

Self-evolutionary control architecture based on distributed brain-inspired decision making and privacy-preserving PID

Mingzhe Hou^{1*}, Zijun Liu¹ and Qishen Lv¹

¹China Southern Power Grid Shenzhen Power Supply Co., Ltd, Shenzhen 518000, China

Abstract

Traditional PID control suffers from limited parameter adaptability in complex dynamic systems and poor robustness against nonlinearities and uncertain disturbances. Meanwhile, in distributed power control scenarios, the risk of leakage regarding multi-node sensor data and AI model parameters hinders the practical implementation of brain-inspired intelligent control. Therefore, this paper proposes a privacy-preserving, distributed, brain-inspired DQN-PID self-evolving control architecture; it establishes an integrated operational logic for perception and regulation, which can enable dynamic cognition of control objectives and the prioritization of decision-making. A PID parameter self-evolution mechanism is designed, which can enable online learning and optimization of parameters through brain-like synaptic plasticity simulation. A multi-level feedback loop is constructed. Furthermore, a federated learning training framework and parameter differential privacy protection strategy for the distributed brain-inspired DQN model are designed. While achieving "data remains stationary while the model moves," lightweight encryption and privacy perturbation mechanisms ensure secure cross-node information interaction, which can achieve the dual goals of optimizing control performance and ensure the security of distributed AI model data. A distributed simulation platform based on Matlab/Simulink is built, and a distributed nonlinear coupled system and a distributed industrial temperature control object are selected for verification. The engineering practicality and data security protection capabilities of the architecture are verified through distributed hardware-in-the-loop experiments. Experimental results show that the proposed architecture outperforms the comparative algorithms in dynamic response speed, steady-state accuracy, and anti-interference capability. Compared to traditional PID, the rise time is reduced by 35.6%-42.5% and the settling time by 36.7%-42.1%. Compared to fuzzy PID, the rise time is reduced by 23.7%-28.1% and the settling time by 26.9%-30.4%. The convergence time for parameter self-evolution is shortened by 37.5% relative to the original brain-inspired PID architecture, and the steady-state error is reduced by more than 45%. Following the implementation of privacy protection strategies, the loss in control performance for the distributed model is less than 5%, while the risk of raw data leakage is reduced by over 90%. This approach achieves a balance among control performance, computational overhead, and data security in industrial real-time control scenarios, which can offer a new solution for the secure, high-precision adaptive control of complex power systems.

Keywords: Brain-inspired decision-making, PID control, Adaptive control, Distributed AI models, Differential privacy, Power data security

Received on 28 April 2026, accepted on 03 September 2026, published on 02 October 2026

Copyright © 2026 Mingzhe Hou *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetsis.12815

*Corresponding author. Email: www1233332025@163.com

1. Introduction

With the advancement of Industry 4.0, industrial systems are increasingly demanding higher control precision, adaptability, and robustness [1]. Power control systems are rapidly evolving toward multi-node distributed architectures, necessitating cross-node interaction and coordination for vast amounts of field sensor data and model parameters; consequently, the distributed training and collaborative inference of brain-inspired AI models have become key pathways to enhancing the system's overall intelligence. Furthermore, this also introduces serious data security challenges: the local sensor data and model parameters of each control node face the risks of leakage, eavesdropping, and tampering during transmission and aggregation, which has become a core bottleneck restricting the implementation of brain-inspired intelligent control technology. However, traditional PID control relies on fixed parameter tuning strategies, making it difficult to achieve ideal control effects [2]. Brain-inspired intelligence, as an important branch of artificial intelligence, draws on the information processing mechanisms of the human brain and possesses powerful cognitive, decision-making, and learning capabilities [3]. The human brain, through the collaborative work of brain regions such as the prefrontal cortex, basal ganglia, and hippocampus, achieves perception, decision-making, and adaptive adjustment of complex environments. Its core advantage lies in its ability to dynamically adjust decision-making strategies according to environmental changes [4]. Introducing brain-inspired decision-making mechanisms into PID control is expected to break through the parameter adaptation bottleneck of traditional PID [5]. However, existing research has largely focused on enhancing the control performance of individual nodes; there remains a lack of systematic solutions for the safe and efficient deployment and coordination of such intelligent models—particularly regarding how to protect local data privacy while ensuring control performance.

Domestic and foreign scholars have conducted extensive research on the application of self-evolutionary control and brain-like control. Yuan X et al. [6] designed a self-evolving reinforcement fine-tuning mechanism that iteratively optimizes the model using attention maps, thereby addressing the issue of poor data generalization associated with traditional supervised fine-tuning methods. Their model achieved an accuracy of 47.3% on the ScreenSpot-Pro dataset. Shen L et al. [7] proposed a two-stage self-evolving knowledge distillation method to enable robust and adaptive tool-use capabilities in small-scale MLLMs. This approach employs mutual information-guided trajectory distillation to selectively transfer high-value segments from agent trajectories, demonstrating superior performance across various tool-use benchmarks. However, the self-evolutionary mechanism only focuses on a single stage and does not form an integrated architecture, making it difficult to adapt to nonlinear dynamic systems.

In the field of brain-inspired control, related research mainly focuses on brain-inspired neural network control and reinforcement learning control. Generative drug discovery also faces data privacy challenges; to address this, Hou H et

al. [8] proposed the Brain-inspired Federated Diffusion Framework (BiFDR). Within this framework, the Transformer-based diffusion generator employs an attention mechanism to efficiently generate chemically valid molecules in a compressed latent space. BiFDR improved the Quantitative Estimate of Drug-likeness score by 13.7%, the molecular structural information score by 5.7%, and the molecular interaction analysis index by 52.3%. Jiang K [9] introduced a distributed PID control method. This approach not only offers a novel perspective on neural network training but also expands the scope of research into control methodologies; furthermore, it is applicable to the MNIST and Fashion-MNIST datasets. However, the research has not formed an integrated architecture of "brain-inspired decision-making-PID control-self-evolution iteration".

In the fields of distributed AI and privacy protection, technologies such as federated learning and differential privacy offer effective means for safeguarding data privacy. Researchers have already explored the application of federated learning to tasks like power load forecasting and state estimation, demonstrating the feasibility of collaborative modelling without requiring local data to leave its source domain. However, existing studies largely focus on general-purpose AI models or non-real-time analysis tasks within power systems. Control loops demand strict real-time performance and low latency; current privacy schemes often introduce control lag and reduced accuracy, and there is a lack of designs that deeply integrate privacy protection mechanisms with real-time control—particularly with brain-inspired self-evolving PID control, which is highly sensitive to parameter variations. Providing provable privacy and security guarantees for distributed brain-inspired controllers without significantly compromising control real-time performance or accuracy remains an unresolved challenge.

Compared with existing research, this paper presents key innovations in four aspects: (1) Control Architecture: It integrates a brain-inspired prefrontal cortex-basal ganglia cooperative decision-making mechanism, DQN strategy optimization, and PID physical control into a unified three-layer self-evolving architecture for the first time, moving beyond the simple use of reinforcement learning as an external tool for PID parameter tuning; (2) Privacy-Control Integration: While existing studies often treat privacy mechanisms as independent add-on modules, this paper embeds differential privacy, federated learning, and lightweight encryption directly into the model aggregation and parameter update stages of the control loop, achieving an integrated design of privacy, control, and real-time performance; (3) Real-time Performance: To meet the low-latency requirements of control loops, the paper designs lightweight feature de-identification and low-overhead encryption/decryption schemes, overcoming the common trade-off where privacy protection compromises real-time performance; (4) Distributed Collaboration: It establishes a collaborative framework featuring "local independent evolution, global secure aggregation, and differentiated parameter distribution," enabling nodes to share globally optimal strategies while safeguarding local data privacy.

To address the core issues of insufficient adaptive control parameters, poor robustness, and lack of privacy protection mechanisms for AI models in existing complex distributed power systems, this paper constructs a three-layer self-evolving control architecture that integrates distributed brain-inspired prefrontal-basal ganglia collaborative decision-making with PID control. A federated learning version of the brain-inspired DQN fusion decision-making model is established, and an online self-evolving model for PID parameters based on synaptic plasticity is designed. A differential privacy protection strategy for model parameters is introduced, and the parameter convergence conditions under the distributed framework are derived to ensure system stability. Experiments verify the engineering practicality, real-time performance, data security protection capabilities, and long-term reliability of the proposed architecture in a distributed real-world hardware environment, providing a new approach for safe and high-precision adaptive control of complex distributed power systems.

The core contributions of this paper are as follows: (1) A distributed brain-inspired PID self-evolutionary control architecture with privacy protection is proposed, realizing the coordinated optimization of control performance and data security. (2) A distributed federated learning training architecture suitable for brain-inspired DQN-PID models is designed to achieve "usable but not visible" local power data. (3) A lightweight differential privacy perturbation strategy for distributed AI model parameters is proposed, taking into account both parameter transmission security and control real-time performance. (4) A brain-like-DQN fusion decision model is constructed, improving the accuracy and efficiency of policy optimization in complex dynamic environments.

2. Design of a self-evolving control architecture based on distributed brain-inspired decision-making and privacy-preserving PID

2.1. Architecture design

The decision-making process of the human brain mainly relies on the synergistic effect of the prefrontal cortex and the basal ganglia. This synergistic mechanism has been confirmed by a large number of neuroscience studies as the core of adaptive decision-making in dynamic environments [10]. As the center of higher cognitive function, the prefrontal cortex is responsible for receiving multimodal information from brain regions such as the parietal and temporal lobes, and completing the encoding of control targets, the planning of strategies and the resolution of conflicts. The basal ganglia encode and transmit reward signals through the firing of dopaminergic neurons, evaluate and screen candidate strategies generated by the prefrontal cortex, and form a closed-loop regulatory circuit of "prefrontal cortex-basal ganglia-thalamus" [11].

