A Physiological Control System for Pulsatile Ventricular Assist Device Using an Energy-Efficient Deep Reinforcement Learning Method

Te Li, *Member, IEEE*, Wenbo Cui, Xingjian Liu, *Member, IEEE*, Xu Li, Nan Xie, and Yongqing Wang

***Abstract*— Pulsatile ventricular assist device (PVAD) is a blood pump used to assist the circulation support of the native heart. Because of patients' complex physiological environment, PVAD's control system requires high adaptive ability and low computational energy. However, traditional PID controllers do not possess sufficient adaptive ability. Although neural network controllers are with high adaptive ability, their extensive energy cost limits the applications. In this study, a PVAD physiological control system based on deep reinforcement learning (DRL) is proposed, which significantly improves the system's adaptive ability. To further reduce its energy cost, a new energy-efficient DRL method, AddTD3, is developed, in which the fully connected network (FC) with high computation complexity is replaced by energy-efficient AdderNet. Experimental results show that the proposed AddTD3 controller is with higher adaptive ability than the traditional PID controller (cumulative absolute error: 237.1 versus 484.3 mmHg) and can be migrated to the mock circulation system (MCS) without fine-tuning. It can reduce the energy cost of the traditional DRL algorithm TD3 to 44.8% without reducing the performance (8773.2 versus 22420.4 pJ).**

***Index Terms*— Adder neural network, deep reinforcement learning (DRL), heart failure, physiological control, pulsatile ventricular assist device (PVAD).**

I. Introduction

VENTRICULAR assist device (VAD) has gradually become an important means to treat patients with heart failure who cannot pump blood properly due to the structural or functional damage of their hearts [1]. Pulsatile VAD (PVAD) is a typical VAD approach, which works similar to native hearts [2]. It intermittently drives the blood flowing through itself with pressure and assists the patient's failing heart with blood perfusion. A study of implantable pneumatic PVAD shows that patients exhibited an 81% improvement in two-year survival with this device as opposed to optimal medical therapy [3].

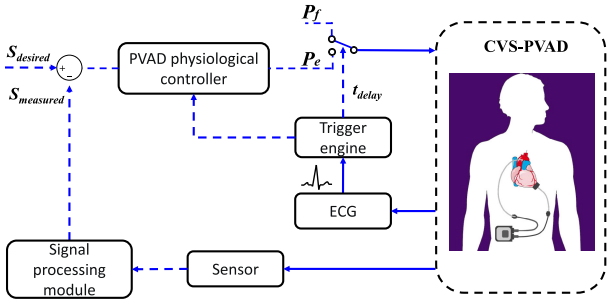


Fig. 1. Block diagram of the synchronous PVAD physiological control system. S_{desired} and S_{measured} are the desired and measured physiological signals, respectively. The trigger engine uses ECG signals to control PVAD ejection and contraction at a frequency synchronized with the heart rate. The physiological signals can be used including but not limited to aortic pressure and aortic flow. Images of the representative heart were taken from "Smart Servier Medical Art" (https://smart.servier.com/).

However, there are some challenges with the PVAD physiological control system, to allow the people discharged from the hospital to live a normal life wearing the PVAD. The framework of the PVAD physiological control system is shown in Fig. 1. Physiological signals (e.g., aortic pressure, aortic flow) are collected from the cardiovascular system (CVS) by sensors. The difference between the processed signal and the expected value is set as the input to the PVAD controller at a rate synchronized with the heart rate to adjust the driving pressure.

One of the challenges is that it requires high adaptive ability and should be able to adaptively adjust the driving pressure according to the patient's blood perfusion needs. Several techniques have been developed to improve the adaptive ability of the PVAD control system [4], [5], [6], [7]. The traditional PVAD physiological controller is based on the PID method [8], [9]. However, the traditional PID-based methods can only work normally under fixed physiological environmental parameters. Furthermore, in dynamic and complex CVS-PVAD, the PID-based methods' adaptive abilities

Manuscript received 13 February 2023; revised 24 April 2023; accepted 2 May 2023. Date of publication 19 May 2023; date of current version 7 June 2023. This work was supported in part by the National Key R&D Program of China under Grant 2022ZD0116700, in part by the National Natural Science Foundation of China under Grant 52105008, in part by the Fundamental Research Funds for the Central Universities under Grant DUT22GF301 and Grant DUT21RC(3)037, in part by the Foundation of State Key Laboratory of Digital Manufacturing Equipment and Technology under Grant DMETKF2022009, and in part by the Young Elite Scientists Sponsorship Program by CAST under Grant YESS20220068. The Associate Editor coordinating the review process was Dr. Gabriele Patrizi. *(Corresponding author: Xingjian Liu.)*

Te Li, Wenbo Cui, Xu Li, Nan Xie, and Yongqing Wang are with the State Key Laboratory of High-performance Precision Manufacturing and the School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China (e-mail: teli@dlut.edu.cn; 1297691410@mail.dlut.edu.cn; imlixu@dlut.edu.cn; xiedanan@mail.dlut.edu.cn; yqwang@dlut.edu.cn).

Xingjian Liu is with the State Key Laboratory of High-Performance Precision Manufacturing and the School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China, and also with the State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: xjliu@dlut.edu.cn).

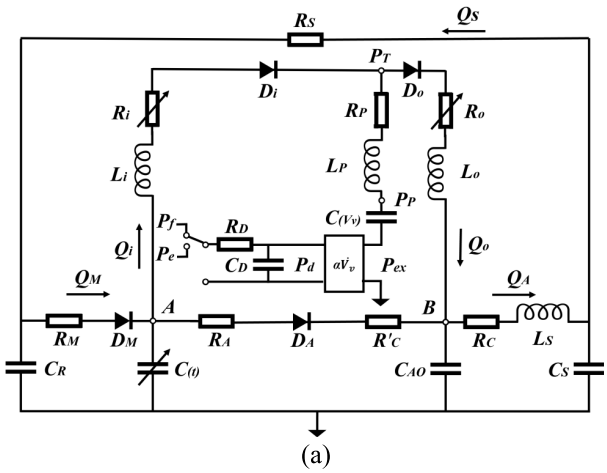
Digital Object Identifier 10.1109/TIM.2023.3277993

cannot satisfy the demand [10], which is prone to oscillation, long adjustment time, etc. The fuzzy controller [11] is also a common VAD controller, and its adaptive ability is slightly stronger than the PID controller. However, the performance of fuzzy controllers in VAD control tasks has not been proved by large-scale experiments [12], [13]. The model-free adaptive control (MFAC) method is a viable way to address this problem, which does not require knowledge about the dynamic and structural information of the controlled system [4], [5]. This controller shows excellent adaptive ability in a large number of simulation experiments with interpatient and intrapatient variations. However, the experimental results show that this method is very time-consuming in the control process. In the switching process of a patient from rest to exercise, the controller costs over 90 s to complete the control task. According to recent studies, network-based learning methods are also proven to be effective tools to improve the control adaptive ability of PVAD, such as the multilayer perceptron (MLP) method [7], recurrent neural network [6], and convolutional neural network (CNN) method [5], [14]. The experimental results show that these methods can effectively improve the robustness of the controller. However, these works [4], [5], [6], [7], [15] are all conducted in the simulation environment, and the control effectiveness of these methods in the real-world mock circulation system (MCS) is unknown at present. In addition, the problem of the high computational energy cost of neural-network-based PVAD physiological controllers has been ignored in previous studies [4].

