

DOI: 10.3901/JME.2022.11.133

# 基于多尺度注意力深度强化学习网络的行星 齿轮箱智能诊断方法\*

王 辉<sup>1</sup> 徐佳文<sup>1</sup> 严如强<sup>1,2</sup>

(1. 东南大学仪器科学与工程学院 南京 210096;  
2. 西安交通大学机械工程学院 西安 710049)

**摘要:** 针对行星齿轮箱在多工况下故障表征具有差异性, 现有方法中存在特征提取不足, 且泛化性和诊断准确率低的问题, 提出一种基于多尺度深度注意 Q 网络(Multiscale deep attention Q network, MSDAQN)的深度强化学习行星齿轮箱多工况智能诊断方法。首先定义分类马尔科夫决策过程描述诊断过程, 建立故障诊断模拟环境; 其次构造 MSDAQN 深度强化学习智能体结构, 其通过多尺度卷积神经网络提取多尺度故障特征, 并利用自适应通道注意力进行加权融合、突出关键信息; 最后依据所建智能体与诊断模拟环境交互的经验, 自主学习最佳诊断策略。通过行星齿轮箱的多工况试验和实际案例的测试与分析, 表明所提方法具有更高的诊断准确率和较强的工况适应性。

**关键词:** 行星齿轮箱; 故障诊断; 深度强化学习; 多尺度特征; 通道注意力

**中图分类号:** TH165; TH17

## Multi-Scale Attention Based Deep Reinforcement Learning for Intelligent Fault Diagnosis of Planetary Gearbox

WANG Hui<sup>1</sup> XU Jiawen<sup>1</sup> YAN Ruqiang<sup>1,2</sup>

(1. School of Instrument Science and Engineering, Southeast University, Nanjing 210096;  
2. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049)

**Abstract:** Given the difference in fault characterization of the planetary gearbox under multiple working conditions, the existing methods have problems of insufficient feature extraction, low generalization, and diagnosis accuracy, an intelligent diagnosis method using multiscale deep attention Q network (MSDAQN) of deep reinforcement learning for planetary gearboxes under multi-working condition is proposed. Classification Markov decision process is defined to describe the diagnosis process and establish the simulation environment of fault diagnosis. The structure of an MSDAQN deep reinforcement learning agent is designed, which extracts multi-scale fault features by a multi-scale convolutional neural network and uses adaptive channel attention to reweight them and highlight key fault information. Finally, using the interaction experience between the developed agent and diagnosis simulation environment, the optimal diagnosis strategy is learned autonomously. The test and analysis of planetary gearbox's multi-working condition test and actual cases show that the proposed method has higher diagnostic accuracy and strong adaptability.

**Key words:** planetary gearbox; fault diagnosis; deep reinforcement learning; multi-scale features; channel attention

## 0 前言

行星齿轮箱(Planetary gearbox, PGB)作为风机、

直升机、舰船等装备中传动系统的核心部件<sup>[1]</sup>, 承载着系统的变速与动力传递<sup>[2]</sup>, 但恶劣的环境常导致其关键零件(轴承、齿轮等)不可避免地发生磨损或疲劳裂纹等故障, 影响装备的安全性<sup>[1,3]</sup>。因此, 研究先进的行星齿轮箱故障诊断技术, 保障其可靠性至关重要。