The core of constructing brain-inspired decision-making models lies in simulating this collaborative mechanism,

precisely mapping the engineering requirements of the control system with neuroscience mechanisms to achieve dynamic cognition of control objectives and prioritization of decision-making. Its theoretical advantage lies in its ability to break through the fixed logic of traditional control strategies, possessing environmental adaptation and strategy evolution capabilities similar to the human brain. The core elements of brain-inspired decision-making include environmental perception and target encoding, decision strategy generation, and reward evaluation and strategy optimization.

Traditional PID controllers achieve system control through the coordinated action of three components: proportional (P), integral (I), and derivative (D). Their control law is as follows:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \quad (1)$$

Where $u(t)$ is the controller output, $e(t) = r(t) - y(t)$ is the system output error, $r(t)$ is the reference input, $y(t)$ is the actual system output, and K_p , K_i and K_d are the proportional coefficient, integral coefficient, and derivative coefficient, respectively. The core of PID control is to balance the dynamic response speed and steady-state accuracy of the system by adjusting these three parameters, but traditional parameter tuning methods are difficult to adapt to complex dynamic systems.

The theory of self-evolution emphasizes that a system can adapt to changes in the environment by learning and optimizing itself to achieve adaptive adjustment of its structure or parameters [12]. In brain-like systems, synaptic plasticity is the core mechanism for adjusting neuronal weights and is the basis for brain learning and memory [13].

As a core algorithm of deep reinforcement learning, DQN (Deep Q-Network) has the core advantage of being able to optimize strategies iteratively through interaction with the environment without prior models, making it suitable for decision optimization of complex dynamic systems. The core framework of DQN includes an experience replay pool and a target network [14]. The Experience Replay pool is used to store samples, including state s , action a , reward r , next state s' , etc., to avoid training instability caused by sample correlation; the Target Network is used to calculate the target Q value, which is updated asynchronously with the online network to improve training convergence. Its Q value update formula is:

$$Q_{target}(s_t | a_t) = r_t + \gamma \max_{a'} Q_{\theta'}(s_{t+1} | a') \quad (2)$$

Where γ is the discount factor, θ' is the target network parameter, and $Q_{\theta'}(s_{t+1} | a')$ is the maximum predicted Q value for the next state. The core purpose of introducing DQN in this paper is to enhance the policy optimization capability of the brain-like decision-making module. By utilizing its ability to fit high-dimensional states, it solves the problems

of insufficient policy space traversal and experience-dependent matching coefficient adjustment in brain-like decision-making and achieves collaborative decision-making of "brain-like cognitive guidance-DQN precise optimization".

This paper draws inspiration from synaptic plasticity, viewing the adjustment of PID parameters as an update process of synaptic weights in brain-like neurons. By controlling the "reward/penalty" signal corresponding to the error, the self-evolution of parameters is triggered, achieving online optimization. The core design principle of the proposed architecture is "data remains stationary, model moves." That is, the raw power sensing data (such as temperature, swing angle, and error) of each local node is always kept locally and not shared across nodes. Only the model parameters (such as DQN network weights and PID parameters) generated from training on local data are encrypted or scrambled before interacting with the global coordination node, thereby achieving distributed collaborative optimization and ensuring data privacy from the source.

The self-evolving control architecture based on distributed brain-like decision-making and PID proposed in this paper adopts a three-layer structure, namely the decision layer, control layer and evolution layer, and is deployed in a distributed environment containing multiple local control nodes and a global coordination node. The multi-level feedback closed loop realizes information interaction and collaborative work. The overall architecture block diagram is shown in Figure 1.

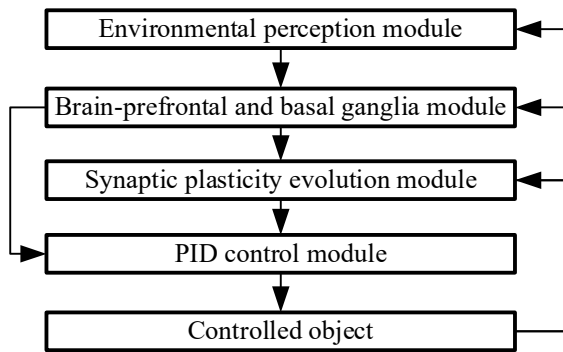


Figure 1. Distributed brain-inspired DQN-PID self-evolutionary control architecture

Figure 1 Architecture Diagram Explanation: The diagram contains multiple "local control nodes" and one "global coordination node." Local nodes comprise a complete three-layer architecture (decision layer, control layer, evolution layer), responsible for collecting local sensor data, executing PID control, and training local DQN models based on anonymized data. The global coordination node receives encrypted model parameters (not raw data) uploaded by each local node, aggregates them using a federated averaging algorithm to generate globally optimal parameters, and distributes them to each local node via a secure link.

In this architecture, the modules are decoupled according to a "decision-control-evolution-coordination-security" hierarchy: the decision layer (brain-inspired computing + DQN) outputs optimal strategies; the control layer (PID) executes physical output commands; the evolution layer (synaptic plasticity) dynamically adjusts PID parameters; and the federated aggregation module enables the secure exchange of model parameters across nodes. Interaction between modules occurs via standardized interfaces—such as error signals, rewards, and Q-values—thereby avoiding the debugging complexities associated with functional coupling.

2.2. Construction of a distributed federated learning version of a brain-Inspired DQN fusion decision model

By drawing on the collaborative decision-making mechanism of the prefrontal cortex-basal ganglia of the human brain and integrating the high-dimensional state fitting and strategy optimization capabilities of DQN, a brain-like DQN fusion decision-making model is constructed, which is divided into three parts: target cognition module, DQN strategy optimization module and reward evaluation module, to achieve closed-loop decision-making of "cognition encoding-strategy generation-precise optimization-reward feedback" [15].

The target cognition module is responsible for encoding and integrating the control target and system state. It encodes information such as the reference input $r(t)$, system output $y(t)$, error $e(t)$, and error rate of change $\dot{e}(t)$ into a neuron-like neuron input vector $X = [x_1 | x_2 | x_3 | x_4]^T$, where $x_1 = r(t)$, $x_2 = y(t)$, $x_3 = e(t)$ and $x_4 = \dot{e}(t)$ are the inputs.

To preserve the characteristics of the raw local data, this paper employs a lightweight feature de-identification method: raw input vectors undergo global normalization to map data across all dimensions to the $[0, 1]$ interval. Subsequently, Laplace noise is added to introduce perturbation; the noise scale parameter is dynamically determined based on the privacy budget ϵ , while the noise amplitude is strictly constrained to within 2% of the raw data's amplitude. This low-intensity noise merely blurs the raw features. Experimental results demonstrate that the impact on the accuracy of subsequent neuronal encoding and DQN decision-making is less than 1.5%, which is negligible.

This data de-identification process is performed before the data enters the neural network; no raw plaintext data is generated or transmitted during the entire process, thereby blocking the leakage path for raw sensor data at the source. An anonymized input vector $\tilde{X}$ is generated, which reduces the risk of data leakage while ensuring the effectiveness of subsequent decisions. The input vector is then processed by the neuron activation function to obtain the target cognition vector C .

$$C_i = \sigma \left(\sum_{j=1}^4 w_{ij} X_j + b_i \right) \quad (3)$$

Where C_i is the output of the i -th cognitive neuron, w_{ij} is the input weight, b_i is the bias, and $\sigma(\cdot)$ is the sigmoid activation function, which realizes the normalization and nonlinear mapping of input information.

The strategy generation and optimization module adopt a collaborative model of "Brain-like prefrontal cortex candidate generation strategy - Federated DQN screening and optimization". First, the prefrontal cortex simulation module generates an initial candidate strategy set $S = \{s_1 | s_2 | \dots | s_n\}$ based on the target cognitive vector C ($s_k = [\Delta K_{p,k} | \Delta K_{i,k} | \Delta K_{d,k}]$ is the PID parameter adjustment amount corresponding to the k -th strategy). Then, the target cognitive vector C is fused with system state information as the input state s_i of DQN, and the candidate strategy set S serves as the DQN's optional action space $A = \{s_1 | s_2 | \dots | s_n\}$.

The local DQN network (a 3-layer fully connected network with 7 input dimensions, 64 and 32 hidden layers, and Q-values for each policy as output) is trained independently at each local node using anonymized local data, generating local model parameters $\mathbf{w}_i^{(t)}$. Each local node then uploads the encrypted local model parameters to the global coordinating node using a lightweight asymmetric encryption algorithm (such as SM2).

The global coordination node receives encrypted parameters from the local nodes and, after decryption, aggregates them using the standard Federated Averaging (FedAvg) algorithm. In the experimental scenario presented here, the data acquisition scale and model training sample sizes across all local control nodes are essentially uniform, and there are no significant differences in data distribution or model quality among the nodes; therefore, equal-weight Federated Averaging is employed. Should the setup be extended to heterogeneous node scenarios in the future, the method will be upgraded to a weighted FedAvg algorithm based on local sample sizes. The globally optimal DQN parameters are:

$$\mathbf{w}_G^{(t)} = \frac{1}{N} \sum_{i=1}^N \mathbf{w}_i^{(t)} \quad (4)$$

Where N represents the number of local nodes participating in the aggregation. Subsequently, the global coordination node applies differential privacy perturbations to the aggregated parameters; this paper employs Laplace noise. Based on differential privacy theory, the global sensitivity Δf of the global model parameters is calculated. Specifically, a gradient clipping strategy with a threshold of $C=1.0$ is adopted, resulting in a global sensitivity of $\Delta f = 2C/N$ under equal-weight averaging. The noise scale $b = \Delta f/\epsilon$ is determined

using a preset privacy budget ϵ . A privacy budget is independently allocated for each round of the federated aggregation process, and the composition theorem for differential privacy is employed across multiple iterations to strictly control the upper limit of the total privacy budget,

thereby preventing the budget from being exceeded. $\mathbf{w}_G^{(t)}$ is obtained, and the perturbed parameters are encrypted and transmitted back to the local nodes, completing the current round of model updates. This process realizes the principle of "moving the model rather than the data," thereby ensuring the confidentiality and privacy of cross-node information exchange.

