medium coal and gangue door, float weight, coal gangue bucket volume, and float value. When the communication is good and the aggregated data volume exceeds half of the 32 \* f sampling value, a control strategy is adopted to generate the operation of the fixture. At the same time, warning messages for gate overload, scouring, and compaction are determined based on the height of the bucket, the amplitude of the float, and the opening of the gate. When the communication blocking amount and convergence data amount are less than or equal to half of the 32 \* f sampling value, use differential evolution algorithm to generate the adjustment strategy for the operation of the jig, and the warning information of network communication problems is fed back. Because network communication is very important for intelligent controlling, we use the sound and light warning to inform the jig driver to deal with the network problems in time.

#### 2.2.1. Regulation Strategy with Deep Reinforcement Learning

#### (1) Basic Definition of DQN

(1) Definition of State. The core elements of deep reinforcement learning are state  $s_t$ , action  $a_t$ , and reward  $r_t$  [15]. Corresponding to the operation process of the jig, we record the key control parameters of the jig as  $s_t$ , including the throughput frequency TF<sub>t</sub>, air pressure AP<sub>t</sub>, water pressure WP<sub>t</sub>, air valve AV<sub>t</sub>, the combination of valve opening VO<sub>t</sub>, hydraulic cylinder HC<sub>t</sub>, gangue gate gg<sub>t</sub>, buoy counterweight bc<sub>t</sub>, coal gangue bucket amount cg<sub>t</sub>, and buoy value bv<sub>t</sub> at time t. Overall,



FIGURE 1: The overall architecture of the hardware foundation.



FIGURE 2: The regulatory model of jig adopting the hybrid deep reinforcement learning and evolutionary computation.

$$s_t = [TF_t, AP_t, WP_t, AV_t, VO_t, HC_t, gg_t, bc_t, cg_t, bv_t].$$
(1)

(2) Definition of Action. Action a<sub>t</sub> corresponds to the single adjustment step size of abovementioned parameters, including the throughput frequency ΔTF<sub>t</sub>, air pressure ΔAP<sub>t</sub>, water pressure ΔWP<sub>t</sub>, air valve ΔAV<sub>t</sub>, the combination of valve opening ΔVO<sub>t</sub>, hydraulic cylinder ΔHC<sub>t</sub>, gangue gate Δgg<sub>t</sub>, buoy

counterweight  $\Delta bc_t$ , coal gangue bucket amount  $\Delta cg_t$ , and buoy value  $\Delta bv_t$ . Overall,

$$a_{t} = [\Delta TF_{t}, \Delta AP_{t}, \Delta WP_{t}, \Delta AV_{t}, \Delta VO_{t}, \Delta HC_{t}, \Delta gg_{t}, \Delta bc_{t}, \Delta cg_{t}, \Delta bv_{t}].$$
(2)

For the convenience of operation, the adjustment step size of  $a_t$  elements is set following operating habit, so the elements of  $a_t$  are all discrete. Different  $a_t$  elements have various step sizes. For example, the adjustment step sizes of the throughput frequency  $TF_t$ , air pressure  $AP_t$ , water pressure  $WP_t$ , air valve  $AV_t$ , the combination of valve opening  $VO_t$ , hydraulic cylinder  $HC_t$ , gangue gate  $gg_t$ , buoy counterweight  $bc_t$ , coal gangue bucket amount  $cg_t$ , and buoy value  $bv_t$  are 1 Hz, 0.001 Mpa, 0.1 MPa, 1%, 1°, 0.1 MPa, 1°, 50 g, 0.1 m, and 1 cm. In engineering applications, the machines are usually not allowed to operate frequently or on a large scale. Hence, a constraint is applied to the performed action as follows:

$$|a_t| < \psi, \tag{3}$$

where  $|a_t|$  means the number of nonzero elements and  $\psi$  is the threshold meeting the practical requirements. This operation also can reduce the space size of the search space and downgrade the pressure of training the mapping models.