近年来, 受传感技术和深度学习(Deep learning,

\* 国家重点研发计划资助项目(2018YFB1702400)。20210829 收到初稿,  
20220206 收到修改稿

DL)<sup>[4]</sup>的驱动发展,机械故障诊断步入了“大数据”时代<sup>[5]</sup>,基于 DL 的智能诊断,如深度信念网络(Deep belief network, DBN)<sup>[6]</sup>、堆叠自编码器(Stacked autoencoder SAE)<sup>[7]</sup>、卷积神经网络(Convolutional neural network, CNN)<sup>[8]</sup>等,无需人工提取特征、专家知识和经验,且具有强大的特征学习能力,已成为故障诊断领域的研究热点<sup>[5]</sup>。尤其, CNN 作为 DL 的代表,在智能诊断中受到极大关注<sup>[9]</sup>。WANG 等<sup>[10]</sup>提出基于短时傅里叶变换和二维 CNN 的电机智能诊断;ZHANG 等<sup>[11]</sup>提出第一层具有宽卷积核的 CNN(Deep convolutional neural networks with wide kernel, WDCNN)用于滚动轴承故障诊断;叶壮等<sup>[12]</sup>提出结合经验模态分解(Empirical mode decomposition, EMD)和多通道加权 CNN 的齿轮箱智能诊断;它们的性能均远超支持向量机(Support vector machine, SVM)、深度神经网络(Deep neural network, DNN)等,但单尺度 CNN 难以学习鲁棒的多尺度特征以适应复杂工况。为此,LIU 等<sup>[13]</sup>提出基于多尺度核残差 CNN 的非平稳状态下感应电机的故障诊断;唐斯等<sup>[14]</sup>提出融合注意力机制的多尺度 CNN(Multiscale CNN, MSCNN)实现轴承智能诊断。以上方法虽取得极好的效果,但多是被动学习样本与故障模式间的强映射关系,缺乏主动探索,且不符合人类认知过程,即知识是通过与环境交互逐渐获取的,一定程度上影响诊断的智能性和泛化性。

深度强化学习<sup>[15]</sup>(Deep reinforcement learning, DRL)不仅具有 DL 的强感知能力,还拥有强化学习中通过与环境进行交互实现自主决策的优势,增强了智能性和通用性,已在控制、游戏等领域取得突破<sup>[16-20]</sup>。如 2016 年,DeepMind 团队推出的 AlphaGo<sup>[16]</sup>,战胜围棋顶级高手李世石。近年来,DRL 也被尝试应用于智能诊断。DING 等<sup>[18]</sup>将 SAE 和深度 Q 网络(Deep Q network, DQN)相结合用于滚动轴承故障诊断,获得较好的效果,但其完全忽略状态间的联系,交互规则并非严格的马尔科夫决策过程(Markov decision process, MDP);WANG 等<sup>[19]</sup>在分类 MDP<sup>[20]</sup>(Classification MDP, CMDP)基础上,利用同步提取变换和 DQN 实现行星齿轮箱的智能诊断,增强了诊断泛化性,但其受预提取时频表征的制约。同时,以上方法均存在单尺度模型在特征学习上的不足。

鉴于此,本文提出一种基于多尺度深度注意 Q 网络(Multiscale deep attention Q network, MSDAQN)的 DRL 行星齿轮箱智能诊断方法。基于 DQN 模型,通过构造多尺度 CNN 增强特征学习的丰富性,并利用通道注意力对特征加权融合,在 DRL 框架下实

现诊断策略的自主学习及故障识别,为解决多工况下行星齿轮箱智能诊断提供新思路。

## 1 理论背景

### 1.1 强化学习(Reinforcement learning, RL)

如图 1 所示,RL 主要由智能体(Agent)和环境(Environment)组成,其依据 MDP 解决序贯决策任务,MDP 主要由元组 $\{S, A, P, R, \gamma\}$ 组成,其中, $S$  为状态空间, $A$  为动作空间, $P$  为状态转移策略, $R$  为奖励函数, $\gamma$  为折扣因子。不同于监督学习,RL 中智能体面对问题时并不知其“答案”,是通过与环境交互、试错,根据环境的奖励进行学习。

据图 1,RL 过程可描述为:智能体感知当前环境状态  $s_t \in S$ ,选择执行动作  $a_t \in A$ ,环境反馈智能体奖励  $r_t$ ,并转至下一状态  $s_{t+1} \in S$ ,可表示为 $\{s_t, a_t, r_t, s_{t+1}\}$ 。同时,智能体根据奖励调整策略  $\pi$ <sup>[21]</sup>,以寻找一个最佳策略  $\pi^*$ ,使其能够在任意状态  $s$  和时间步  $t$  下,均可获得最大的累积奖励

$$\pi^* = \arg \max_{\pi} E[R_t | s_t = s], \forall s \in S, \forall t \geq 0 \quad (1)$$