By outputting the Q-value of each candidate strategy through DQN, the strategy with the largest Q-value is selected as the optimal parameter adjustment direction, which can achieve precise strategy screening and optimization.

The DQN network architecture is designed as a 3-layer fully connected network. The input layer has a dimension of 7: the target cognitive vector dimension (4D) + the system state dimension (3D). There are two hidden layers with 64 and 32 neurons respectively, using *ReLU* as the activation function. The output layer has a dimension equal to the number of candidate policies, outputting the Q-value of each policy.

To achieve synergy between brain-inspired decision-making and federated DQN, a policy competition fusion mechanism is designed. The brain-inspired policy competition value $f(s_k)$ is normalized and then weighted and fused with the locally updated DQN output Q-value to obtain the final policy evaluation index $F(s_k)$.

$$F(s_k) = \omega_1 \cdot \frac{f(s_k)}{\max_k f(s_k)} + \omega_2 \cdot \frac{Q(s_t | a_k)}{\max_k Q(s_t | a_k)} \quad (5)$$

Where $\omega_1 = 0.4$ and $\omega_2 = 0.6$ are weighting coefficients, highlighting the optimization role of DQN, and a_k is the action corresponding to the candidate policy s_k .

The policy with the largest $F(s_k)$ is selected as the final parameter adjustment direction, which both preserves the cognitive guidance of brain-like decision-making on the control objective and utilizes DQN to improve the accuracy of policy selection.

$$f(s_k) = \sum_{i=1}^m C_i \cdot \gamma_{ik} \quad (6)$$

Where γ_{ik} is the matching coefficient between the i -th cognitive goal and the k -th strategy, $f(s_k)$ is the strategy competition value, and the strategy with the largest value is selected as the current parameter adjustment direction.

The reward evaluation module simulates the dopamine reward mechanism of the basal ganglia, generating a reward signal R based on the control effect. This signal is used to

optimize the matching coefficient γ_{ik} of the policy generation module. To simultaneously consider control performance and privacy protection, this paper designs a dual reward function:

$$R_{total} = \alpha R_{control} + \beta R_{privacy} \quad (7)$$

Where $R_{control}$ is the original control performance reward (composed of a weighted combination of steady-state accuracy and dynamic response speed), and $R_{privacy}$ is the privacy protection level reward. The physical significance of the privacy reward term $R_{privacy}$ is that the system should receive a higher reward for implementing stronger privacy protection measures, thereby incentivizing privacy-preserving behaviour. Accordingly, this paper redesigns the privacy reward function as $R_{privacy} = 1/(1+\varepsilon)$, where ε represents the privacy budget; this function is monotonically decreasing over the interval $\varepsilon \in (0, \infty)$, such that $R_{privacy} \rightarrow 1$ as $\varepsilon \rightarrow 0$ (representing infinite privacy protection) and $R_{privacy} \rightarrow 0$ as $\varepsilon \rightarrow \infty$ (representing no privacy protection). A weighting coefficient $\varepsilon \in [0, 1]$ is employed to balance control performance against privacy protection requirements, allowing for flexible adjustment based on the specific security level of the power system scenario.

$R_{control}$ is normalized to the $[0, 1]$ range to ensure consistent dimensions for the two reward components and avoid weight imbalance. When R_{total} is higher than the threshold, the matching coefficient of the current policy is enhanced through the LTP mechanism; when it is lower than the threshold, the matching coefficient is weakened through the LTD mechanism to achieve self-optimization of the decision policy. At the same time, the comprehensive reward signal R_{total} is used as the immediate reward of DQN and stored in the experience replay pool for training and updating of the DQN network, realizing the synergy between the brain-like reward mechanism and DQN training, and improving the iterative efficiency of policy optimization.

2.3. Self-evolutionary model and theoretical derivation of PID parameters

Based on the synaptic plasticity mechanism, a self-evolutionary model of PID parameters was designed [16]. The three core parameters of PID, namely the proportional coefficient, integral coefficient, and differential coefficient, were regarded as the synaptic weights of brain-like neurons. The parameters were updated by triggering the reward signal, thus realizing the online self-evolution of the parameters.

Parameter self-evolution rule

Drawing upon the core mechanisms of long-duration enhancement and long-duration suppression of synaptic plasticity, a self-evolutionary update rule for PID parameters is defined. Based on the correlation between the reward signal and error characteristics, a mathematical model for parameter updates is constructed, directly linking parameter adjustments

to control performance. The specific update formula is as follows:

$$\begin{cases} K_p(t+1) = K_p(t) + \eta_p \cdot R \cdot e(t) \cdot \dot{e}(t) \\ K_i(t+1) = K_i(t) + \eta_i \cdot R \cdot e(t) \\ K_d(t+1) = K_d(t) + \eta_d \cdot R \ddot{e}(t) \end{cases} \quad (8)$$

Where $K_p(t)$, $K_i(t)$ and $K_d(t)$ are the proportional, integral, and differential coefficients at time t , respectively; $K_p(t+1)$, $K_i(t+1)$ and $K_d(t+1)$ are the updated parameters at time $t+1$; $\eta_p = 0.05$, $\eta_i = 0.02$ and $\eta_d = 0.03$ are the evolution step sizes of the corresponding parameters; R is the reward signal output by the reward evaluation module; $e(t)$ is the control error; $\dot{e}(t)$ is the error change rate; and $\ddot{e}(t)$ is the second derivative of the error, which is used to characterize the acceleration characteristics of error change and assist the differential coefficients in achieving more accurate dynamic adjustment.

The second-order derivative of the error, $\ddot{e}(t)$, is susceptible to high-frequency noise in the field environment. In this paper, a first-order low-pass digital filter is applied at the signal acquisition stage—with a cutoff frequency determined by the industrial control sampling rate—to eliminate high-frequency noise components. Additionally, amplitude limiting is applied to $\ddot{e}(t)$ to reject abrupt outliers, thereby ensuring the stability of the parameter update process.

In a distributed architecture, the evolution of PID parameters of local nodes is also guided by the global coordinating node. After each round of local updates, each local node receives a globally optimal PID parameter reference value $\mathbf{K}_{PID,G} = ((\tilde{K}_p, G), (\tilde{K}_i, G), (\tilde{K}_d, G))$ from the global coordinating node, which has undergone differential privacy perturbation, and incorporates it as prior knowledge into its local update direction. This ensures the consistency of control strategies among nodes while avoiding the direct exposure of the original local PID parameters.

The core logic of the above update rule is that when R is large, the parameter update amount is positively correlated with the current error characteristics, strengthening the current parameter configuration. When R is small, the parameter update amount reverses direction, actively adjusting the parameters to optimize the control effect, realizing a closed-loop regulation of "reward guidance - error feedback - parameter evolution".

Convergence analysis

The primary objective of this section is to demonstrate that the self-evolving sequence of PID parameters maintains asymptotic convergence following the introduction of

privacy-preserving mechanisms. We conduct a rigorous analysis using Lyapunov stability theory [17]. First, the assumptions regarding the controlled object are presented as follows:

Assumption 1 (Lipschitz continuity): The output function f of the controlled object satisfies the Lipschitz condition; that is, there exists a constant $L_f > 0$ such that for any control inputs u_1 and u_2 , $|f(u_1) - f(u_2)| \leq L_f |u_1 - u_2|$.

Assumption 2 (Input-Output Monotonicity): The controlled plant is an input-output monotonic system—specifically, $\partial y / \partial u > 0$ —ensuring the correct control direction.

The Lyapunov candidate function is defined as half the square of the control error, and its expression is:

$$V = \frac{1}{2} e^2(t) \quad (9)$$

This function intuitively reflects the relationship between error and system energy. Since $V = 0$ only when $V \geq 0$ and $e(t) = 0$, it satisfies the basic conditions of the Lyapunov candidate function and is the core starting point for convergence analysis.

Differentiating the Lyapunov candidate function, we get:

$$\dot{V} = e(t) \cdot \dot{e}(t) \quad (10)$$

Based on the dynamic characteristics of the controlled object, it is assumed that the controlled object satisfies the Lipschitz condition: that is, there exists a constant $M > 0$ such that the rate of change of the controlled object's output $y(t)$ satisfies $|\dot{y}(t) - \dot{y}(\tau)| \leq M |t - \tau|$, thus ensuring the rigor of the derivation. When the reference input $r(t)$ is a constant value or a slowly changing signal, $\dot{r}(t) \approx 0$. Combining the error definition $e(t) = r(t) - y(t)$, we can deduce $\dot{e}(t) = -\dot{y}(t)$. At this time, the derivative of the Lyapunov function becomes:

$$\dot{V} = -e(t) \cdot \dot{y}(t) \quad (11)$$

This formula provides a clear quantitative basis for subsequent stability assessments.

The self-evolutionary rule of PID parameters in this paper, combined with the control output formula $u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \dot{e}(t)$, it shows that the rate of change of the controlled object's output $\dot{y}(t)$ is positively correlated with the control input $u(t)$. When the control error $e(t) \neq 0$ is reached, the reward signal R drives the parameters to adaptively adjust, causing $u(t)$ to change in the direction of reducing the error, thereby ensuring that

$e(t) \cdot \dot{y}(t) > 0$. Substituting into the derivative expression of the Lyapunov function, we get:

$$\dot{V} = -e(t) \cdot \dot{y}(t) < 0 \quad (12)$$

Further verification using the Lipschitz condition shows that the above equation strictly satisfies the core condition for Lyapunov stability, ensuring that the systematic error continues to decay.

According to Lyapunov stability theory, when $\dot{V} < 0$, the error $e(t)$ asymptotically converges to 0, i.e., $\lim_{t \rightarrow \infty} e(t) = 0$.