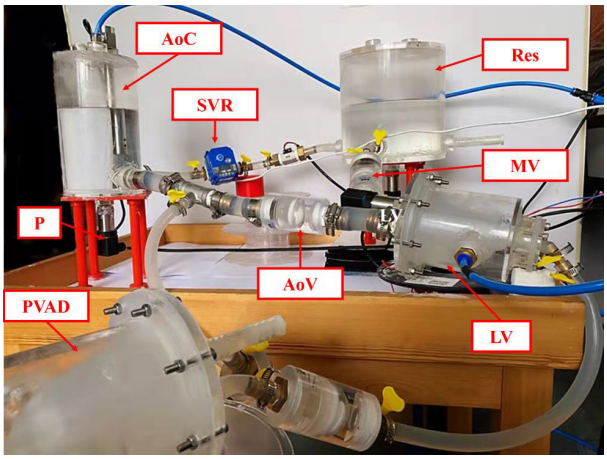
Another challenge with the PVAD physiological control system is reducing its computational energy cost. To improve the longevity of a wearable medical device, it needs to operate at a low power consumption [16], [17]. Although neural-network-based methods show high adaptive ability in PVAD control tasks [5], [6], [7], [14], they require a large number of multiplication operations and consume a lot of computing energy. A series of works have been proposed in the field of model compression, such as network pruning [18], [19], and knowledge distillation [20], to reduce computation by shrinking or shallowing the network [21]. However, these methods still involve a lot of multiplication operations, which cannot tackle the root of the problem to reduce the device's energy cost. In addition, quantization methods [22] have been proposed to reduce computational complexity while maintaining the original neural network structure. Although it can reduce energy costs, the accuracy of the network cannot be guaranteed. Hence, a method with high network accuracy and low computational energy cost is needed for the PVAD physiological control system.

The contributions of this article are summarized as follows.

1) To solve the first challenge, this article proposes a PVAD physiological control system based on deep reinforcement learning (DRL), which significantly improves the adaptive ability of the system. The experimental results show that the proposed method has higher adaptive ability than the traditional PID controller (cumulative absolute error: 148.8 versus 423.5 mmHg).



(a)



(b)

Fig. 2. (a) Electrical equivalent circuit simulation model of the CVS with PVAD. *P*, pressures; *R*, resistances; *C*, compliances; *L*, inertances; and *D*, diodes. (b) MCS. AoC, aortic compliance chamber; SVR, blood resistance simulator; Res, fluid storage chamber; LV, left ventricle; PVAD, pulsatile VAD; AoV, aortic valve; MV, mitral valve; and P, pressure sensor.

2) To solve the second problem and further reduce the computational energy consumption of the algorithm, this article proposes a new energy-efficient DRL algorithm AddTD3, in which the fully connected network (FC) with high computation complexity is replaced by energy-efficient AdderNet. The experimental results show that the algorithm can effectively reduce the computational energy consumption (34835.2 versus 77721.6 pJ) without reducing the performance of the algorithm.

The rest of this article is organized as follows. Section II introduces the general situation of the experimental system. Section III introduces the proposed PVAD physiological controller based on DRL, the energy-efficient DRL algorithm AddTD3, and the agent training method based on imitation learning and automatic domain randomization (ADR). Section IV presents the experimental results and analysis. Finally, Section V is summarized.

II. System Overview

A. Numerical Model of CVS-PVAD

In this work, the CVS-PVAD numerical model is referred to [23] and implemented in MATLAB [see Fig. 2(a)]. This

simulation system greatly improves the training speed of the DRL and avoids damage to the physical system. The blood flow starts from point A, flowing to PVAD (Q_i) and aortic valve (represented by resistance R_A and diode D_A), respectively. After converging at point B, blood flows to the aorta (Q_A), and after circulating through the body, it flows to the mitral valve (represented by resistance R_M and diode D_M) and finally returns to point A. This system is used in this study to train DRL agents and to verify the performance of different controllers.

B. MCS Experimental Setup

The MCS is constructed, as shown in Fig. 2(b), which can verify the control effect of the proposed PVAD control system in the real-world system. The designed MCS consists of a left ventricular simulator, a left VAD, an aortic compliance chamber, a blood resistance simulator, a fluid storage chamber, and a valve equipped with a pressure sensor in the aorta. The left ventricular simulator and PVAD production process reference from [24].

C. Operation Sequence

Fig. 3 illustrates the overall framework and process of the study. In the training mode, the agent first adopts the PID controller as the expert strategy under the fixed environmental parameters and carries on the imitation learning [25] in the simulation environment. Then the ADR method [26] is used to train the agent. Finally, the control system is transferred to MCS for verification experiments. In the control loop, the network used for control in AddTD3, AddPolicyNet, outputs the control signal P_e and inputs P_e into the simulation environment with a delay of t_{delay} after detecting the R-wave. The simulation environment then processes the detected signal into a group of memories and stores them in the replay buffer for AddTD3 training. In the working mode, the trained AddPolicyNet is directly transferred to the MCS. The sensor in MCS transmits the collected aortic pressure signal to AddPolicyNet for processing. AddPolicyNet outputs the control signal P_e. After detecting the R-wave, the control signal is transmitted to programmable logic controller (PLC) by the upper computer for control with a delay of t_{delay}.

III. Methodology

To meet the needs of PVAD physiological control systems with high adaptive ability and low computational energy cost, two key technologies are developed in this work. A PVAD physiological control system based on the DRL algorithm and a new DRL algorithm AddTD3 with lower computational energy cost are proposed. In addition to these two key technologies, imitation learning and ADR are adopted to accelerate the training speed and transferability of agents.

A. DRL-Based PVAD Physiological Control System With High Adaptive Ability

Fig. 1 shows a classic PVAD physiological control system, which is based on synchronous control. The use of synchronizing PVAD with the native heart might bring substantial treatment benefits to patients with severe heart failure [15] and also adapt the PVAD ejection frequency to the patient's heart rate. In previous studies, the PID method was used as the PVAD physiological controller to adjust PVAD ejection pressure, but the adaptive ability of the PID method was insufficient. We propose a DRL-based PVAD control method, that is, the PVAD physiological controller in Fig. 1 is defined as the DRL controller. It can significantly improve the adaptive ability of the system.

As shown in Fig. 1, the PVAD synchronous control system uses the ECG signal to detect the R-wave. After each detection of the R-wave, the trigger delays for a while and then controls the PVAD to ejection. Where the R-wave is a signal extracted from the ECG that can be used for heartbeat detection [27]. The DRL-based PVAD control system uses an improved differential threshold method proposed in [28] to detect the R-wave. As shown in Fig. 4, the R-wave in ECG is used to detect the heartbeat. k denotes the kth heartbeat cycle, and the R-wave signal $\delta_R[k]$ and PVAD ejection signal $\delta_{\text{PVAD}}[k]$ were defined as follows:

$$\delta_R[k] = \begin{cases} 1, & \text{R-wave detected} \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

$$\delta_{\text{PVAD}}[k] = \begin{cases} 1, & t = t_R[k] + \lambda T_c[k-1] \\ 0, & \text{Otherwise.} \end{cases} \tag{2}$$

As shown in Fig. 4 and (1), $t_R[k]$ represents the moment when the kth R-wave signal is detected, i.e., the start of the kth heartbeat cycle. $\lambda \in [0, 1)$ is the ejection delay ratio factor, and $T_c[k]$ is the duration of a heartbeat cycle.