(3) Definition of Reward. It is worth noting that due to the inertia of the operation process, the change of cleaned coal ash content within T<sub>0</sub> minutes after executing strategy a<sub>t</sub> has higher stability and credibility, that is, more accurate. The cleaned coal ash needs to be controlled within a given range [Ob - ε, Ob + ε], where Ob is the expected cleaned coal ash and ε is the tolerance. In the evaluation period, assume that the initial cleaned coal ash is cx and the adjusted cleaned coal ash after performing a<sub>t</sub> is given as follows:

$$r_t = \left(\max\left(1 - \frac{|\mathrm{Ob} - \mathrm{nx}|}{|\mathrm{Ob} - \mathrm{cx}|}, 0\right)\right)^2. \tag{4}$$

It is worth noting that if cx is within  $[Ob - \epsilon, Ob + \epsilon]$ , there is no need performing an  $a_t$ , so |Ob - cx| is ensured to be larger than 0. 1 - |Ob - nx|/|Ob - cx|means that if the cleaned coal ash is adjusted to Ob, |Ob - nx| is 0, and 1 - |Ob - nx|/|Ob - cx| equals to 1. If nx is away from the expected Ob, 1 - |Ob - nx|/|Ob - cx| is negative, and  $\max(1 - |Ob - nx|/|Ob - cx|, 0)$  assigns  $r_t$  to 0. In this way,  $r_t$  is controlled between [0, 1]. In addition, to motivate the DQN to quickly adopt efficient actions,  $r_t$  is further developed to  $r_t = (\max(1 - |Ob - nx|/|Ob - cx|, 0))^2$ , where  $r_t$  is more clearly divided into two parts of lower and higher ranges. This operation can reduce the probability of action selection faults caused by reward value calculation errors.

(2) Training of DQN. With the generalization ability of deep neural network, the mapping model between  $(s_t, a_t)$  and  $r_t$  can be established. The mapping model has ability to recommend the best  $a_t$  to maximize  $r_t$ , that is, to realize the regulation of cleaned coal ash fastest. As shown in Figure 3, this paper uses the typical deep reinforcement learning

model (Deep Q-learning Network, DQN) to build the mapping relationship with  $(s_t, a_t)$  and  $r_t$ . The first half of Figure 3 is the training of DQN, and the second half is the application of DQN. The training data come from the automatically collected data during the operation of jig and drivers' experience. To transfer the drivers' experience to the formatted training data,  $a_t$  is determined by drivers with experience, and during the period of collecting the training data, the machine state  $s_t$  and the action  $a_t$  are recorded, and the reward  $r_t$  is also archived. The jig operation data are continually collected before the application of DQN. In this way, once an operation is conducted by the jig driver, a piece of operation data  $d_t = (s_t, a_t, r_t)$ , namely, training data, is collected. When a certain amount of training data is collected, the DQN is trained and a simulator is used to illustrate the performance of the trained DQN. The effectiveness of initially trained DQN is tested with two aspects: (1) whether an effective action can be recommended; (2) whether the actual reward is near to the expected reward. Considering that the recommended action is not safe to be directly performed on the jig, the effectiveness of DQN is manually judged. When the trained DQN is not sufficient as a controller, the training data are still recorded and the DQN is constantly updated. Once the trained DQN can recommend effective action and predict relatively correct reward, the trained DQN is employed. To reduce the error of verifying the effectiveness of DQN, overall  $\tau$  tests are conducted. If the success rate meets the requirements, the DQN is applied. Although DQN is applied as a controller, we still set an emergency measurement. After repeated regulation and control, if the jig does not meet the expected requirements, jig will exit the automatic control with alarming and will be controlled manually.