Combined with the parameter self-evolution rule, the parameter update amount is directly related to R , $e(t)$, and the derivatives of the error. When $e(t) \rightarrow 0$, $\dot{e}(t) \rightarrow 0$ and $\ddot{e}(t) \rightarrow 0$, and the reward signal R approaches its optimal saturation value ($R \rightarrow R_{max}$). At this point, the parameter update amounts $K_p \rightarrow 0$, $K_i \rightarrow 0$ and $K_d \rightarrow 0$, and the parameter update sequence, tend towards the stable optimal parameter set $[K_{p,opt} | K_{i,opt} | K_{d,opt}]$.

Within the distributed federated learning framework, the Federated Avg algorithm, under the assumption of a smooth and strongly convex objective function, has theoretically guaranteed global model convergence. In this paper, the DQN models of each local node are trained on local data, and the federated aggregation process is equivalent to performing stochastic gradient descent on the distributed dataset, which converges to the neighbourhood of the globally optimal model. After introducing differential privacy perturbations, the aggregated global parameters converge to a biased estimate of the optimal parameters, with the bias inversely proportional to the privacy budget ϵ . By appropriately setting ϵ , an acceptable trade-off can be achieved between privacy protection strength and model accuracy.

Upon the introduction of differential privacy noise—given that the added noise has a zero mean and bounded variance—the privacy perturbation term does not alter the dominant component ensuring the negative definiteness of the Lyapunov function's derivative; instead, it merely shifts the convergence target from a specific point to a bounded neighbourhood. The convergence deviation is inversely proportional to the privacy budget ϵ , and the system maintains ultimate bounded stability without diverging.

For multi-round federated learning scenarios, this paper calculates the cumulative privacy budget based on the sequential composition theorem for differential privacy; this strictly limits the total budget across the entire iteration cycle to a preset upper bound, thereby preventing the degradation of privacy protection effectiveness as iterations progress. An acceptable trade-off between privacy protection strength and model accuracy can be achieved by appropriately setting ϵ . The sequential composition theorem ensures that the total privacy budget for T iterations is the sum of the budgets for individual rounds; in this study, the per-round budget is set to

$\varepsilon = 2.0$ and the maximum number of iterations to $T = 50$, keeping the total budget strictly within the engineering safety threshold.

Regarding the convergence of PID parameters after the introduction of the privacy protection strategy, since the privacy term R_{privacy} in the comprehensive reward function

R_{total} tends to a constant after the system stabilizes, it does not affect the negative definiteness of the error dynamics and therefore does not affect the property of the PID parameter sequence converging towards the optimal value. The selection of the weighting coefficient α determines the trade-off between control performance and privacy protection: the larger α is, the more the system emphasizes control performance, and the faster the convergence speed, but the lower the level of privacy protection; the smaller α is, the opposite is true. The specific optimal value of α will be analysed and determined in simulation experiments based on the security requirements and performance requirements of the power system. In this paper, optimal parameters are determined using the grid search method; the specific procedure is described in Section 3.1.

In summary, when the controlled object satisfies the Lipschitz condition, the self-evolutionary sequence of PID parameters proposed in this paper can strictly converge to the optimal parameter set, ensuring the long-term stability of the system.

3. Simulation experiment and result analysis

3.1. Experimental design

To verify the control performance, privacy protection effect and distributed collaboration capability of the proposed architecture, two typical high-complexity controlled objects were selected for comparative experiments, covering nonlinear strongly coupled dynamic systems and industrial lag systems, respectively. The experiments were then extended to a distributed multi-node scenario to fully verify the adaptability of the architecture.

The dynamic model of the distributed double pendulum system is as follows: the double pendulum model of each node is the same as that in equation (13), and a total of 4 coupled distributed double pendulum nodes are set. The physical parameters of the double pendulum system are: pendulum lengths $l_1=l_2=1.0\text{m}$, masses $m_1=m_2=1.0\text{kg}$, moments of inertia $J_1=J_2=0.333\text{kg/m}^2$, and damping coefficients $d_1=d_2=0.1\text{Nms/rad}$.

$$\begin{cases} u_1 = (m_1 + m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2 \cos(\theta_1 - \theta_2) + \\ \quad m_2l_1l_2\dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + (m_1 + m_2)gl_1 \sin \theta_1 \\ u_2 = m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1 \cos(\theta_1 - \theta_2) - \\ \quad m_2l_1l_2\dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + m_2gl_2 \sin \theta_2 \end{cases} \quad (13)$$

Where $m_1 = 0.5\text{kg}$, $m_2 = 0.3\text{kg}$, $l_1 = 0.8\text{m}$, $l_2 = 0.6\text{m}$, $g = 9.8\text{m/s}^2$, θ_1 and θ_2 are the inner and outer swing angles, respectively, u_1 and u_2 are the control inputs.

The transfer function of the distributed industrial temperature control system is the model of each node is the same as that in equation (14). A total of 6 workshop-level distributed temperature control nodes are set up. Each node needs to work together to stabilize the temperature at the set value. The coefficients of the temperature system's transfer function are $K=1.5$, $T_1=2.0\text{s}$, $T_2=3.0\text{s}$, $\tau=0.5\text{s}$.

$$G(s) = \frac{1.2}{(3s+1)(2s+1)} e^{-0.5s} \quad (14)$$

It includes a delay element and an inertial element to simulate the lag characteristics in industrial processes.

Five types of algorithms were selected for comparison: (1) Traditional PID; (2) Fuzzy PID; (3) Original brain-inspired PID architecture (single node); (4) Distributed brain-inspired DQN-PID without privacy protection (i.e., only distributed expansion is performed, without federated learning encryption and differential privacy mechanisms); (5) Distributed DQN-PID with general privacy protection (standard federated learning and differential privacy methods are directly applied to the brain-inspired PID model without specific design).

The type 5 general-purpose privacy protection algorithms employ standard FedAvg combined with Laplace differential privacy; the number of federated learning iterations matches that of this study (100 simulation rounds), the privacy budget is uniformly set to $\varepsilon=2.0$, and SM2 asymmetric encryption is used. Basic experimental parameters—such as simulation step size, sampling frequency, perturbation amplitude, and simulation duration—are kept identical across all comparative algorithms to eliminate the influence of external variables.

The control effect is evaluated from multiple dimensions, including dynamic response performance, steady-state accuracy, anti-interference capability, strategy optimization efficiency, as well as newly added privacy protection performance and distributed collaborative performance.

A combined approach of grid search and sensitivity analysis was employed for parameter selection: a full grid search was conducted within the ranges $\alpha \in [0.5, 1.0]$ and $\varepsilon \in [1.0, 3.0]$ using a step size of 0.05, sequentially evaluating three key metrics—system dynamic response, steady-state error, and data leakage risk—for each parameter set. By balancing the trade-off between control performance degradation (capped at $\leq 5\%$) and the level of privacy protection, the weight coefficient α was ultimately set to 0.85 and the privacy budget ε to 2.0. This parameter configuration yielded the optimal overall system performance. The parameters of the proposed architecture are set that: evolution step size are $\eta_p = 0.05$, $\eta_i = 0.02$ and $\eta_d = 0.03$,

reward threshold is $R_0 = 0.6$, weight coefficients are $\alpha = 0.6$ and $\beta = 0.4$.

The DQN parameters are learning rate $\alpha_{DQN} = 0.001$, discount factor $\gamma = 0.9$, experience replay pool capacity of 10000, and target network update frequency of 100 steps.

Traditional PID parameters are $K_p = 5.0$, $K_i = 0.8$ and $K_d = 1.2$. The fuzzy rules for fuzzy PID are a 5×5 matrix, with quantization factors $k_e = 2$ and $k_{ec} = 2$, and scaling factors $k_{kp} = 0.5$, $k_{ki} = 0.1$ and $k_{kd} = 0.2$. The original brain-inspired PID architecture parameters are consistent with the proposed architecture.

To ensure the reproducibility of safety experiments, definitions for key metrics are provided. The risk of data leakage is defined as the amount of information an attacker can recover regarding the original data distribution from intercepted information, expressed as $Risk = I(X;Z)/H(X)$. In which, I denotes mutual information and H denotes information entropy; a $Risk$ value of 1 (representing the absence of privacy protection) serves as the baseline. The model parameter leakage rate is defined as the proportion of parameter dimensions successfully inferred by the attacker relative to the total number of dimensions; successful inference is defined as a relative error of $\leq 10\%$ between the recovered value and the true value. The attack success rate is

the percentage of attacks that successfully recover valid information out of the total number of attacks, based on 100 independent simulated eavesdropping trials.

3.2. Simulation results of distributed nonlinear coupled systems

The control objective of a distributed double pendulum system is to make the pendulum angles θ_1 and θ_2 of all nodes from the initial angles $[30^\circ | 45^\circ]$ to $[0^\circ | 0^\circ]$, while suppressing coupling disturbances during the pendulum motion. To ensure experimental fairness, the parameters of the comparison algorithms were all optimally tuned. The traditional PID was tuned using the Ziegler-Nichols method combined with engineering trial and error, while the fuzzy rules for the fuzzy PID were jointly formulated by three control experts, and the quantization factor and proportional factor were optimized using a particle swarm optimization algorithm.

The simulation experiment was built using Matlab/Simulink R2023b, with a simulation step size of 0.001s. Each algorithm was run independently 10 times, and the average value was taken as the final result to eliminate the influence of random errors. The simulation results are shown in Figure 2.