The aortic pressure signal P_A is used as the physiological signal to input into the PVAD physiological controller, and the controller outputs the ejection pressure P_e. The PVAD pneumatic driving pressure P_d has two options: injection pressure P_e and filling pressure P_f. Filling pressure $P_f = -35$ mmHg is a fixed negative pressure that helps PVAD return to its initial state. When $\delta_{\text{PVAD}}[k]$ is 1, the pressure signal P_d is switched to the injection pressure P_e, and after maintaining $\sigma T_c[k-1]$ s, P_d is switched back to P_f. Where $\sigma \in [0, 1)$ is the coefficient of PVAD ejection duration.

The traditional PVAD physiological control system is based on the PID method. Its output is calculated according to

$$u(k) = K_a e(k) + K_b e(k-1) + K_c e(k-2) \tag{3}$$

where $u(k)$ is the output of the PID controller, namely, PVAD drive pressure P_d. $K_a = K_P + K_I + K_D$, $K_b = K_P + 2K_D$, and $K_c = K_D$ are calculated by PID parameters K_P, K_I, and K_D, respectively, and $e(k)$ is the error value of the input PID controller at time k. The PID-based physiological control systems for PVAD rely only on the error e_k and fixed parameters K_P, K_I, and K_D, without understanding the potential interactions between the PVAD and CVS. The constrained, nonlinear, observation error, and time-varying nature of the CVS-PVAD model parameters may cause the controller to become unreliable in the event of sudden changes in cardiovascular conditions [10].

Unlike the PID method, DRL is a learning-based algorithm. DRL has been widely used in tracking control problems [29],

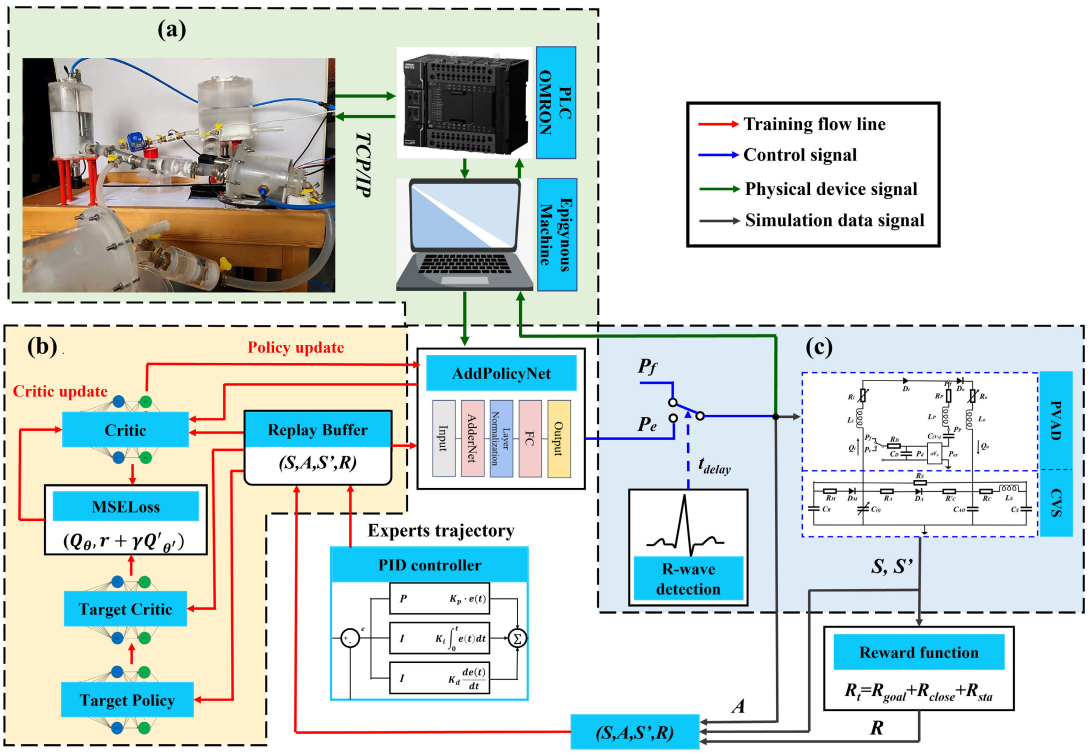


Fig. 3. Overall framework of the proposed method. The whole framework is divided into three modules and two modes. The three modules include (a) MCS module, (b) AddTD3 algorithm module, and (c) PVAD control system module. AddPolicyNet belongs to (b) and (c). The two modes include the training mode and working mode. In the training mode, the data flow through (b) and (c) via the red, blue, and gray lines. In the working mode, the data flow through (c) and (a) via the blue and green lines.

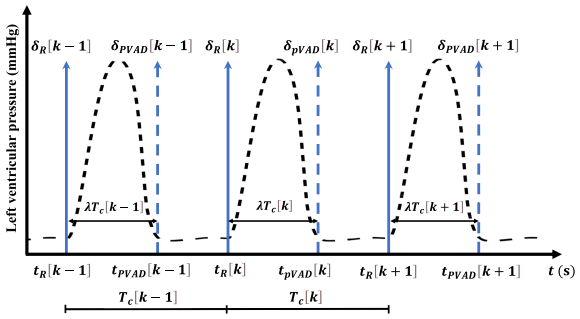


Fig. 4. Diagram of PVAD ejection delay time. $\delta_R[k]$, kth heartbeat cycle native heart ventricular contraction detection, i.e., kth heartbeat cycle R-wave; $\delta_{\text{PVAD}}[k]$, PVAD ejection time in the kth heartbeat cycle; t_{delay}, PVAD ejection delay time; $T_c[k-1]$, the length of the kth heartbeat cycle. During each heartbeat cycle, the PVAD injection moment is always synchronized but delayed with respect to the native heart R-wave.

[30]. TD3 [31] is one of the most advanced continuous control DRL algorithms, which solves the problem of overestimation of the DDPG [32] value function. TD3 consists of six networks, which are PolicyNet, target PolicyNet, two CriticNet, and two target CriticNet. During training, target CriticNet calculates action values $Q_{\phi'}$ at moment t based on the state s_t which is calculated by $P_A[k]$ and the target PolicyNet's output a_t. The meaning of Q_ϕ estimated from target CriticNet, is the cumulative discounted reward expectation of the PVAD physiological control system. In other words, $Q_{\phi'}$ is the target CriticNet's evaluation of the output of the current PVAD physiological control system. The loss function of Q_ϕ is defined as

$$y = r + \gamma \min_{i=1,2} Q_{\phi_i'}(s_{t+1}, \pi_{\theta'}(s_{t+1})) + \text{clip}(\epsilon, -c, c)) \quad (4)$$

$$L_{\phi_i} = E\left[\left(Q_{\phi_i}(s_t, a_t) - y\right)^2\right], \ i = 1, 2 \quad (5)$$

where ϕ_i is the parameter of the ith CriticNet, and ϕ_i' is the parameter of the ith target CriticNet. θ is a parameter of the PolicyNet, θ' is a parameter of the target PolicyNet, and $\pi_{\theta'}(s_t)$ is the action that the PolicyNet network outputs based on input s_t. $\epsilon \sim N(0, \sigma)$ is the noise used to smooth the target policy, and $\gamma = 0.9$ is the discount factor. The action a_t output by the PolicyNet will be input to the CriticNet, and the CriticNet will give Q_{ϕ_i} of a_t. The gradient of PolicyNet's parameters is updated according to this Q_{ϕ_i} using the derivative chain rule. The gradient of PolicyNet is expressed as