Considering that fully connected neural networks are capable of learning complex patterns and relationships in data, the neural network implementing DQN is a fully connected neural network. In the training data, a number of pieces of data are same, due to the same initial machine states. So, b samples with the biggest difference are selected in the specific training process and normalized from the recorded operation dataset M. Although equation (2) shows that  $s_t$  just contains 10 elements, parts of elements, like  $\Delta AP_t$ , contain 4 members from different operating rooms. So, actually 32 elements are contained in  $s_t$  as mentioned in Section 2.2. Also,  $a_t$  also contains 32 elements, and the input of trained neural network, i.e.,  $(s_t, a_t)$ , is with 64 dimensions. That means the neural network should have deep layers and big number of neurons. In this paper, the number of neurons in all the hidden layers is [64, 32, 16, 8, 4], respectively. For the learning rate, when the learning rate is too high, the model may experience oscillations or divergence during the training process, resulting in the inability to converge to the optimal solution. In this case, the model may skip the optimal solution or wander around the optimal solution, unable to achieve good training results. When the learning rate is too low, the training speed of the model will become very slow, and it may require more iterations to converge to the optimal solution. So, in our training process, we set two training stages with higher and lower learning rates,



FIGURE 3: The training and application of DQN mapping  $(s_t, a_t)$  and  $r_t$ .

respectively. In the first stage, the learning rate is 0.01, while in the second stage, the learning rate is 0.001.

When the input of the sigmoid function approaches positive or negative infinity, the output approaches 0 or 1, which helps to improve the stability and convergence of the neural network. So, the hidden layers are connected by the sigmoid function, and the activation function of the output layer is a linear function. Then, predictive values of the training samples are obtained by forward propagation  $z^{(l)} = W^{(l)}c^{(l-1)} + b^{(l)}$ , where  $c^{(l)} = f_l(z^{(l)}), c^{(0)} = (a_l, s_l)$ .  $c^{(l-1)}$  is the input of  $l^{\text{th}}$  layer from  $(l-1)^{\text{th}}$  layer, and  $W^{(l)}$ and  $b^{(l)}$  are the weights and bias of  $l^{\text{th}}$  layer. Calculate the prediction error  $L(W,b) = \sum_{t=1}^{N} (r_t - Q(s_t, a_t; \theta))^2$ , and adjust the weights W and bias b of the network along the negative gradient direction of the prediction error, where  $\theta$  is the union of weights W and bias b,  $Q(s_t, a_t; \theta)$  is the output of the DQN, and N is the number of trained samples. When the error requirements are met, stop the training on the DQN.

(3) Application of DQN. After the training of DQN, the running state of jig  $s_t$  at the time t is taken as the input of DQN to obtain the predicted performance value of different regulated actions. For all  $a_t$ , the control action with the largest predicted value is  $a_t^{\max} = \arg \max\{r_t^{\prime}\}$ . Then,  $a_t^{\max}$  is the subsequent regulation strategy for jig. In order to balance exploration and exploitation, the Epsilon greedy strategy is employed, which strikes a balance between exploring new options (with a probability of  $\epsilon$ ) and exploiting the best-known option (with a probability of  $1 - \epsilon$ ). The Epsilon greedy strategy is given as follows:

$$a_t = \begin{cases} \text{rand\_strategy,} & \text{if rand} < \epsilon, \\ a_t^{\max}, & \text{otherwise,} \end{cases}$$
(5)

where rand\_strategy means selecting a random regulatory strategy, which needs to consider the parameter constraints in terms of the safe use of the jig. Furthermore, to avoid the significant adjustment of the jig parameters, the parameter nearest to the current setting is first adjusted as selected  $a_t$ . Only one parameter is adjusted at a time, and the other parameters are temporarily not adjusted. If the previous adjustment is not effective, the other parameters will be adjusted in turn.

2.2.2. Regulatory Strategies with Auto-Differential Evolution. The generation of jig regulation strategy is essentially a solution of a high-dimensional optimization problem. Unlike the deep reinforcement learning generating regulatory strategies, evolutionary computation does not require known jig states, i.e.,  $s_t$ , but only a surrogate model evaluating candidate solutions and parameter boundaries is needed. Evolutionary algorithms, such as differential evolution (DE) [16-19], are effective in solving the highdimensional optimization problems. However, to avoid large adjustment of equipment parameters, DE can only locally fine-tune the control parameters. Moreover, considering the timeliness requirement of online processes, DE needs to have the ability of rapid convergence. In view of this, this paper proposes auto-differential evolution (Au-DE) to improve the autonomous learning ability of the control parameters and accelerate the search convergence. The basic operation of Au-DE is to randomly generate NP individuals, and in each generation, the mutation operator, cross over operator, and selection operator are constantly performed to the population until the last generation  $G_{\text{max}}$ , until the evolutionary process is over. Assuming that the optimization problem is a constrained minimization problem with D dimensions, the optimization objective is min f(x), where f(x) is the objective function. A solution is shown as

$$x = (x_1, \cdots, x_j, \cdots, x_D).$$
(6)