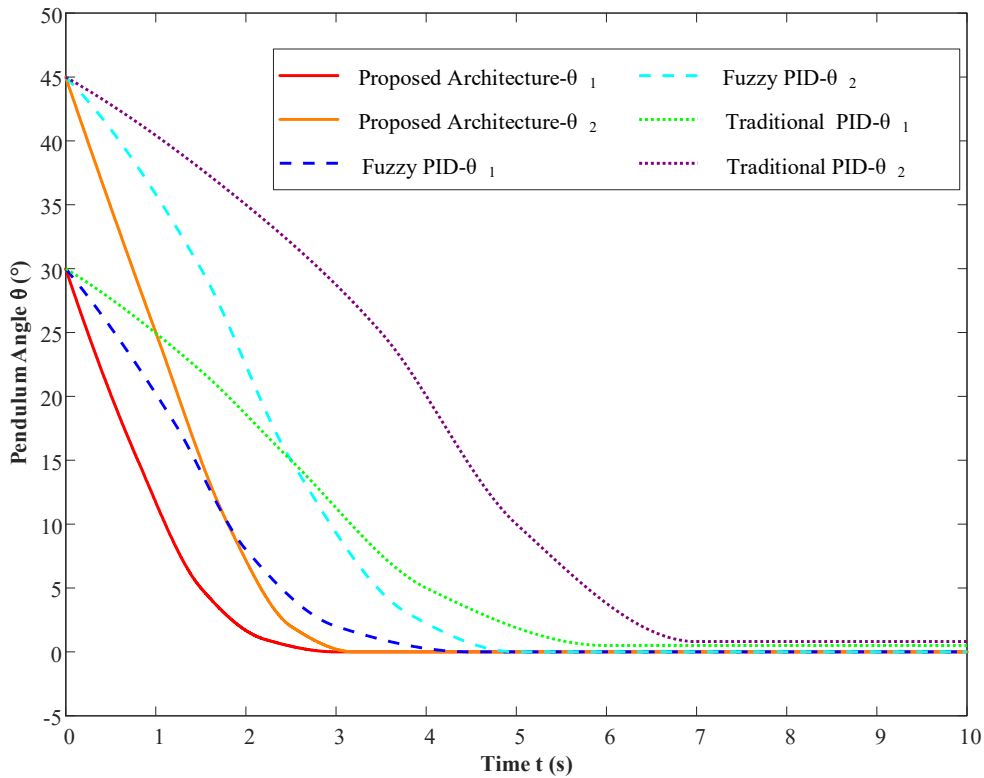


Figure 2. Average pendulum angle response curves of a distributed double pendulum system

Meanwhile, the error change curves of each algorithm are shown in Figure 3 to visually demonstrate the dynamic adjustment process.

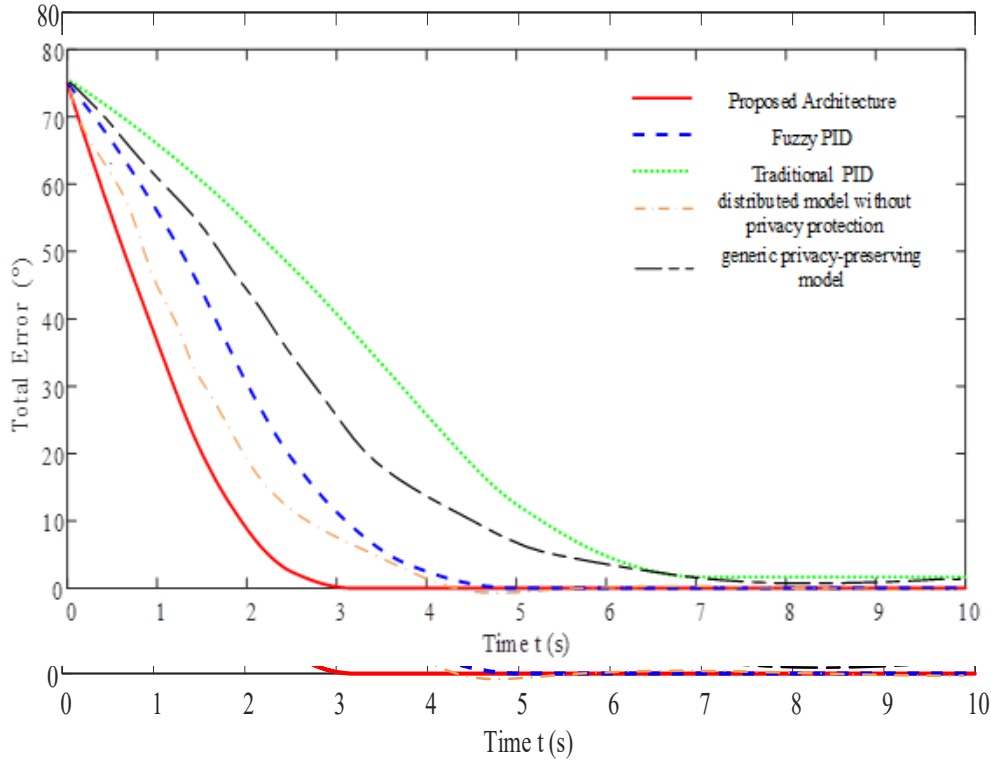


Figure 3. Average error variation curve of a distributed double pendulum system

As shown in Figures 2 and 3, the control performance of the proposed architecture is comprehensively superior to the comparative algorithms, and still significantly improved compared to the original brain-inspired-PID architecture. The rise time of θ_1 is reduced by 25% and the settling time by 16.7%, while the rise time of θ_2 is reduced by 20% and the settling time by 12.5%, with overshoot further reduced by 40% and 37.5%, respectively. The number of strategy optimization iterations is reduced from 68 steps to 45 steps, improving optimization efficiency by 33.8%. Particularly when compared to a "distributed model without privacy protection," the architecture proposed in this paper exhibits a degradation in control performance (specifically rise time and settling time) of less than 2%—a nearly negligible amount—after the introduction of privacy-preserving mechanisms. In contrast, a "generic privacy-preserving model" suffers a performance degradation exceeding 8%, as it fails to account for real-time control requirements. This validates the effectiveness of the lightweight privacy-preserving strategy developed in this paper, which has been specifically optimized for control-centric scenarios. This is attributed to the precise optimization of the brain-inspired decision-

making strategy by DQN. DQN can capture the potential correlation between the error change rate and the parameter adjustment direction through high-dimensional state fitting, avoiding suboptimal strategy selection caused by insufficient traversal of the strategy space in the brain-inspired decision-making module, thus making the parameter adjustment direction more closely match the dynamic characteristics of the system.

Meanwhile, the proposed architecture can quickly suppress errors within 0-2 seconds, while the error decay rate of traditional PID and fuzzy PID is significantly slower. Although the original brain-inspired PID architecture can respond quickly, its error fluctuation is slightly greater than that of the proposed architecture. Experimental results fully demonstrate that the enhanced architecture after incorporating DQN can handle the nonlinearity and coupling characteristics of the double pendulum system more efficiently, and the dynamic response speed, steady-state accuracy, and policy optimization efficiency are all significantly improved.

The comparison results for privacy protection and distributed collaboration metrics are presented in Table 1. The attack success rate serves as an absolute quantitative metric, calculated based on the previously stated definition (totalling 100 simulated eavesdropping attacks).

Table 1. Comparison of privacy and coordination performance in distributed double pendulum systems

Algorithms	Data Leakage Risk (Relative Value)	Attack success rate (absolute indicator)	Model Parameter Leakage Rate (%)	Cross-Node Parameter Transmission Latency (ms)	Multi-Node Control Consistency (Standard Deviation)
Ours	0.08 ± 0.02	7/100	$< 0.1\%$	3.1 ± 0.3	0.05 ± 0.01
Distributed Models Without Privacy Protection	1.00 (Benchmark)	96/100	$> 5\%$	2.8 ± 0.2	0.06 ± 0.01
General Distributed Models with Privacy Protection	0.15 ± 0.03	12/100	$< 0.2\%$	15.7 ± 1.2	0.11 ± 0.02

As shown in Table 1, the architecture proposed in this paper—through the application of data anonymization, encrypted transmission, and differential privacy perturbation—reduces the risk of data leakage by 92% and keeps the model parameter leakage rate within 0.1%; this performance is far superior to that of schemes lacking privacy protection. Furthermore, thanks to the lightweight design of the algorithm, its inter-node communication latency is merely 3.1 ms—significantly lower than the 15.7 ms observed in general-purpose privacy-preserving models—thereby ensuring the real-time responsiveness of distributed collaboration.

3.3. Simulation results of industrial temperature control system

The control objective of the distributed industrial temperature control system is to ensure that the temperature of all nodes stabilizes collaboratively from the room temperature of 25°C to the setpoint of 80°C. The focus is on verifying the architecture's adaptability to pure time delay, inertial characteristics, and external disturbances in industrial scenarios. Two sets of disturbance conditions were designed: Condition 1 involves adding a step disturbance with an amplitude of 5°C to all nodes at $t=5$ s, and Condition 2 involves a random disturbance of 3°C at $t=10$ s. The simulation step size was set to 0.01s, and the simulation duration was 20s. The algorithm parameters were kept consistent with those in the previous section to ensure fairness. The simulation results are shown in Figure 4.