$$\begin{aligned} \nabla_\theta J_\theta &= E\left[\nabla_\theta Q_{\phi_i}(s_t, a_t)|_{a_t=\pi_\theta(s_t|\theta)}\right] \\ &= E\left[\nabla_a Q_{\phi_i}(s_t, a_t)|_{a_t=\pi_\theta(s_t)} \nabla_\theta \pi_\theta(s_t)\right] \end{aligned} \quad (6)$$

where $E[\cdot]$ represents the mathematical expectation, $\nabla_\theta Q_{\phi_i}(s_t, a_t)$ represents the gradient of CriticNet on parameters of PolicyNet, $\nabla_a Q_{\phi_i}(s_t, a_t)$ represents the gradient of CriticNet on the action a_t, and $\nabla_\theta \pi_\theta(s_t)$ represents the gradient of the policy $\pi_\theta(s_t)$ on network parameter θ. In continuous iteration and learning, PolicyNet's output strategy π_θ is optimized. In this process, the action value Q_ϕ output by CriticNet guides PolicyNet to gradually understand how the ejection pressure P_e of PVAD should change when the PVAD physiological system is in different states. This will significantly improve the adaptive ability of the system.

The proposed physiological control system based on DRL is shown in Fig. 3. AddTD3 constantly interacts with the PVAD control system and determines the optimal control strategy of the PVAD driver according to the reward. Convert the PVAD control problem into a DRL problem and construct the following three crucial elements.

1) State Space: Data collection for wearable medical devices is limited because the sensors in CVS are mostly invasive [5], and too many sensors will cause additional damage to the human body. The mature method of measuring the aortic pressure as the input signal source was selected. The original sensor data are P_A, the aortic pressure signal within a heartbeat cycle. The state s_t is represented as

$$s_t = \left\{E[k], \bar{P}_A[k], F_{\text{sign}}\right\} \tag{7}$$

$$E[k] = \left|P_{\text{target}} - \bar{P}_A[k]\right| \tag{8}$$

$$\bar{P}_A[k] = \frac{1}{T_c[k-1]} \int_{t_{\text{PVAD}}[k-1]}^{t_{\text{PVAD}}[k]} P_A(t)dt \tag{9}$$

where F_{sign} is the sign used to determine whether the system error is small enough, when $E[k] < 1$, $F = 1$; otherwise, $F = 0$.

2) Action Function: To reduce the left ventricular load and ensure adequate aortic blood flow to reduce aortic valve dysfunction, ejection delay ratio factor $\lambda = 0.4$ and ejection duration coefficient $\sigma = 0.3$ were determined to remain constant. The controller only needs to adjust the PVAD ejection pressure P_e.

3) Reward Space: In the PVAD controller, divide the reward into three modules, $R_t = R_{\text{goal}} + R_{\text{close}} + R_{\text{sta}}$. R_{goal} is a step reward function. When the CVS is in a healthy state, that is, $\bar{P}_A[k]$ is approximately equal to the expected value, R_{goal} is set to 1; if $\bar{P}_A[k]$ is too far away from the expected value, $R_{\text{goal}} = -5$ is given and the current simulation ends; otherwise, the reward is 0. $R_{\text{close}} = \alpha S[k]$ is a piecewise linear reward function that constraints $\bar{P}_A[k]$ as close to the expected pressure P_{target} as possible, where α is a reward normalization factor. R_{sta} is a penalty reward for suppressing PVAD controller oscillation

$$R_t = \begin{cases} 1 - \alpha E[k] - \text{action}^2, & \text{if } 0 \le E[k] \le 1 \\ -\alpha E[k] - \text{action}^2, & \text{if } 1 \le E[k] \le 20 \\ -5 - \alpha E[k], & \text{if } 20 < E[k]. \end{cases} \tag{10}$$

B. AddTD3: An AdderNet-Based DRL Algorithm With High Energy Efficiency

Although the application of the DRL algorithm in the PVAD control system enhances the adaptive ability of the system, it pays the price for extra computational energy. As a tool of policy approximation, neural networks have a high computational burden and memory requirement in the inference stage. This is a problem to be solved for the PVAD that is not rich in computing resources and inconvenient to carry extra batteries. A simple solution is to replace the policy network in DRL directly with AdderNet [33]. This algorithm is proposed in the field of model compression and has achieved good results in many tasks of deep learning [21], [34], [35]. But the new algorithm with simple substitution is difficult to converge. To solve this problem, we modified the structure of AdderNet and propose a DRL algorithm AddTD3 based on AdderNet. Compared with TD3, the proposed AddTD3 should achieve the following two points: 1) increase the algorithm energy efficiency and 2) maintain the performance of the original algorithm.

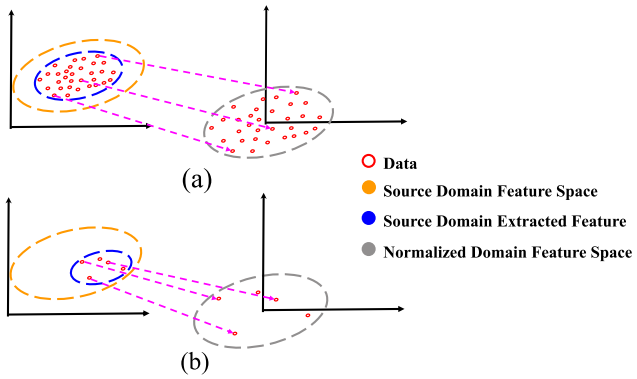


Fig. 5. Influence of the number of output elements in the output layer on normalization. (a) During the migration of features from the source domain to the normalized domain, if the number of samples is large enough, the estimated source domain features from the samples will be very close to the actual source domain feature space, and the original relationship of features will remain after migration to the normalized domain. (b) If the sample data are insufficient, the gap between the source domain feature space estimated only from the sampled data and the actual source domain feature space is too large, which will destroy the feature distribution in the normalization process and lead to the failure of normalization.

In TD3, the convolution kernel is expressed as $F \in \mathbb{R}^{d\times d\times C_{\text{in}}\times C_{\text{out}}}$, where kernel size is d, and C_{in} and C_{out} are the input and output channels, respectively, The input feature is defined as $X \in \mathbb{R}^{H\times W\times C_{\text{in}}}$, where H and W are the length and width of the feature, respectively. The PolicyNet's output at each layer is

$$Y(m,n,t) = f\left(\sum_{i=0}^{d}\sum_{j=0}^{d}\sum_{k=0}^{C_{\text{in}}}(X(m+i,n+j,k)\times F(i,j,k,t))\right) \tag{11}$$

where $f(\cdot)$ is the activation function of the current layer. The difference between the hidden layer and the output layer of PolicyNet is only in the activation function, with ReLU for the hidden layer and tanh for the output layer.