Then, the population individual (candidate solution) can be further expressed as

$$x_{i,G} = (x_{1,i,G}, \cdots, x_{j,i,G}, \cdots, x_{D,i,G}),$$
 (7)

where  $i = 1, 2, \dots, NP, G = 1, 2, \dots, G_{max}$ . *G* is the current generation number, and Gmax is the maximum generation number.

For the individual  $x_{i,G}$  in generation *G*, the corresponding mutation operator is

$$v_{i,G} = x_{r1,G} + F \times (x_{r2,G} - x_{r3,G}),$$
(8)

where  $r_1, r_2, r_3$  are integers,  $r_1, r_2, r_3 \in [1, \text{NP}], r_1 \neq r_2 \neq r_3 \neq i, F \in (0, 1]; F$  is a scaling factor controlling the difference vector  $x_{r_2,G} - x_{r_3,G}$ .

If  $v_{i,G}$  exceeds the boundary, it is repaired by

$$v_{j,i,G} = \max(L_j, \min(v_{i,G}, U_j)), \tag{9}$$

where  $L_j$  and  $U_j$  are the lower and higher boundaries on the  $j^{\text{th}}$  dimension, respectively.

The goal of the crossover operator is to produce intermediate solutions:

$$u_{i,G} = \left( u_{1,i,G}, \cdots, u_{j,i,G}, \cdots, u_{D,i,G} \right),$$
(10)

and

$$u_{j,i,G} = \begin{cases} u_{j,i,G}, & \text{if } \operatorname{rand}(0, 1) < \operatorname{CR} \operatorname{or} j = \operatorname{rand}(1, D), \\ x_{j,i,G}, & \text{otherwise,} \end{cases}$$
(11)

where CR is the crossover probability.

The selection operator is used to select individuals with better fitness among parent and child individuals by

$$x_{i,G} = \begin{cases} u_{i,G}, & \text{if } f(u_{i,G}) < f(x_{i,G}), \\ x_{i,G}, & \text{otherwise.} \end{cases}$$
(12)

In original DE, F and CR are randomly generated within [0, 1]. Though this operation has strong global search ability, it is not suitable for quickly finding optimal solutions and returning the regulatory strategies in time. In our proposed Au-DE, the successful F and CR helping a population individual in generation G to find the better child solution are recorded and reused. In this mode, F is generated by

$$F_i = \operatorname{rand} c_i (M_{F,ri}, 0.1), \tag{13}$$

where  $M_{F,ri}$  is the memory of successful  $F, ri \in [1, H], H$  is the total number of memory entries, and  $randc_i(*)$  is the Cauchy distribution function. If  $F_i > 1$ ,  $F_i \leftarrow 1$ ; if  $F_i \le 0$ ,  $F_i$  is generated by repeatedly executing equation (9), until it meets requirements.  $M_{F,k}$  ( $k = 1, \dots, H$ ) is initialized to 0.5. From the second generation,  $M_{F,k}$  is updated by

$$M_{F,k,G+1} = \begin{cases} \operatorname{mean}_{\mathrm{WL}}(S_F), & \text{if } S_F \neq \phi, \\ M_{F,k,G}, & \text{otherwise,} \end{cases}$$
(14)

where k is the memory site, initialized by 1.  $\operatorname{mean}_{WL}(S_F)$ means the weighted sum of  $S_F$  which is the successful F.  $\operatorname{mean}_{WL}(S_F)$  is computed by

$$\operatorname{mean}_{WL}(S_F) = \frac{\sum_{k=1}^{|S_F|} w_k \times S_{F,k}^2}{\sum_{k=1}^{|S_F|} w_k \times S_{F,k}},$$

$$w_k = \frac{\Delta f_k}{\sum_{k=1}^{|S_F|} \Delta f_k},$$

$$\Delta f_k = f(x_{k,G}) - f(x_{k,G+1}).$$
(15)