式中,  $R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'}$  表示累积奖励; $T$  为终止时间步; $\gamma \in [0, 1]$ ;  $E$  为期望函数; $t'$  为未来时间步。

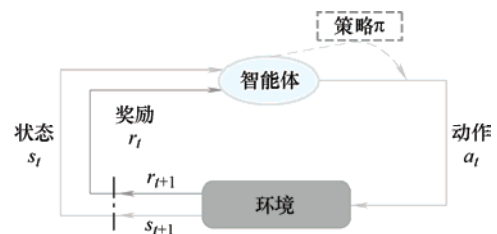


图 1 RL 基本框架

### 1.2 Q-learning

Q-learning 作为一种免模型 RL 算法,通过创建一个  $h_1 \times h_2$  的表  $Q \in \mathbf{R}^{h_1 \times h_2}$  计算状态-动作的最大期望未来奖励, $Q$  表中每个  $Q$  值是通过计算动作值函数( $Q^\pi(s, a)$ )进行更新学习,可定义为

$$Q^\pi(s, a) = E[R_t | s_t = s, a_t = a, \pi] \quad (2)$$

根据贝尔曼方程<sup>[15]</sup>,式(2)可以被表达为

$$Q^\pi(s, a) = E[r_t + \gamma Q^\pi(s_{t+1}, a_{t+1}) | s_t = s, a_t = a, \pi] \quad (3)$$

再依据经验样本  $e = \{s, a, r, s'\}$ ,通过时间差分进行动作值函数  $Q(s, a)$  的更新

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_a Q(s', a) - Q(s, a)] \quad (4)$$

式中,  $\alpha$  表示学习率; $s'$  为下一状态。根据式(1)

和(3), 最佳动作值函数  $Q^*(s, a)$  对应于在最佳策略  $\pi^*$  下的最大期望回报, 可定义为

$$Q^*(s, a) = E_{s' \sim \varepsilon} \left[ r + \gamma \max_{a'} Q^*(s', a') | s, a \right] \quad (5)$$

式中,  $a'$  是下一动作;  $\varepsilon$  表示环境模型。

根据式(5)最佳动作值函数  $Q^*(s, a)$ , 可得最佳策略  $\pi^*$  指导智能体在状态  $s$  下选择最佳动作  $a^*$

$$\pi^*(s) = \arg \max_a Q^*(s, a) \quad (6)$$

Q-learning 虽在简单任务中取得不错的效果, 但对于复杂的高维状态空间、动作空间等, Q 表往往难以适用<sup>[15, 17]</sup>。

### 1.3 深度 Q 网络

受益于 DL 的发展和 Q-learning, MNIH 等<sup>[15]</sup>提出了第一个 DRL 算法, DQN, 在 Atari 2600 游戏中超越专业玩家, 且有较强的通用性。DQN 算法主要利用 DNN 估计动作值函数, 并依据经验回放 (Experience replay, ER)<sup>[15]</sup>保障学习的稳定性, 算法更新如图 2, 其中当前 Q 网络 (Eval-Net) 和目标 Q 网络 (Target-Net) 具有相同结构, 分别用  $Q(s, a; \theta_i)$  和  $Q(s', a'; \theta_i^-)$  表示,  $\theta_i$  和  $\theta_i^-$  分别为上述网络的参数。在学习过程中, 最小化每次迭代  $i$  的均方误差 (Mean square error, MSE) 损失  $L(\theta_i)$  进行网络训练

$$L(\theta_i) = E_{e \sim U(D)} \left[ \left( r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s, a; \theta_i) \right)^2 \right] \quad (7)$$