Using AdderNet to replace FC in TD3 can effectively reduce the calculation energy cost. In AdderNet, use the ℓ_1 norm instead of cross correlation operations, and AdderNet output characteristics can be expressed as

$$Y(m,n,t) = -\sum_{i=0}^{d}\sum_{j=0}^{d}\sum_{k=0}^{C_{\text{in}}}|X(m+i,n+j,k) - F(i,j,k,t)|. \tag{12}$$

Without other modifications, the output of PolicyNet using only AdderNet is defined as

$$Y(m,n,t) = f\left(\text{BN}\left(-\sum_{i=0}^{d}\sum_{j=0}^{d}\sum_{k=0}^{C_{\text{in}}}|X-F|\right)\right). \tag{13}$$

The use of AdderNet instead of FC in (13) is called AddPolicyNet; although this approach achieves higher energy efficiency, the performance of the algorithm is decreased. AdderNet uses batch normalization (BN) to normalize the output to solve the problem that the output of (12) is constantly negative. However, BN requires each mini-batch of input data to be independently and evenly distributed. Because AddTD3 uses the experience replay buffer, the data extracted here are randomly selected from the data interacting with the different policies. In addition, with the improvement of the strategy, the agent will collect a new batch of data, which will cause the data distribution in the replay buffer to be more inconsistent, and the stability of mini-batch data cannot be guaranteed, thus leading to the failure of the BN method in AddTD3. To solve this problem, we introduced layer normalization (LN) instead of BN. Different from BN, all the hidden units in the same layer of LN share the same set of mean and variance, so it is not necessary to consider whether the data between the mini-batch are independently and equally distributed. However, AddPolicyNet only with LN is still not suitable for single-output or low-output DRL tasks, such as PVAD physiological control tasks. As shown in Fig. 5, AddPolicyNet cannot guarantee the same performance as PolicyNet due to too few output elements of the PVAD control task. To address this issue, the AddPolicyNet output features must be weighted before output.

Unlike the PolicyNet of TD3 defined by (13), the hidden layer output Y_h and the output layer output Y_o in AddTD3 are calculated by the following equations, respectively,

$$Y_h = \mathrm{ReLU}\left(\mathrm{LN}(-\sum_{i=0}^{d}\sum_{j=0}^{d}\sum_{k=0}^{C_{\mathrm{in}}}|X_h - F_h|)\right) \tag{14}$$

$$Y_o = \tanh\left(W_o\left(\mathrm{ReLU}\left(\mathrm{LN}(-\sum_{i=0}^{d}\sum_{j=0}^{d}\sum_{k=0}^{C_{\mathrm{in}}}|X_o - F_o|)\right)\right) + b_o\right) \tag{15}$$

where X_h and F_h are the inputs and filters of the hidden layer, respectively, and X_o and F_o are the inputs and filters of the output layer, respectively. W_o and b_o are the weight matrix and bias of the output layer, respectively, and tanh and ReLU are the activation functions. AddTD3 can effectively reduce the calculation energy cost without reducing the algorithm performance.

C. Agent Training by Imitation Learning and ADR

Imitative learning [25] and ADR [26] methods were used during the agent training process. In the PVAD control task, the agent without elaborate reward shaping often oscillates near the expected pressure and the convergence rate is very slow. Therefore, imitation learning is introduced into the training process. Referring to [8], we reproduce the PID-based synchronous PVAD controller. Under default parameters, this well-performing PID controller is used to collect the control trajectory and store it in the replay buffer, and the agent is pretrained with reference to [25].

TABLE I
COMPARISON OF SIMULATION EXPERIMENT RESULTS

Controller	Experiment	SAE	$N_{failure}$
AddTD3	Single-factor	32.1±21.1	0
TD3	experiment	33.0±21.8	1
PID	(R_s)	93.0±92.8	9
fuzzy-PID		62.8±70.0	7
ANN-PID		67.1±35.4	0
AddTD3	Single-factor	19.8±21.6	1
TD3	experiment	16.6±10.4	0
PID	(E_{max})	122.7±93.6	14
fuzzy-PID		61.2±80.1	8
ANN-PID		35.1±24.6	0
AddTD3	Single-factor	22.8±12.8	0
TD3	experiment	24.2±14.1	0
PID	(HR)	23.3±16.8	4
fuzzy-PID		23.3±19.1	1
ANN-PID		28.9±15.6	0
AddTD3	Single-factor	39.8±38.1	0
TD3	experiment	42.3±42.2	0
PID	(P_{target})	86.8±89.7	8
fuzzy-PID		70.3±92.1	5
ANN-PID		64.7±45.7	0
AddTD3	Mixed-factor	35.9±22.1	0
TD3	experiment	31.6±16.9	0
PID	(Mix)	121.6±87.1	11
fuzzy-PID		58.0±53.4	6
ANN-PID		62.7±43.4	1

In addition, the ADR method was introduced to train the agents because the gap between the mathematical model and the actual system cannot be ignored. ADR is a method that uses the idea of course learning to gradually increase the difficulty of the environment as the agent strategy improves [26]. Specifically, in the PVAD control task, the difficulty of the environment increases successively as follows: only HR and P_{target} change in the process of round initialization, but not in the experiment. The four parameters HR, E_{max}, R_s, and P_{target} are randomized at the beginning of each round, but the parameters remain the same in each round. These four parameters are randomly initialized at the beginning and during the round. Add a ±20% disturbance to each of the patient's default parameter values until training is complete. Trained PVAD agents can be adapted to different patient physiological conditions and can transfer to MCS without fine-tuning.

IV. EXPERIMENTAL RESULTS

A. Evaluation of Adaptive Ability Performance

To evaluate the adaptive ability performance of the proposed DRL controller, the PID controller, the fuzzy PID controller, and the artificial neural network (ANN)-PID controller are used for comparison. The number of controller failures (N_{failure}) and the sum of absolute errors (SAE) were used to evaluate the performance of different controllers. N_{failure} includes controller oscillations, failure to track the target value, etc. SAE is defined as

$$\mathrm{SAE} = \sum_{i=1}^{n}\left|\overline{P}_{\mathrm{target}} - \overline{P}_{\mathrm{measure}}\right| \tag{16}$$

where $\overline{P}_{\mathrm{expect}}$ is the expected value of mean aortic pressure, $\overline{P}_{\mathrm{measure}}$ is the actual measured mean aortic pressure, and n is

TABLE II
FUZZY RULES OF k_p

e \ ec / $\triangle k_p$	NB	NS	Z	PS	PB
NB	PB	PS	PS	PS	Z
NS	PS	PS	PS	Z	NS
Z	PS	PS	Z	NS	NS
PS	PS	Z	NS	NS	NS
PB	Z	NS	NS	NS	NB

TABLE III
FUZZY RULES OF k_i

e \ ec / $\triangle k_i$	NB	NS	Z	PS	PB
NB	NB	NS	NS	NS	Z
NS	NB	NS	NS	Z	PS
Z	NS	NS	Z	PS	PS
PS	NS	Z	PS	PS	PB
PB	Z	PS	PS	PS	PB

TABLE IV
FUZZY RULES OF k_d

e \ ec / $\triangle k_d$	NB	NS	Z	PS	PB
NB	PS	NB	NB	NB	PS
NS	Z	NS	NS	NS	Z
Z	Z	NS	NS	NS	Z
PS	Z	Z	Z	Z	Z
PB	PB	PS	PS	PS	PB

the number of samples. The smaller the SAE value, the better the performance of the PVAD physiological controller.