In general, once new memories are updated, k = k + 1. If k > H, k = 1. If no new memories are provided, the memory of the current entry is not updated. Similarly, CR is generated by

$$CR_i = randn_i (M_{CR,ri}, 0.1),$$
(16)

where  $M_{\text{CR,ri}}$  is the memory of successful CR,  $ri \in [1, H]$ , H is the total number of memory entries, and  $\operatorname{rand} n_i(*)$  is the normal distribution function.  $M_{\text{CR},k}$  ( $k = 1, \dots, H$ ) is initialized to 0.5. From the second generation,  $M_{\text{CR},k}$  is updated by

$$M_{F,k,G+1} = \begin{cases} \operatorname{mean}_{WA}(S_{CR}), & \text{if } S_{CR} \neq \phi, \\ M_{CR,k,G}, & \text{otherwise,} \end{cases}$$

$$\operatorname{mean}_{WA}(S_{CR}) = \sum_{k=1}^{|S_{CR}|} w_k \times S_{CR,k}.$$
(17)

m

Since Au-DE can only solve the problem of minimization or maximization and the cleaned coal ash needs to be controlled within a given range [Ob - 1, Ob + 1], the optimization objective needs to be modified to min f'(x), where  $f'(x) = |f(x) + \mu_t - Ob|$ .  $\mu_t$  is the cleaned coal ash at the time t. This paper constructs the operation data of jig and each piece of manual experience as a  $(D+1) \times Y$  matrix, where D + 1 corresponds to a group of jig operation parameter (input) and a cleaned coal ash parameter (output), respectively, and Y is the number of entries in the data. The mapping model with a BP neural network is shown in Figure 4, where the hidden layer is connected by a sigmoid function, and the activation function of the output layer is linear function. The  $l_1^{\text{th}}$  hidden layer is denoted as  $h^{(l_1)}$ . During the training process, the weights W1 and bias b1, the maximum number of training times  $q_1$ , the hidden layer and its number of neurons, and the network learning rate lr are first initialized. Then, for the training samples, predicted values of the training samples are obtained by forward propagation  $z_1^{(l_1)} = W_1^{(l_1)}c_1^{(l_1-1)} + b_1^{(l_1)}$ . Calculate the prediction error  $L_1(W_1, b_1) = \sum_{t=1}^{N_1} (r_t - f(x))^2$ , and adjust the weights W1 and bias b1 along the negative gradient direction of the prediction error, where  $c_1^{(l_1)} = f_{l_1}(z^{(l_1)})$ ,  $c_1^{(0)} = (a_t, s_t)$ . f(x) is the output of the BP neural network,  $c_1^{(l_1-1)}$  is the input of  $l_1^{\text{th}}$  layer from the  $(l_1 - 1)^{\text{th}}$  layer, and  $W_1^{(l_1)}$  and  $b_1^{(l_1)}$  represent the weights and bias values of the  $l_1^{\text{th}}$  layer.  $N_1$  is the total number of samples participating in the training, and  $r_t$  is the actual performance value of the strategy. Finally, the training of the BP neural network is stopped when the error requirements  $L_1$  are met or the training time  $g_1$  is reached. Assuming that the parameter adjustment step size is  $\lambda$ , according to the current parameter set  $h_t$ , set the parameter search range to  $[h_t - k\lambda, h_t + k\lambda]$ , where k is the control parameter for search accuracy.

#### 3. Results

In this paper, taking a washing workshop as an example, we set the test running time of one month to verify the effectiveness of the proposed method. In Section 3.1, the parameter setting of DEIS system including the initial parameters, warning threshold parameters, and the feasible interval are given. In Section 3.2, the accuracy of jig fault warning in DEIS system is given. In Section 3.3, the average response time of cleaned coal ash from unsatisfying to satisfying production demand is given via the automatic control of DEIS system. The qualified rate of cleaned coal ash in the DEIS system is given in Section 3.4.