式中,  $e \sim U(D)$  表示从经验回放池中抽取经验样本,  $U(D)$  表示存储经验数据的经验回放池, 目标 Q 网络 (Target-Net) 的目标 Q 值计算通过下式

$$y_i = r + \gamma \max_{a'} Q(s', a'; \theta_i^-) \quad (8)$$

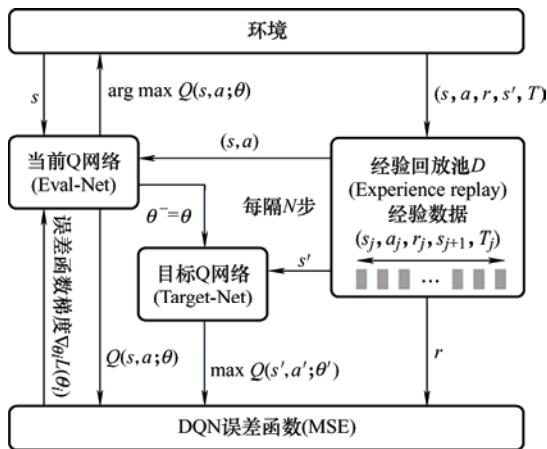


图 2 DQN 算法更新流程

最后, 通过从  $U(D)$  随机抽取小批量经验数据  $e$ , 根据式(7)中损失函数  $L(\theta_i)$  对参数  $\theta$  的偏导数

$\nabla_{\theta_i} L(\theta_i)$  更新模型, 其通过下式获得

$$\nabla_{\theta_i} L(\theta_i) = E_{e \sim U(D)} [y_i - Q(s, a; \theta_i) \nabla Q(s, a; \theta_i)] \quad (9)$$

同时, 利用随机梯度下降法进行网络参数的更新, 并采用延迟  $N$  步更新 Target-Net 网络, 其是通过复制 Eval-Net 的网络参数直接给 Target-Net 网络。

受此启发, 本文将利用一维 CNN 的感知优势, 立足于 DQN 深度强化学习算法, 以实现行星齿轮箱的智能诊断。

## 2 所提基于 MSDAQN 的行星齿轮箱智能诊断方法

本节定义一种用于描述故障诊断的 CMDP, 在此基础上, 设计一种多尺度深度注意 Q 网络智能诊断模型, 并给出详细的诊断流程。

### 2.1 分类马尔科夫决策过程 (CMDP)

由节 1 可知, DQN 是基于 MDP 解决序贯决策任务, 为此, 本文将行星齿轮箱故障诊断看作一个识别游戏, 采用由元素  $\{S, A, R, P, \gamma\}$  组成的 CMDP, 将其转化为一个简单的序贯决策任务<sup>[19-20]</sup>。假设训练数据集为  $D_{train} = \{(x_1, l_1), (x_2, l_2), \dots, (x_n, l_n)\}$ , 其中  $x_i$  表示第  $i$  个样本,  $l_i$  表示  $x_i$  相应的标签,  $n$  为样本总量。CMDP 中相关变量定义如下所述。

**S:** 环境状态  $s_t$  取决于训练集  $D_{train}$  中样本  $x_t$ , 环境初始状态为  $x_1$ , 每次开始新的情节 (Episode) 时, 随机打乱训练集  $D_{train}$ 。

**A:** 智能体的动作  $a_t$  是猜测或预测一个样本的类别, 与训练集  $D_{train}$  中标签相关联。对于  $K$  分类任务,  $A = \{0, 1, \dots, K-1\}$ 。

**R:** 奖励  $r_t$  来自环境反馈, 用于衡量智能体在状态  $s_t$  时执行动作  $a_t$  的优劣, 指导智能体学习最佳的识别策略  $\pi_\theta^*$ 。其定义如下

$$R(s_t, a_t, l_t) = \begin{cases} 1 & a_t = l_t \\ -1 & \text{其他} \end{cases} \quad (10)$$

式(10)可解释为智能体正确猜测样本类别时, 获得一个正奖励, 否则得到一个惩罚。

**P:** 此处  $P$  是确定性的, 表示从当前状态  $s_t$  到下一状态  $s_{t+1}$  是根据  $D_{train}$  中样本顺序进行转移。

$\gamma$ :  $\gamma \in [0, 1]$ , 用于平衡当前和未来奖励, 鉴于状态间相关性较小, 智能体更注重当前及时奖励, 因此设置  $\gamma=0.1$ 。