In the simulation experiment, the total length of a single simulation experiment was 80 s. The simulation experiment was divided into the single-factor experiment and the mixed-factor experiment. In the single-factor experiment, one of the parameters of R_s, HR, $E_{\max}$, and P_{target} was randomly initialized at the beginning of the experiment and at the 40th second, respectively, while in the mixed-factor experiment, all the parameters will be randomly initialized at the beginning and the 40th second, respectively. The change at the 40th second was a step-change because we assume that if the controller responds correctly to this extreme case, it will also respond correctly to small changes [5]. To verify the adaptive ability of the algorithm and simulate the interpatient and intrapatient variations, we conducted 20 experiments for each group of experiments, 100 experiments in total. Since the simulation environment of this article is the same as that of [8], to ensure the effectiveness of the comparison experiment, the three parameters of PID in (3) are set to $K_a = 5.44$, $K_b = -2.44$, and $K_c = 0.56$, referring to [8] and then fine-tuned. To further reflect the superiority of the DRL controller, we compare the current advanced fuzzy PID controller [12] and ANN-PID controller [36]. The principle and implementation of fuzzy PID refer to [12]. The initial value of fuzzy PID is the same as the PID parameters, and its fuzzy rules are shown in Tables II–IV. The ANN-PID controller [36] is similar to the fuzzy PID controller which is a controller that adaptively adjusts the parameters of the PID using ANN. In this experiment, there are three hidden layers of the neural network in ANN-PID, and the number of units is 16, 32, and 16, respectively.

Fig. 6 and Table I show the simulation results of the AddTD3 controller, TD3 controller, PID controller, fuzzy PID controller, and ANN-PID controller in the PVAD control task. It can be seen that there is almost no difference in the control effect of AddTD3 and TD3 controllers, whether SAE or N_{failure}. Although the cumulative absolute error of the fuzzy PID controller and the ANN-PID controller is smaller than that of the PID controller, its SAE is still twice that of the DRL controller. Notably, changing the patient's heart rate had little effect on the performance of all the controllers, as all of them use synchronous control modes. The controller is controlled according to the sensor signal of the whole cardiac cycle. The value of the aortic pressure signal did not change regardless of the size of the cardiac cycle. In addition, the DRL controller and the ANN-PID controller showed almost no control failures, but in contrast, the fuzzy PID control has a 30% probability of failure. And the PID controller is even worse, failing in nearly half of the experiments. This is because PID control has a weak adaptive ability and may perform well in the face of tasks with fixed patient parameters. However, the PID controller is prone to failure in complex and dynamic environments. Although the fuzzy PID controller can adaptively change the three parameters of the PID according to the fuzzy rules, the artificial fuzzy rules are difficult to be effective in various complex environments. ANN-PID relies on the neural network to adjust PID parameters, but it mainly extracts knowledge from training data and cannot predict the state of the future environment. Unlike them, the DRL controller is based on the Bellman equation and neural network to fit complex multidimensional nonlinear functions. The estimation of the Q value in the training process ensures that the DRL controller can predict the future state to a certain extent. Therefore, the PVAD physiological controller based on DRL is more suitable for CVS-PVAD control tasks where the environment changes frequently and complexly.

B. Evaluation of Energy Efficiency Performance

To evaluate the performance of AddTD3, two indexes should be considered. One is the computational power of the algorithm. The algorithm's energy cost calculation is described in [37]. Each algorithm model in the experiment is implemented by 32-bit floating point numbers, and the energy cost of one addition and multiplication is 0.9 and 3.7 pJ, respectively. The model compression performance of the proposed algorithm is compared by calculating the energy cost of each algorithm during one forward propagation. Another index is SAE, which reflects the control performance of each algorithm, and only mixed-factor experiments were evaluated. This work conducted ablation experiments to compare the performance of TD3, TD3 + Add (BN) (which replaces TD3's FC with AdderNet and uses BN), TD3 + Add (LN) (which replaces TD3's FC with AdderNet and uses LN, but the output layer elements are not reweighted), and the AddTD3 algorithm proposed in this article.

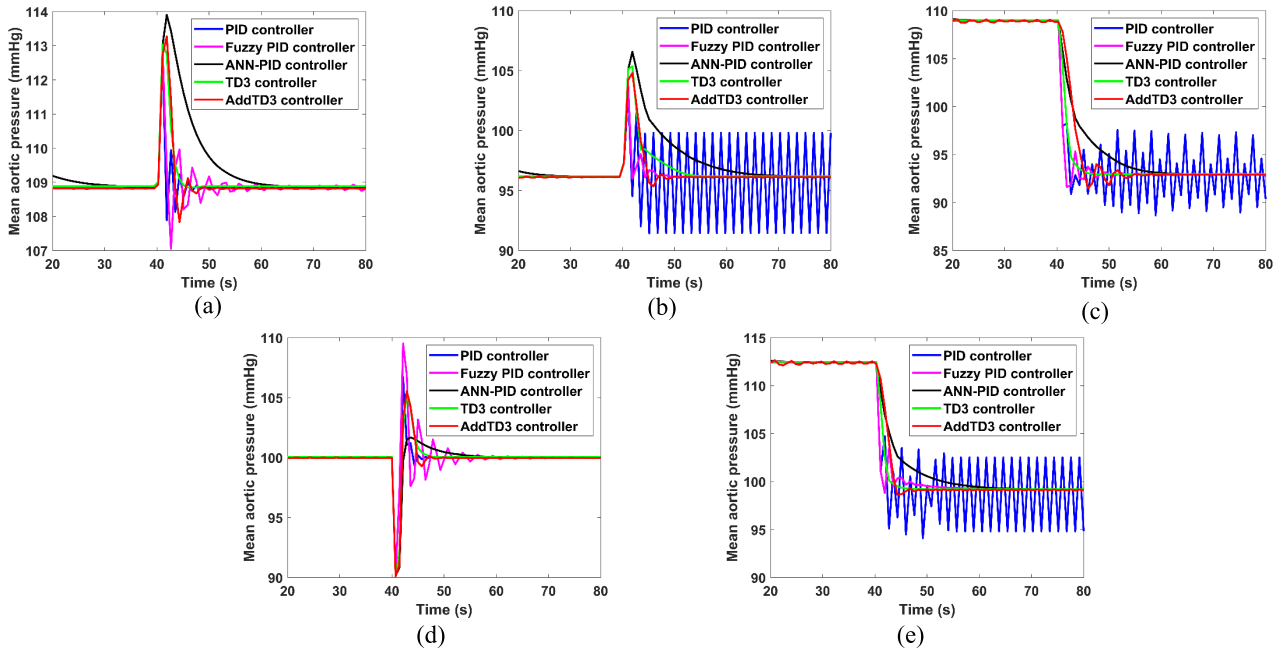


Fig. 6. Control effects of PID, fuzzy PID, ANN-PID, TD3, and AddTD3 controllers in five random experiments. The blue line represents the PID controller, the pink line represents the fuzzy PID controller, the black line represents the ANN-PID controller, the green line represents the TD3 controller, and the red line represents the AddTD3 controller. (a)–(d) Single-factor experiments, and $E_{\max}$, R_s, P_{target}, and HR were changed at the 40th second, respectively. (e) Mixed-factor experiment. The four parameters $E_{\max}$, R_s, P_{target}, and HR were changed at the 40th second.