3.1. System Parameter Setting. In this paper, the empirical threshold parameters for the key module warning are set as follows. The height of the coal in bucket is less than 0.39 m, water tank liquid level is above 4.2 meters, the throughput frequency is controlled between 23 and 30 Hz, the optimal bed looseness is controlled between 0.4 and 0.45, the buoy height is set between 7 and 9, hydraulic cylinder pressure is greater than 1.6 MPA, wind pressure is set between 0.036 and 0.042 MPa, air valve intake period ratio is 23%–32%, air valve exhaust period ratio is 17%-25%, air valve expansion period ratio is 50%-60%, the critical value of gangue valve opening is 22, the critical value of medium coal valve opening is 18, and the critical value of the valve opening amplitude is 3. Due to the randomness in the working process of jig and the error of sensor data, judging whether a module is abnormal also needs to judge the duration of the jig fault. The default duration thresholds of all module parameters are set as follows: the lowest point of buoy is higher than the set value for 0.8 seconds, buoy stops beating for 16 seconds, bucket load exceeds limit for 18 seconds, gate keeps opening for 0.8 seconds, wind pressure is lower than the preset value for 16 seconds, hydraulic value is lower than the dedfault set for 18 seconds, water tank pressure value is lower than the default set for 0.8 minutes. The duration of the abnormal looseness is 18 minutes, the upper limit of the buoy height is 0.8 seconds, and the upper limit of the cleaned coal ash exceeds the limit of 10 minutes. In addition, considering that the coal height in bucket is obtained by machine vision and the coal shape is usually irregular, the accuracy of the computation result on



FIGURE 4: Surrogate model for mapping jig operating parameters and adjustment effectiveness of clean coal ash content.

coal height in bucket is significantly influenced. So, it is only considered as the abnormal bucket amount when the proportion of abnormal bucket is higher than 0.8 in multiple bucket lifting.

In addition, the intake period ratio of the 6 air chambers is initialized at 24%, 19%, 19%, 18%, 16%, 16%, and 21%, respectively; the throughput frequency is initialized at 27 Hz. All parameters are passed to the background data analysis model through the software interface. Adjustment step length of throughput frequency, buoy, intake (exhaust) period and expansion period, and valve opening is set to  $\lambda = 1$ ; search accuracy control parameter *k* is set to 3;  $\epsilon = 0.1$ .

3.2. The Warning Accuracy of DEIS. At present, the DEIS system mainly considers the overlimit warning, bed overturning warning, stuck buoy warning, stuck gate warning, and stockpile warning. During the specific test, according to the warning information displayed by the DEIS system, the warning information is verified manually. According to the proportion of warning types in the test duration given in Figure 5, it can be seen that the system warning in this period mainly includes overlimit warning, bed overturning warning, stuck buoy warning, and stuck gate warning. Stuck buoy and stuck gate are two main types of warning information, which exceed 50% of all warning information. Error warning refers to the warning issued by the DEIS system, but no abnormality is found in actual verification. Among all the warning results, the false warning accounted for 0.82%. The error warning is mainly caused by two reasons: (1) the data returned by the sensor have obvious error, which affects the accuracy of data analysis; (2) after the warning was issued, the jig returned to normal, resulting in inconsistent verification results.