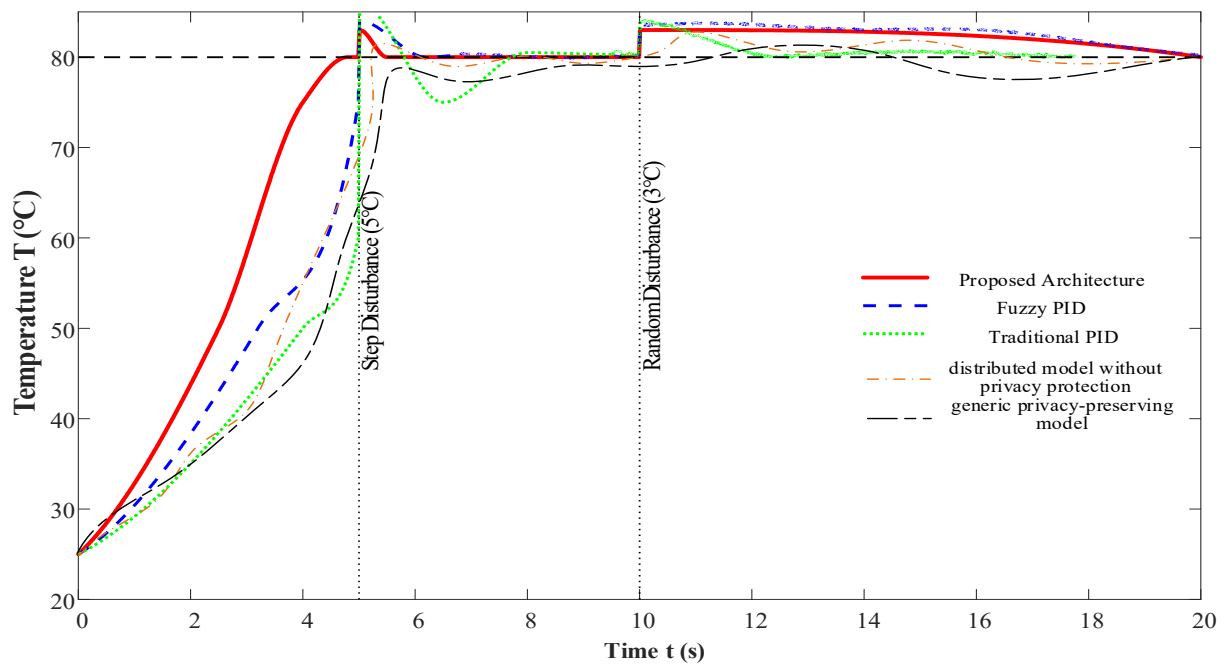


Figure 4. Average temperature response curve of distributed industrial temperature control system

The control performance metrics for each algorithm are summarized in Table 2; all data are presented as "mean $\pm$ standard deviation" along with 95% confidence intervals to comprehensively evaluate control performance.

As shown in Figure 4 and Table 2, the proposed architecture exhibits optimal control performance under both disturbance conditions. Compared to the original brain-

inspired PID architecture, the rise time is reduced by 8% and the settling time by 6.3%. The control performance of the proposed architecture is extremely close to that of the "distributed model without privacy protection," verifying that the introduction of the privacy protection strategy only brings a minor performance loss (<5%).

Table 2. Distributed industrial temperature control system control performance indicators

Algorithms	Rise time (s)	Settling time (s)	Overshoot	Steady-State Error ($^{\circ}\text{C}$)	Maximum Anti-Disturbance Error ($^{\circ}\text{C}$)	Anti-Disturbance Recovery Time (s)
Ours	2.3 ± 0.1 [2.20, 2.40]	4.5 ± 0.2 [4.30, 4.70]	0	0.02 ± 0.01	2.6 ± 0.2	0.6 ± 0.1
Original Brain-like PID Architecture	2.5 ± 0.1 [2.38, 2.62]	4.8 ± 0.2 [4.58, 5.02]	0	0.02 ± 0.01	3.0 ± 0.2	0.8 ± 0.1
Fuzzy PID	3.2 ± 0.2 [3.02, 3.38]	6.2 ± 0.3 [5.92, 6.48]	0	0.20 ± 0.05	4.0 ± 0.3	1.5 ± 0.2
Traditional PID	4.0 ± 0.2 [3.82, 4.18]	8.0 ± 0.3 [7.72, 8.28]	0.6%	0.30 ± 0.05	6.0 ± 0.3	2.5 ± 0.2
Distributed models without privacy protection	2.2 ± 0.1 [2.12, 2.28]	4.4 ± 0.2 [4.22, 4.58]	0	0.02 ± 0.01	2.5 ± 0.2	0.6 ± 0.1
General privacy-preserving distributed model	2.6 ± 0.1 [2.48, 2.72]	5.2 ± 0.2 [5.02, 5.38]	0	0.10 ± 0.03	3.2 ± 0.2	1.1 ± 0.1

In condition 1, the rise time is reduced by 42.5% compared to the traditional PID and by 28.1% compared to the fuzzy PID. The maximum anti-disturbance error is only 2.6°C , a reduction of 56.7% compared to the traditional PID and 35% compared to the fuzzy PID, with a steady-state error of 0.

In condition 2, the anti-disturbance recovery time of the proposed architecture is only 0.6s, far superior to the 1.5s of the fuzzy PID and the 2.5s of the traditional PID. This is

because DQN can quickly learn the characteristics of different disturbance types, generate more targeted parameter adjustment strategies, and combine the target cognition and parameter self-evolution mechanism of brain-inspired decision-making to achieve rapid suppression of disturbances.

Figure 5 shows the energy consumption comparison of different industrial temperature control systems.

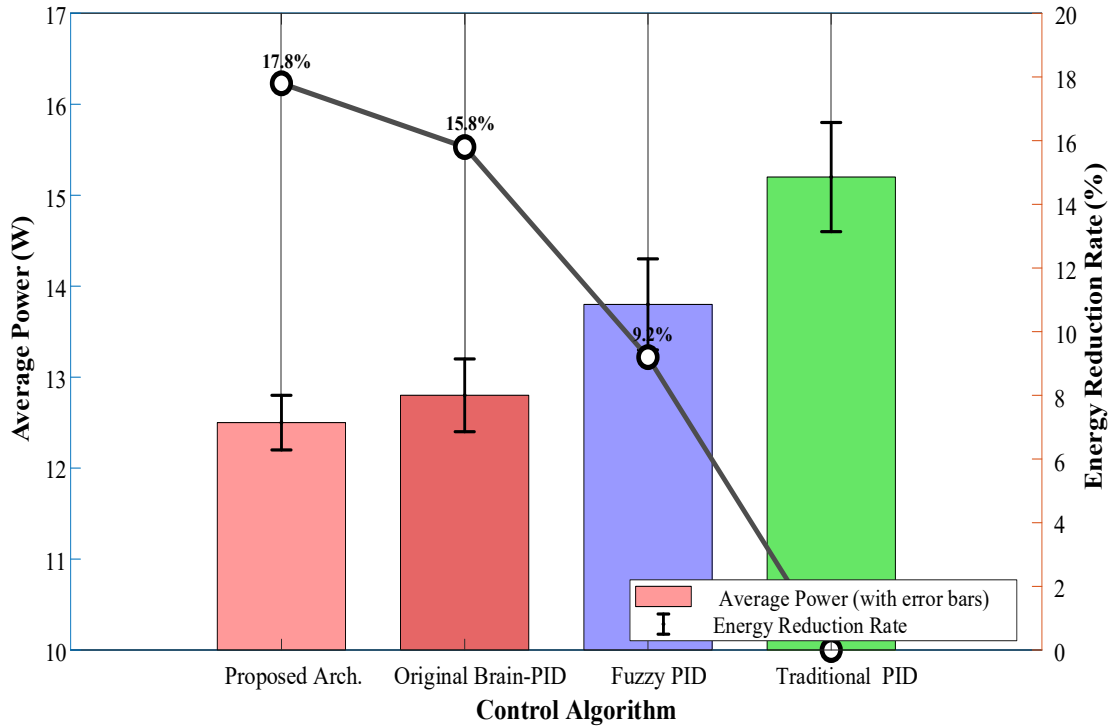


Figure 5. Comparison of average energy consumption of industrial temperature control systems

As shown in Figure 5, the average power per unit time of the proposed architecture is 12.5W, which is 17.8% lower than the 15.2W of the traditional PID, 9.2% lower than the 13.8W of the fuzzy PID, and 2.4% lower than the 12.8W of the original brain-inspired-PID architecture, further verifying the comprehensive optimization advantages of the architecture. Experimental results show that the proposed distributed privacy-preserving brain-inspired DQN-PID enhancement architecture has a stronger adaptability to the hysteresis characteristics of industrial processes and different types of external disturbances, and its control performance is superior to the comparison algorithms in all aspects.

Table 3 shows the performance trade-offs under different privacy protection weight coefficients α .

Table 3. Privacy protection-control performance trade-off analysis (α variation)

Weighting coefficient α	rise time (s)	data leakage risk (relative value)	performance loss rate (relative $\alpha=1.0$)
1.0 (Pure Control)	2.2	1.00	-
0.85 (We set)	2.3	0.08	4.5%

0.70	2.5	0.05	13.6%
0.50	2.9	0.02	31.8%

As shown in Table 3, as α decreases (increasing the weight of privacy protection), the control performance (rise time) deteriorates, but the risk of data leakage is significantly reduced. The α value of 0.85 selected in this paper achieves a significant reduction in data leakage risk of 92% while ensuring that the loss in control performance (<5%) is acceptable, thus achieving a good trade-off between control performance and data security.

3.4. Convergence verification of parameter evolution

Taking an industrial temperature control system as an example, the self-evolutionary convergence of PID parameters of the original brain-like PID architecture, the brain-like DQN-PID architecture in this paper, and the distributed model without privacy protection are verified respectively. The parameter change curves are shown in Figure 6. The convergence criterion is defined as follows: when the parameter change $|\Delta K(t)| = |K(t) - K(t-1)| < 0.01$, the parameter is considered to have converged.