TABLE V
COMPARISON OF CALCULATED ENERGY COST

Model	#Add	#Mul	Energy cost (pJ)	SAE
AddTD3	36420	536	34835.2	35.9±22.1
TD3	16896	16896	77721.6	31.6±16.9
TD3+Add(BN)	34306	514	32777.2	1550.2±445.4
TD3+Add(LN)	34306	514	32777.2	1550.1±445.4

Considering the nonlinear dynamics of CVS, the PolicyNet for all the algorithms contains two hidden layers each with 128 hidden units. The output layer of AddPolicyNet is additionally reweighted with an FC layer with a hidden unit of 8. All the algorithms' CriticNet is set as a single hidden layer, the hidden unit is 128, and activated with ReLU. In the experiment, the discount factor was set as $\gamma = 0.90$, the learning rate of the Adam optimizer was set as $\text{lr} = 0.001$, the soft update rate of the target network was set as $\tau = 0.005$, the mini-batch size was set as 256, and the size of the experience replay buffer was set as 10^6.

As shown in Table V, the control performance (SAE) of AddTD3 and TD3 is similar, but the calculation energy cost of AddTD3 and TD3 in a forward propagation is 34835.2 and 77721.6 pJ, respectively, and the former is only 44.8% of the latter. In one forward propagation, AddTD3 multiplications are reduced from 16 896 to 556. However, only eight of these multiplication operations are brought by the output layer reweighting, and the rest are brought by the LN. It is worth noting that as the network becomes more complex, the proportion of LN's floating point operations in the whole network will decrease, which means that the proposed AddTD3 can save more computational energy. In addition, as analyzed above, the networks of TD3 + Add (BN) and TD3 + Add (LN) broke down during the training process. The SAE of these two controllers is almost the same because they control

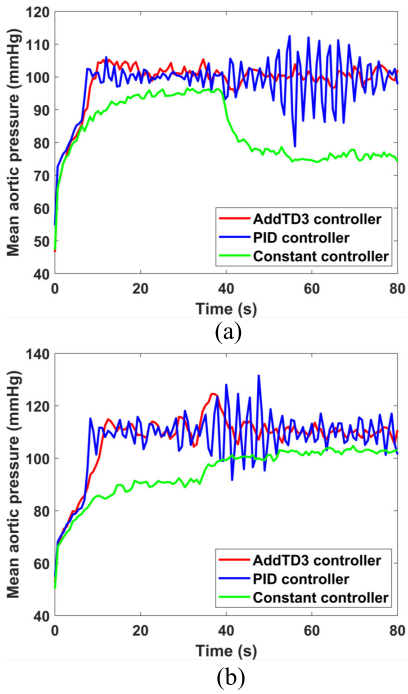


Fig. 7. Control effects of AddTD3, constant, and PID controllers in the MCS experiment. The red line represents the AddTD3 controller, the green line represents the constant controller, and the blue line represents the PID controller. (a) Heart failure experiment, which reduced the maximum myocardial elasticity at the 40th second and increased the vascular resistance. (b) Rest-to-exercise experiment, which increased the heart rate at 40th second and reduced the vascular resistance.

TABLE VI
COMPARISON OF MCS EXPERIMENTAL RESULTS

Controller	Experiment	SAE
AddTD3	Heart failure experiment	148.8
PID		423.5
Constant		1644.7
AddTD3	Rest to exercise	237.1
PID		484.3
Constant		1002.18

the driving pressure to keep decreasing until it touches the minimum drive pressure (which is much lower than the normal pressure) and finally keep working with the minimum drive pressure. The results show that our proposed AddTD3 only reduces the computational energy cost and has no impact on the performance of the algorithm.

C. Overall Evaluation in MCS

To further verify the adaptive ability and transferability of the proposed PVAD controller in the MCS, the PID controller and constant pressure controller are used for comparison. SAE defined by (16) was used as the evaluation index of the MCS experiment. The constant pressure controller is that the controller adopts constant driving pressure to drive PVAD based on the synchronous mode and delayed injection. In the MCS experiment, the total length of a single experiment was 80 s, and the physiological parameters of patients were initialized at the beginning of the experiment and changed at the 40th second. The MCS experiment was divided into the heart failure experiment and the rest-to-exercise experiment. The heart failure experiment reduced the heart simulator's pumping capacity at the 40th second and increased the patient's vascular resistance. The exercise experiment increased the patient's heart rate while reducing the vascular resistance. Similar to the simulation experiment, the step-change was also used in the MCS experiment.

The MCS experimental results are shown in Table VI and Fig. 7. AddTD3 in the experiment was directly transferred to MCS after the simulation training without fine-tuning. All kinds of physiological control systems proposed for VAD in recent years [4], [5], [6], [7], [15] are mostly set or designed based on the simulation models, and their performance in the real world is unknown. To solve this problem, we introduced ADR into algorithm training. With the high adaptive ability of AddTD3 and ADR, the PVAD physiological controller designed by us can be directly transferred to MCS for pressure regulation tasks without fine-tuning. It can be seen from the results that the proposed controller has higher adaptive ability and control accuracy in the physical system, and its SAE is only about 42% of the PID controller and can stabilize at a faster speed to complete the adjustment task in the MCS experiment.

V. CONCLUSION

This article reports a PVAD physiological control system based on the energy-efficient DRL algorithm AddTD3. The proposed DRL-based PVAD physiological control system has a high adaptive ability and can adapt to different physiological parameters of patients. The proposed AddTD3 can effectively reduce the energy cost of the TD3 algorithm. The experimental results show that the proposed AddTD3 controller is with lower cumulative absolute error than the traditional PID controller and can be migrated to the MCS without fine-tuning. It can reduce the energy cost of the traditional DRL algorithm TD3 to 44.8% without reducing the performance.

REFERENCES

[1] C. P. Kyriakopoulos et al., "LVAD as a bridge to remission from advanced heart failure: Current data and opportunities for improvement," *J. Clin. Med.*, vol. 11, no. 12, p. 3542, Jun. 2022.

[2] E. L. Wu, M. C. Stevens, J. P. Pauls, and U. Steinseifer, "First-generation ventricular assist devices," in *Mechanical Circulatory and Respiratory Support*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 93–115.

[3] E. A. Rose et al., "Long-term use of a left ventricular assist device for end-stage heart failure," *New England J. Med.*, vol. 345, no. 20, pp. 1435–1443, 2018.

[4] M. Fetanat, M. Stevens, C. Hayward, and N. H. Lovell, "A physiological control system for an implantable heart pump that accommodates for interpatient and intrapatient variations," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1167–1175, Apr. 2020.

[5] M. Fetanat, M. Stevens, C. Hayward, and N. H. Lovell, "A sensorless control system for an implantable heart pump using a real-time deep convolutional neural network," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 10, pp. 3029–3038, Oct. 2021.

[6] B. C. Ng et al., "Application of multiobjective neural predictive control to biventricular assistance using dual rotary blood pumps," *Biomed. Signal Process. Control*, vol. 39, pp. 81–93, Jan. 2018.

[7] T. Li et al., "Intelligent and strong robust CVS-LVAD control based on soft-actor-critic algorithm," *Artif. Intell. Med.*, vol. 128, Jun. 2022, Art. no. 102308.

[8] T. D. Cordeiro, D. L. Sousa, I. A. Cestari, and A. M. N. Lima, "Controlling the cardiovascular response under synchronized ventricular assist device support," in *Proc. IEEE Int. Conf. Autom. (ICA-ACCA)*, Oct. 2016, pp. 1–6.

[9] M. Bakouri, "Developing a feedback physiological control for ventricular assist devices: A simulation study," in *Proc. 4th Int. Conf. Appl. Autom. Ind. Diag. (ICAAID)*, vol. 1, Mar. 2022, pp. 1–4.

[10] M. C. Stevens, A. Stephens, A.-H. H. AlOmari, and F. Moscato, "Physiological control," in *Mechanical Circulatory and Respiratory Support*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 627–657.