**Episode:** 一个情节 Episode 包含一轮识别过程, 即  $Episode = \{s_1, a_1, r_1, s_2, a_2, r_2, \dots, s_t, a_t, r_t\}$ ,  $t \leq n$ , 当

智能体识别错误或  $t=n$  时, 便结束此情节。

$\pi_\theta$ : 含参数  $\theta$  的诊断策略, 指导智能体在状态  $s_t$  时选择相应的动作  $a_t$ , 完成诊断行为。

基于上述 CMDP 设定的交互规则, 设计相应的 DRL 学习模型便可自主学习最佳诊断策略  $\pi_\theta^*$ :  $S \rightarrow A$ , 从而完成行星齿轮箱的智能诊断。

## 2.2 MSDAQN 模型结构设计

基于前述 DQN 原理、定义的 CMDP 和一维 CNN, 本文设计了一种多尺度深度注意 Q 网络 (MSDAQN) 的智能体结构, 总体网络结构如图 3 所示, 主要包括多尺度特征融合模块、自适应通道注意力 (Adaptive channel attention, ACA) 模块和通用卷积、池化模块等。具体如下

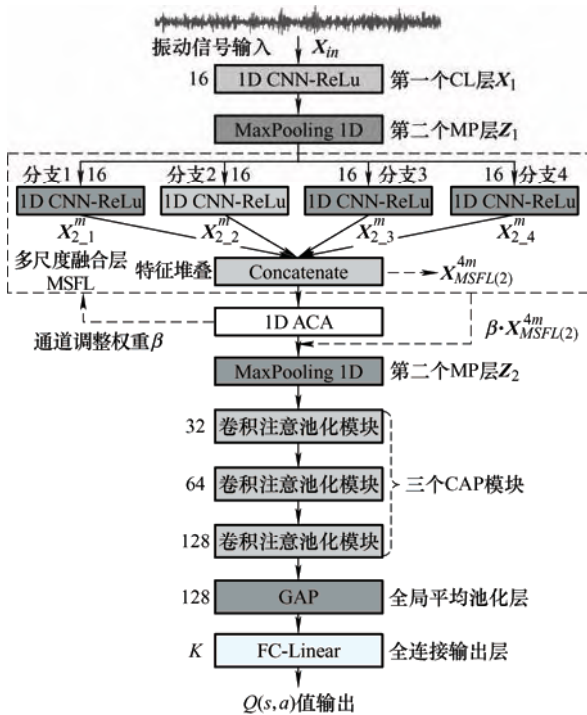


图 3 MSDAQN 网络结构

输入层 (Input) 为一维振动信号  $X_{in} \in \mathbf{R}^{d \times 1}$  的环境状态, 首先经过第一个卷积-池化组合操作进行初特征提取, 这里采用宽的卷积核增大感受视野, 提取低频特征, 抑制高频噪声<sup>[11]</sup>, 再利用最大池化 (Max pooling, MP) 层进行特征缩放映射, 减少特征维度, 降低模型参数量, 其中, 卷积层 (Convolutional layer, CL) 和 MP 层数学上可分别表示为

$$X_l^i = \text{Act} \left( \sum_{c=1}^C W_l^{(c,i)} * X_{l-1}^{(c)} + B_l^{(i)} \right) \quad (11)$$

式中,  $\text{Act}(\bullet)$  是 ReLu 激活函数,  $X_{l-1}^{(c)}$  是  $l-1$  层第  $c$  通道特征映射,  $C$  为通道数,  $X_l^i$  为输出  $l$  层第  $i$  通道特征映射, “\*” 为卷积运算,  $W_l^{(c,i)}$  为权重矩阵,

$B_l^{(i)}$  为偏置项。

$$Z_{l(P,j)}^i = \text{MP}_{(P,j)}(X_l^i) \quad (12)$$

式中,  $\text{MP}_{(P,j)}(\bullet)$  表示第  $j$  个池化区域以池化尺寸  $P$  取最大池化运算,  $Z_{l(P,j)}^i$  为  $l$  层 MP 池化输出。