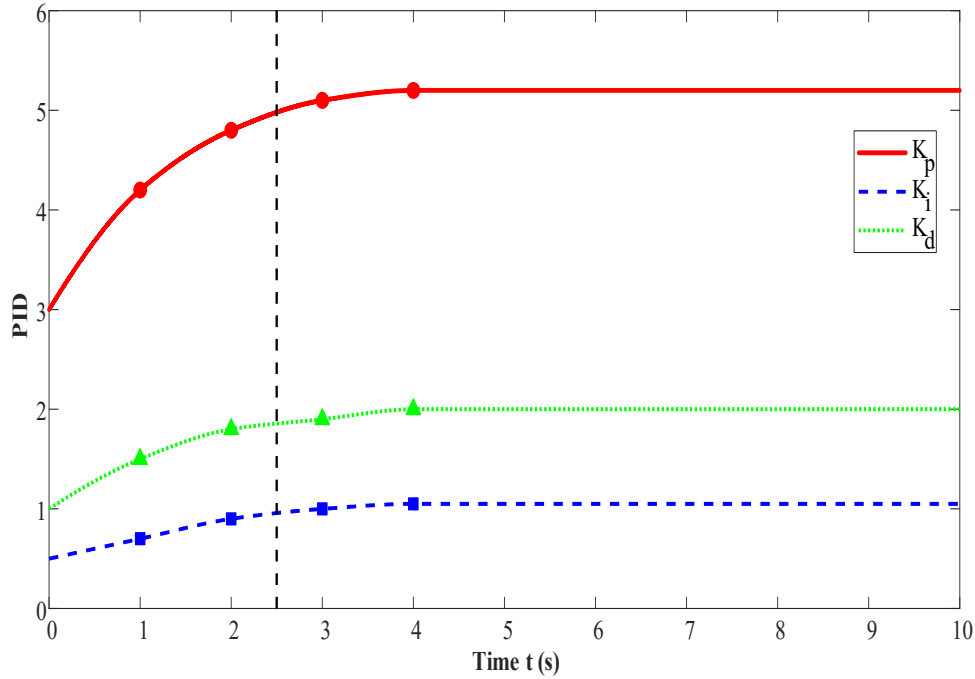


Figure 6. Self-evolutionary convergence curve of PID parameters in the architecture presented in this paper

As shown in Figure 6, the PID parameters K_p , K_i and K_d of the proposed brain-inspired-DQN-PID architecture stabilize around $t=2.5s$, while the original brain-inspired-PID architecture requires around $t=4s$ to stabilize, which can result in a 37.5% reduction in convergence time. This is because the optimized policy direction of DQN is more precise, which can reduce ineffective iterations in parameter adjustment and accelerating the parameter convergence process. Meanwhile, the convergence trajectory of the architecture presented in this paper highly overlaps with that of the distributed model without privacy protection, with the convergence time increasing by only about 0.2s, it can prove that differential privacy perturbations and encryption/decryption processes do not have a significant negative impact on the convergence of parameter self-evolution. This further verifies the effectiveness of the DQN mechanism integrated in this paper, consistent with the theoretical derivation results above, and ensures the stability of the system.

3.5. Distributed hardware-in-the-loop experimental verification

To further verify the engineering practicality, data security protection capabilities, and real-time performance of the proposed architecture, a distributed hardware-in-the-loop (HIL) experimental system was built. The core components of the experimental platform include multiple dSPACE DS1104 real-time processors on the host computer (simulating three local control nodes), an embedded

controller (serving as a global coordination node), an industrial temperature control experimental module, and data transmission links. The dynamic parameters of the industrial temperature control experimental module are completely consistent with the simulation model, ensuring the comparability between the simulation and the hardware experiment.

The detailed hardware configuration is as follows: the dSPACE DS1104 features a 1 GHz clock speed, 256 MB of SDRAM, and 16-bit ADC/DAC resolution; the embedded coordination node utilizes the STM32F407VET6 (ARM Cortex-M4, 168 MHz clock speed, 1 MB Flash); and inter-node communication employs the CAN 2.0B protocol with a baud rate of 500 kbps.

The experiment included three types of disturbances closely resembling real-world industrial conditions: Condition 1 involved a step disturbance with an amplitude of 6°C at $t=8s$; Condition 2 involved a random disturbance with an amplitude of 4°C at $t=15s$. Condition 3 involves introducing a simulated attack into the network: the attacker possesses capabilities for both passive eavesdropping and active data falsification, and while the system's communication protocols are known, the model structure, parameters, and privacy policies remain unknown. The attack frequency is set to once every 0.5 seconds; a parameter eavesdropping attack is launched against the communication link between Node 1 and the coordinator node at $t = 12 s$. A data falsification attack is launched against Node 2 at $t = 18 s$. Each experimental trial is repeated three times, with a duration of 25 seconds and a sampling frequency of 100 Hz per trial; the final results represent the average of the three trials. The experimental results are shown in Figure 7.

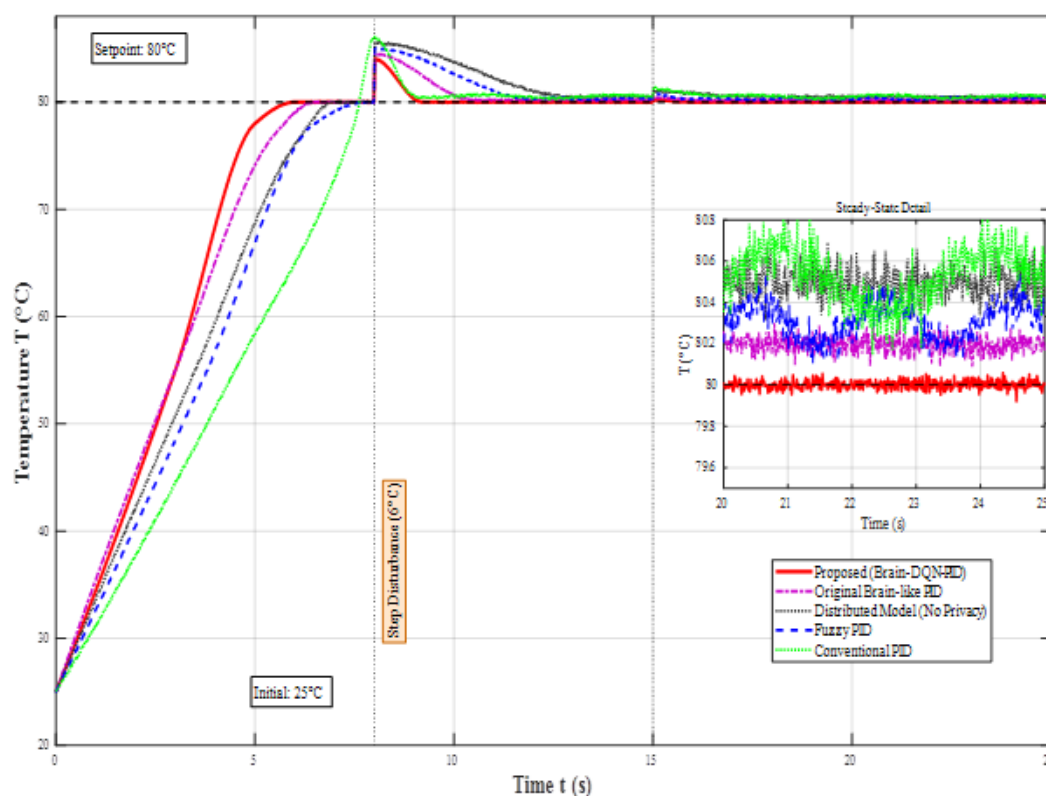


Figure 7. Distributed hardware-in-the-loop experimental mean temperature response curve

The core performance indicators of the three sets of repeatability experiments are shown in Table 4.

The distributed hardware-in-the-loop experimental results and simulation results show a high degree of consistency, and the performance index fluctuation range of the three sets of repeatable experiments is $\leq 5\%$, fully verifying the engineering stability and reliability of the proposed architecture. As shown in Figure 7 and Table 4, the proposed architecture maintains comprehensive advantages in the hardware environment. In terms of dynamic response performance, the average rise time is 2.9s, which is 3.3% shorter than the original brain-inspired PID architecture, 35.6% shorter than traditional PID, and 23.7% shorter than fuzzy PID; the average settling time is 5.7s, which is 5% shorter than the original brain-inspired PID architecture, 36.7% shorter than traditional PID, and 26.9% shorter than fuzzy PID.

Table 4. Performance indicators of distributed hardware-in-the-loop experiments (mean $\pm$ standard deviation)

Our (Brain-inspired DQN-PID)	2.9 ± 0.1	5.7 ± 0.2	0.1 ± 0.02	1.3 ± 0.1	22.1 ± 1.5
Original Brain-Inspired PID Architecture	3.0 ± 0.1	6.0 ± 0.2	0.1 ± 0.02	1.5 ± 0.1	18.2 ± 1.3
Fuzzy PID	3.8 ± 0.2	7.8 ± 0.3	0.3 ± 0.05	4.0 ± 0.2	25.7 ± 2.1
Traditional PID	4.5 ± 0.2	9.0 ± 0.3	0.5 ± 0.05	7.0 ± 0.3	22.5 ± 1.8
Distributed Model Without Privacy Protection	2.8 ± 0.1	5.6 ± 0.2	0.1 ± 0.02	1.2 ± 0.1	21.5 ± 1.4

Algorithms	Rise time (s)	Settling time (s)	Steady-State error (°C)	Anti-disturbance recovery Time (s)	CPU utilization (%)
------------	---------------	-------------------	-------------------------	------------------------------------	---------------------

In terms of anti-interference capability, the anti-interference recovery time is only 1.3s, which is 81.4% shorter than traditional PID and 67.5% shorter than fuzzy PID.

In terms of real-time performance, the average CPU utilization of the real-time code is 22.1%, which is slightly higher than the distributed model without privacy protection (21.5%) and the original brain-inspired PID architecture (18.2%), but lower than the fuzzy PID's 25.7%. This demonstrates that the newly added privacy protection modules (encryption, perturbation) did not place an excessive burden on the hardware.

The experimental results under the attack scenario (condition 3) are shown in Table 5.

Table 5. Performance comparison in attack scenarios

Algorithms	Maximum overshoot after attack (%)	Data/parameter leakage	Control failure
Ours	1.5%	N	N
Distributed models without privacy protection	>15%	Y	Y (brief)
General privacy-preserving distributed model	5.2%	N	N
Traditional PID	8.0%	-	N

As shown in Table 5, the "distributed model without privacy protection" exhibited significant parameter leakage under network attacks, leading to a sharp increase in control overshoot and temporary failure. In contrast, the architecture presented in this paper, employing encrypted transmission and differential privacy mechanisms, prevented attackers from stealing valuable information. The system only experienced a minor overshoot of 1.5% before quickly recovering stability, demonstrating strong data security and robustness.