3.3. Response Duration of the DEIS. To test the regulation ability of the system, we set the ideal interval of cleaned coal ash as [16, 18]. After implementating the regulation strategy, the time duration from the cleaned coal ash meeting requirements to being stable in the ideal interval, i.e., [16, 18] is the response time. Within each test day, three response durations of regulatory strategies were randomly recorded and the mean was calculated. It is worth noting that the



FIGURE 5: Statistics on the proportion of alarm types during the testing period.

recorded response time only considers the jig regulation under normal working conditions, and other special cases, such as insufficient coal quantity and jig faults, are not within the statistical range. As can be seen from the response duration given in Figure 6, the average response duration after the implementation of the regulatory strategies is not more than 40 minutes, and the specific response time is affected by the following main factors:

- (1) Under the influence of actual working conditions, the change trend of cleaned coal ash obtained by jig washing has different inertia. When the inertia is large, the control strategy adopted needs a long time to control the ash content of cleaned coal to the ideal range. The inertia here means that under the joint influence of multiple parameters of the jig, the cleaned coal ash has the property of continuously increasing, decreasing, or not easily changing.
- (2) The randomness of deep reinforcement learning and evolutionary computation makes the regulation strategy not unique. Although different control strategies have the effect of adjusting the ash content of cleaned coal, the control speed is obviously different, that is, the response time is different.
- (3) The quality of communication will affect the speed of receiving and transmitting real-time data, especially affect the system to perceive the trend change of cleaned coal ash in time. When the communication quality is good, the system can quickly sense the change of cleaned coal ash and make timely regulatory response; otherwise, it needs to adopt a longterm multistage regulatory strategy.

3.4. Qualified Rate of Cleaned Coal Ash. During the test cycle, the measured value of cleaned coal ash was recorded every day, the duration of cleaned coal ash at [16, 18] and the working duration of the jig were counted, and the qualified rate (percentage) of cleaned coal ash was characterized by the proportion of cleaned coal ash at [16, 18] and the

working duration of the jig, as shown in Figure 7. In the test period, the qualified rate of cleaned coal ash is between 32% and 64%, and the average qualified rate of cleaned coal ash is about 47%. On the whole, the qualified rate of cleaned coal ash has a large uncertainty and a large range of changes, mainly due to the following:

- (1) After the implementation of the control strategy, the cleaned coal ash cannot reach the expected range within the expected time, which is related to the operation inertia of the jig and has a high randomness. This leads to a long time of equipment control still can not make the cleaned coal ash to meet the requirements, even if it has a good trend.
- (2) The jig has been used for a long time, leading to wear or misalignment of some components, so that the control effect in the actual control process is lower than expected, and continuous regulation is needed to achieve the expected regulation effect, affecting the qualified rate of cleaned coal ash.
- (3) The failure of the jig installed in the sample collection will affect the normal operation of the jig. For example, the coal sampler may be stuck by the coal, resulting in the jig having to shut down the cleaning equipment, which affects the washing effect.

## 4. Discussion

Aiming at the problem of fault monitoring and intelligent adjustment of production state of jig in coal washing process, an intelligent washing method guided by deep reinforcement learning and evolutionary computing is proposed in this paper. In the test period, the proposed method has a shorter strategy response time and a higher washing quality adjustment ability. At the same time, in order to reduce the influence of jig faults on the washing efficiency, the equipment fault warning considered in the proposed method can accurately identify the fault of the key module and notify the worker to deal with it in time, which improves the washing efficiency to a certain extent. Although the response time of the strategy and the qualified



FIGURE 6: Statistical results of the response time of the control strategies.



FIGURE 7: Statistical results of qualified rate of cleaned coal wash.

rate of cleaned coal ash need to be further improved, it has certain reference significance for the construction of intelligent washing, especially the construction of expert system. Due to many uncertain factors in the actual production process, it is necessary to adopt advanced data interpolation technology to deal with the incomplete data received when the communication environment is poor and further explore the potential application value of the data.

## **Data Availability**

The data used to support the findings of this study are included within the article.

## **Conflicts of Interest**

The author declares that there are no conflicts of interest.

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant no. 62303465) and Natural Science Foundation of Shandong Province (grant no. ZR2022LZH017).

### References

 Z. Wang, Y. Kuang, J. Deng, G. Wang, and L. Ji, "Research on the intelligent control of the dense medium separation process in coal preparation plant," *International Journal of Mineral Processing*, vol. 142, pp. 46–50, 2015.