[11] J. Li, R. Ji, X. Liang, S. S. Ge, and H. Yan, "Command filter-based adaptive fuzzy finite-time output feedback control of nonlinear electrohydraulic servo system," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022.

[12] M. Khaledi, M. Dehghani, M. Mohammadi, and R. Abolpour, "Controller design for left ventricular assist devices in patients with heart failure," in *Proc. 27th Nat. 5th Int. Iranian Conf. Biomed. Eng. (ICBME)*, Nov. 2020, pp. 326–332.

[13] M. Azizkhani and Y. Chen, "Supervised adaptive fuzzy control of LVAD with pulsatility ratio modulation," in *Proc. IEEE 18th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2022, pp. 2429–2434.

[14] M. Fetanat, M. Stevens, C. Hayward, and N. Lovell, "Adaptive sensorless control of LVAD using deep convolutional neural network," *J. Heart Lung Transplantation*, vol. 40, no. 4, p. S172, Apr. 2021.

[15] T. D. Cordeiro, D. L. Sousa, I. A. Cestari, and A. M. N. Lima, "A physiological control system for ECG-synchronized pulsatile pediatric ventricular assist devices," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101752.

[16] A. Dionisi, D. Marioli, E. Sardini, and M. Serpelloni, "Autonomous wearable system for vital signs measurement with energy-harvesting module," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 6, pp. 1423–1434, Jun. 2016.

[17] S. Shirmohammadi, K. Barbe, D. Grimaldi, S. Rapuano, and S. Grassini, "Instrumentation and measurement in medical, biomedical, and healthcare systems," *IEEE Instrum. Meas. Mag.*, vol. 19, no. 5, pp. 6–12, Oct. 2016.

[18] A. Ding, Y. Qin, B. Wang, L. Jia, and X. Cheng, "Lightweight multi-scale convolutional networks with adaptive pruning for intelligent fault diagnosis of train bogie bearings in edge computing scenarios," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–13, 2023.

[19] B. Wang, F. Ma, L. Ge, H. Ma, H. Wang, and M. A. Mohamed, "Icing-EdgeNet: A pruning lightweight edge intelligent method of discriminative driving channel for ice thickness of transmission lines," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[20] J. Deng, W. Jiang, Y. Zhang, G. Wang, S. Li, and H. Fang, "HS-KDNet: A lightweight network based on hierarchical-split block and knowledge distillation for fault diagnosis with extremely imbalanced data," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.

[21] D. Song, Y. Wang, H. Chen, C. Xu, C. Xu, and D. Tao, "AdderSR: Towards energy efficient image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 15643–15652.

[22] A. De Angelis, G. De Angelis, and P. Carbone, "Low-complexity 1-bit detection of parametric signals for IoT sensing applications," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–8, 2021.

[23] A. Z. Hunsberger, "Modeling and analysis of interactions between a pulsatile pneumatic ventricular assist device and the left ventricle," Ph.D. dissertation, School Eng., Univ. Pittsburgh, Pittsburgh, PA, USA, 2005.

[24] T. Li, H. Li, W. Cui, N. Xie, X. Li, and Y. Wang, "Design and intelligent control of mock circulation system to reproduce patient-specific physiological indexes," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103987.

[25] M. Vecerik et al., "Leveraging demonstrations for deep reinforcement learning on robotics problems with sparse rewards," 2017, *arXiv:1707.08817*.

[26] I. Akkaya et al., "Solving Rubik's cube with a robot hand," 2019, *arXiv:1910.07113*.

[27] F. Tueche, Y. Mohamadou, A. Djeukam, L. C. N. Kouekeu, R. Seujip, and M. Tonka, "Embedded algorithm for QRS detection based on signal shape," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[28] D. Lai, F. Zhang, and C. Wang, "A real-time QRS complex detection algorithm based on differential threshold method," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 129–133.

[29] S. Sun, Y. Yin, X. Wang, and D. Xu, "Robust visual detection and tracking strategies for autonomous aerial refueling of UAVs," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4640–4652, Dec. 2019.

[30] Y. Sun, J. Xu, C. Chen, and W. Hu, "Reinforcement learning-based optimal tracking control for levitation system of maglev vehicle with input time delay," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022.

[31] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proc. 35th Int. Conf. Mach. Learn.*, 2018, pp. 1587–1596.

[32] Z. Wu, Y. Deng, J. Liu, and L. Wang, "A reinforcement learning-based reconstruction method for complex defect profiles in MFL inspection," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[33] H. Chen et al., "AdderNet: Do we really need multiplications in deep learning?" in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1465–1474.

[34] Y. Wang et al., "AdderNet and its minimalist hardware design for energy-efficient artificial intelligence," 2021, *arXiv:2101.10015*.

[35] Y. Xu, C. Xu, X. Chen, W. Zhang, C. Xu, and Y. Wang, "Kernel based progressive distillation for adder neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 12322–12333.

[36] G. Han, W. Fu, W. Wang, and Z. Wu, "The lateral tracking control for the intelligent vehicle based on adaptive PID neural network," *Sensors*, vol. 17, no. 6, p. 1244, May 2017.

[37] M. Horowitz, "Computing's energy problem (and what we can do about it)," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 10–14.

Te Li (Member, IEEE) received the B.S. degree in automation from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2010, and the Ph.D. degree in mechanical engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, in 2016.

He is currently an Associate Professor with the School of Mechanical Engineering, Dalian University of Technology, Dalian, China. His research interests include software/bionic robot innovative design and manufacturing methods, deep reinforcement learning intelligent control, intelligent equipment, and industrial robots.

Wenbo Cui received the B.S. degree in mechanical design, manufacturing, and automation from Northeast Agricultural University, Harbin, China, in 2020. He is currently pursuing the M.A.Sc. degree in machinery manufacturing and automation with the Dalian University of Technology, Dalian, China.

His research interests include left ventricular assist device (LVAD) control and deep reinforcement learning.

Xingjian Liu (Member, IEEE) received the B.E. degree in electronic information engineering and the Ph.D. degree in materials processing engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2013 and 2018, respectively.

He was a Post-Doctoral Fellow with the Department of Mechanical and Industrial Engineering and the Robotics Institute, University of Toronto, Toronto, ON, Canada. He is currently an Associate Professor with the State Key Laboratory of High-Performance Precision Manufacturing and School of Mechanical Engineering, Dalian University of Technology, Dalian, China. His research interests include computer vision, 3-D sensing and metrology, and robotics.

Xu Li received the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, Dalian, China, in 2018.

He is currently an Associate Professor with the School of Mechanical Engineering, Dalian University of Technology. His research interests include time-delay systems, switched systems, and precision motion.

Nan Xie received the M.Eng. degree in measuring techniques and instruments from the Dalian University of Technology, Dalian, China, in 2018, where he is currently pursuing the Ph.D. degree in machinery manufacturing and automation.

His research interests are bionic heart design and hydrodynamics.

Yongqing Wang received the M.S. degree in electromechanical control and automation, and the Ph.D. degree in mechatronic engineering from the Dalian University of Technology, Dalian, China, in 1994 and 2002, respectively.

He is currently a Professor with the School of Mechanical Engineering, Dalian University of Technology. His research interests include measurement-processing integrated manufacturing theory and technology, equipment in machine measurement theory and technology, and ultralow temperature cooling processing theory.