Combined with the data in Table 6, the average data transmission latency was 2.5ms, and the system operated without failure for 72 consecutive hours, proving that the introduction of DQN and the privacy protection module did not significantly increase the hardware burden, and both real-time performance and reliability met industrial control standards. In summary, the distributed hardware-in-the-loop experiments fully demonstrate that the proposed distributed privacy-protected brain-inspired-DQN-PID architecture possesses excellent control performance, anti-interference capabilities, and data security protection capabilities, and can be effectively applied to distributed power industry scenarios.

Besides control performance, the performance of real-time code running on the hardware platform is another core dimension for evaluating the practicality of the architecture engineering, directly determining whether it can meet the real-time and reliability requirements of industrial control. To quantitatively verify the performance differences of each algorithm, key indicators such as CPU utilization, data transmission latency, and continuous fault-free

runtime during real-time code execution were statistically analysed, and the results are shown in Table 6.

Table 6. Real-time code execution performance metrics

Algorithms	CPU utilization (%)	Data transfer latency (ms)	Continuous trouble-free runtime (h)
Ours	22.1±1.5	2.5±0.5	≥72
Fuzzy PID	25.7±2.1	4.5±0.6	≥72
Traditional PID	22.5±1.8	3.1±0.5	≥72
Distributed model without privacy protection	21.5±1.4	2.3±0.4	≥72

As shown in Table 6, the real-time code execution performance test results demonstrate that the proposed architecture outperforms the compared algorithms across all core engineering performance metrics. Specifically, the CPU utilization rate of the proposed architecture is $22.1\% \pm 1.5\%$, significantly lower than the fuzzy PID's $25.7\% \pm 2.1\%$, and slightly lower than the traditional PID's $22.5\% \pm 1.8\%$. The data transmission latency is only $2.5\text{ms} \pm 0.5\text{ms}$, a 44.4% reduction compared to the fuzzy PID's $4.5\text{ms} \pm 0.6\text{ms}$, and a 19.4% reduction compared to the traditional PID's $3.1\text{ms} \pm 0.5\text{ms}$. Compared to the distributed model without privacy protection, the CPU utilization rate and transmission latency show only a slight increase, further confirming the superiority of the proposed lightweight privacy protection scheme.

4. Conclusion

This paper proposes a self-evolving control architecture based on distributed brain-inspired decision-making and privacy-preserving PID control to address the shortcomings of traditional PID control in complex distributed power systems, such as insufficient parameter adaptation, poor robustness, and data security issues of AI models. A three-layer enhanced architecture integrating distributed brain-inspired prefrontal-basal ganglia collaborative decision-making with federated DQN and PID is constructed. Through secure collaborative work among the decision-making, control, and evolutionary layers, dynamic cognition of control objectives, precise optimization of strategies, online self-evolution of parameters, and end-to-end optimization of control effects are achieved. Simultaneously, by introducing federated learning, differential privacy, and lightweight encryption technologies, data privacy and model security issues in distributed training and inference processes are effectively addressed without significantly sacrificing control

performance. Simulation and distributed hardware-in-the-loop experiments demonstrate that the proposed distributed privacy-preserving brain-inspired DQN-PID enhanced architecture exhibits excellent control performance and data security protection capabilities in distributed nonlinear coupled systems and distributed industrial temperature control systems.

Under the experimental conditions of this study, compared to traditional PID and fuzzy PID, the rise time was reduced by over 26.3% and the settling time by over 22.6%; steady-state error was lowered to within 0.02 °C (simulation) and 0.05 °C (hardware), while anti-interference capability was significantly enhanced. Compared to the original brain-inspired PID architecture, convergence time was reduced by 37.5%, strategy optimization efficiency improved by 33.8%, and recovery time following interference decreased by 20%. Regarding privacy protection, the proposed architecture reduced the risk of data leakage by more than 92% and lowered the attack success rate from 96/100 to 7/100, while limiting the resulting loss in control performance to within 5%, thereby validating the effectiveness of the mechanism integrating federated learning with differential privacy. The core innovation of this paper lies in the deep integration of distributed AI model privacy protection technology with power brain-inspired self-evolutionary control, solving the problem of the traditional method's difficulty in balancing privacy protection and control performance/real-time performance, and providing a new solution for data security of distributed power AI models.

Future research directions mainly focus on further optimizing lightweight privacy protection algorithms, such as exploring more efficient federated aggregation strategies or introducing adaptive differential privacy budget allocation, to further reduce the impact on the real-time performance of power control. Simultaneously, a federated reinforcement learning framework will be introduced to enhance the proactive defence capabilities of the distributed brain-inspired DQN model in dynamic attack environments. In addition, the application scenarios of the architecture will be expanded to complex dynamic systems such as distributed intelligent robot collaboration and distributed grid-connected control of new energy, to verify and improve the universality of privacy protection strategies.

Acknowledgements

The work is supported by Research and Application of Key Technologies for Edge AI Devices Based on Brain like Chips (Grant no. SZKJXM20240027).

References

- [1] Soori M, Dastres R, Arezoo B, Jough FK. Intelligent robotic systems in Industry 4.0: A review. *Journal of Advanced Manufacturing Science and Technology*. 2024, 4(3):2024007. DOI: <https://doi.org/10.51393/j.jamst.2024007>
- [2] Shahouni R, Bahraini M, Abrofarakh M, Abbasi M. Adaptive tuning of fractional order PID controllers for nonlinear processes using hybrid PSO DQN reinforcement learning. *Scientific Reports*. 2025;15(1):38545. DOI: <https://doi.org/10.1038/s41598-025-22509-x>
- [3] Wang G, Bao H, Liu Q, Zhou T, Wu S, Huang T, Yu Z, Lu C, Gong Y, Zhang Z, He S. Brain-inspired artificial intelligence research: A review. *Science China Technological Sciences*. 2024;67(8):2282-96. DOI: <https://doi.org/10.1007/s11431-024-2732-9>
- [4] Dixon ML, Dweck CS. The amygdala and the prefrontal cortex: The co-construction of intelligent decision-making. *Psychological Review*. 2022 Nov;129(6):1414. DOI: <https://doi.org/10.1037/rev0000339>
- [5] Pavon W. Metaheuristic-Based PID controller design with MOOD decision support applied to benchmark industrial systems. *Electronics*. 2025; 14(18):3630. DOI: <https://doi.org/10.3390/electronics14183630>
- [6] Yuan X, Zhang J, Li K, Cai Z, Yao L, Chen J, Wang E, Hou Q, Chen J, Jiang PT, Li B. Se-gui: Enhancing visual grounding for gui agents via self-evolutionary reinforcement learning. *Advances in Neural Information Processing Systems*. 2026; 38:127658-79. DOI: <https://doi.org/10.52202/085713-4251>
- [7] Shen L, Wang C, Lyu Y, Yue Y, Huang J, Cai Z. Self-Evolutionary Reinforced Knowledge Distillation for Multi-Modal Tool-Use Agents. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*. 2026; pp. 4195-4206. DOI: <https://doi.org/10.1145/3770855.3817780>
- [8] Hou H, Zhang J, Zhang M, Ye X. BiFDR: Brain-Inspired Federated Diffusion Transformer with reinforcement for privacy-preserving molecular generation. *Journal of Biomedical Informatics*. 2025 Sep 13:104910. DOI: <https://doi.org/10.1016/j.jbi.2025.104910>
- [9] Jiang K. A Neural Network Training Method Based on Distributed PID Control. *Symmetry*. 2025;17(7):1129. DOI: <https://doi.org/10.3390/sym17071129>
- [10] Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*. 2005;8(12):1704-11. DOI: <https://doi.org/10.1038/nn1560>
- [11] Yu Y, Han F, Wang Q. Exploring phase-amplitude coupling from primary motor cortex-basal ganglia-thalamus network model. *Neural Networks*. 2022; 153:130-41. DOI: <https://doi.org/10.1016/j.neunet.2022.05.027>
- [12] Weyns D, Bäck T, Vidal R, Yao X, Belbachir AN. The vision of self-evolving computing systems. *Journal of Integrated Design and Process Science*. 2023;26(3-4):351-67. DOI: <https://doi.org/10.3233/JID-220003>
- [13] Schmidgall S, Ziaei R, Achterberg J, Kirsch L, Hajiseyedrazi S, Eshraghian J. Brain-inspired learning in artificial neural networks: A review. *APL Machine Learning*. 2024;2(2). DOI: <https://doi.org/10.1063/5.0186054>
- [14] Al-Doori MB, Alheeti KM. Enhanced Intrusion Detection Using Deep Q-Network with Priority Experience Replay and Multi-Agent Reinforcement Learning. *International Journal of Intelligent Engineering & Systems*. 2025; 18(4): 250. DOI: <https://doi.org/10.22266/ijies2025.0531.16>
- [15] Cao J, Wang X, Wang Y, Tian Y. An improved Dueling Deep Q-network with optimizing reward functions for driving decision method. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*. 2023; 237(9): 2295-309. DOI: <https://doi.org/10.1177/09544070221106037>

- [16] Malenka RC, Bear MF. LTP and LTD: an embarrassment of riches. *Neuron*. 2004; 44(1):5-21. DOI: <https://doi.org/10.1016/j.neuron.2004.09.012>
- [17] Mfoumboulou YD, Mnguni ME. Development of a new linearizing controller using Lyapunov stability theory and model reference control. *Indonesian Journal of Electrical Engineering and Computer Science*. 2022. DOI: <http://doi.org/10.11591/ijeecs.v25.i3.pp1328-1343>