- [2] H. Villanueva and H. Lamba, "Operator guidance system for coal washing," *Artificial Intelligence in Engineering*, vol. 12, no. 3, pp. 261–273, 1998.
- [3] H. Villanueva, "AshMod: knowledge model based operator guidance system for coal washing," WIT Transactions on Information and Communication Technologies, vol. 16, 1970.
- [4] X. Yin, Z. He, Z. Niu, and Z. S. Li, "A hybrid intelligent optimization approach to improving quality for serial multistage and multi-response coal preparation production systems," *Journal of Manufacturing Systems*, vol. 47, pp. 199–216, 2018.
- [5] S. K. Das, K. Godiwalla, L. Panda, K. Bhattacharya, R. Singh, and S. Mehrotra, "Mathematical modeling of separation characteristics of a coal-washing spiral," *International Journal* of *Mineral Processing*, vol. 84, no. 1-4, pp. 118–132, 2007.
- [6] C. Zheng, "Technical architecture of big data cloud platform for intelligent washing factory," in *Proceedings of the 2021 4th International Conference on Artificial Intelligence and Big Data*, Xi'an, China, June 2021.
- [7] E. Zhou, G. Yan, X. Weng, Z. Zhang, P. Zhao, and B. Zhang, "A novel and low cost coal separation process: combination of deep screening classification and gravity separation," *Powder Technology*, vol. 367, pp. 568–575, 2020.
- [8] C. Zhou, X. Liu, Y. Zhao et al., "Recent progress and potential challenges in coal upgrading via gravity dry separation technologies," *Fuel*, vol. 305, Article ID 121430, 2021.
- [9] S. Sriramoju, D. Kumar, S. Majumdar, P. Dash, D. Shee, and R. Banerjee, "Sustainability of coal mines: separation of clean coal from the fine-coal rejects by ultra-fine grinding and density-gradient-centrifugation," *Powder Technology*, vol. 383, pp. 356–370, 2021.
- [10] T. Hirajima, H. Petrus, Y. Oosako, M. Nonaka, K. Sasaki, and T. Ando, "Recovery of cenospheres from coal fly ash using a dry separation process: separation estimation and potential

application," International Journal of Mineral Processing, vol. 95, no. 1-4, pp. 18-24, 2010.

- [11] D. Bandyopadhyay and J. Sen, "Internet of things: applications and challenges in technology and standardization," Wireless Personal Communications, vol. 58, no. 1, pp. 49–69, 2011.
- [12] S. Sagiroglu and D. Sinanc, "Big data: a review," in Proceedings of the 2013 International Conference on Collaboration Technologies and Systems, San Diego, CA, USA, July 2013.
- [13] Z. Wang, W. Zhao, W. Du, N. Li, and J. Wang, "Data-driven fault diagnosis method based on the conversion of erosion operation signals into images and convolutional neural network," *Process Safety and Environmental Protection*, vol. 149, pp. 591–601, 2021.
- [14] G. Wang, J. Fu, and L. Meng, "Development of innovation team construction and key technology research in coal mine intelligence," *Journal of Mine Automation*, vol. 48, no. 12, pp. 1–15, 2022.
- [15] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Playing atari with deep reinforcement learning," 2013, https://arxiv.org/abs/ 1312.5602.
- [16] R. Storn and K. Price, "Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [17] M. Zuo, D. Gong, Y. Wang, X. Ye, B. Zeng, and F. Meng, "Process knowledge-guided autonomous evolutionary optimization for constrained multiobjective problems," *IEEE Transactions on Evolutionary Computation*, vol. 52, 2023.
- [18] M. Zuo, G. Dai, L. Peng, M. Wang, Z. Liu, and C. Chen, "A case learning-based differential evolution algorithm for global optimization of interplanetary trajectory design," *Applied Soft Computing*, vol. 94, Article ID 106451, 2020.
- [19] M. Zuo, G. Dai, L. Peng, Z. Tang, D. Gong, and Q. Wang, "A differential evolution algorithm with the guided movement for population and its application to interplanetary transfer trajectory design," *Engineering Applications of Artificial Intelligence*, vol. 110, Article ID 104727, 2022.