通过上述第一个卷积-池化组合操作后进行初特征提取, 再将其经过多尺度融合层 (Multi-scale fusion layer, MSFL) 实现多尺度特征学习, 其通过采用 4 个不同大小的卷积核, 根据式 (11) 中卷积操作提取不同频率段特征, 并将其沿通道维度进行堆叠, 通过 MSFL 增加了多工况下特征学习的多样性, 更好适应多工况下故障诊断, 其可表示为

$$X_{MSFL(l)}^{4m} = \text{Concat}(X_{l-1}^m, X_{l-2}^m, X_{l-3}^m, X_{l-4}^m) \quad (13)$$

式中,  $X_{MSFL(l)}^{4m}$  多尺度融合后的特征,  $X_{l-i}^m$  为  $l$  层不同尺度分支提取的特征映射, 此处  $i=1, \dots, 4$ ,  $\text{Concat}(\bullet)$  表示沿通道维度进行特征堆叠。

同时, 采用 ACA 模块对多尺度特征  $X_{MSFL(l)}^{4m}$  沿通道维度加权融合, 对不同尺度下提取特征进行重要程度分配, 以突出关键故障特征、降低信息冗余, 获得判别性强的故障特征。ACA 结构如图 4, 具体为

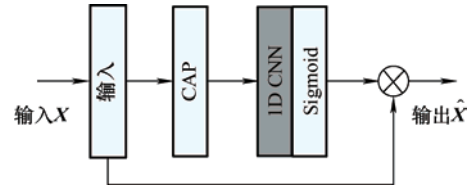


图 4 1D ACA 结构图

假设输入特征映射为  $X \in \mathbf{R}^{L \times 1 \times C}$ ,  $L$  为特征长度, 将  $X$  经过全局平均池化 (Global average pooling, GAP) 后, 通过卷积核大小为  $k$  的 1D 卷积运算捕获跨通道信息, 再经过 Sigmoid 激活函数  $\sigma$  获取特征通道调整权重  $\beta$ , 以分配不同尺度下每个特征通道一个权值, 从而对  $X$  沿通道维度加权融合, 输出新特征映射  $\hat{X}$ 。

$$X_{\text{GAP}} = \frac{1}{L} \sum_{j=1}^L X_{c,j} \quad (14)$$

$$\beta = \sigma(\text{Conv1d}(X_{\text{GAP}})) = [\beta_1, \beta_2, \dots, \beta_C] \quad (15)$$

$$\hat{X} = \beta \cdot X \quad (16)$$

式中,  $X_{\text{GAP}}$  为 GAP 特征向量,  $\text{Conv1d}(\bullet)$  为一维卷积操作, 这里  $C$  的大小为  $4m$ , 其卷积核大小  $k$  根据式 (17) 自适应选择<sup>[22]</sup>。



$$k = \left\lfloor \frac{\log_2(C)}{\lambda} + \frac{b}{\lambda} \right\rfloor_{\text{odd}} \quad (17)$$

式中,  $\lfloor \cdot \rfloor_{\text{odd}}$  表示向下取奇数操作,  $\lambda$  和  $b$  为映射函数的参数。

然后, 将经过 ACA 加权融合后的特征依次经过一个 MP 层, 和 3 个卷积-注意-池化(Convolution attention pooling, CAP)模块, 进一步实现深层故障特征学习, 其中 CAP 模块是前述卷积层、ACA 和 MP 层的依次组合操作; 其次采用一个 GAP 操作进行特征整合, 获取全局故障特征, 且减少网络参数; 最后, 通过一个  $K$  维全连接层(Fully-connected layer, FC)输出动作值函数  $Q(s, a)$ , 完成值函数估计。

### 2.3 所提方法步骤流程

基于前述 DQN 算法更新流程(图 2)和建立的 MSDAQN 深度强化学习智能体结构(图 3), 所提方法用于行星齿轮箱智能诊断的流程如图 5 所示, 主要包括以下步骤。

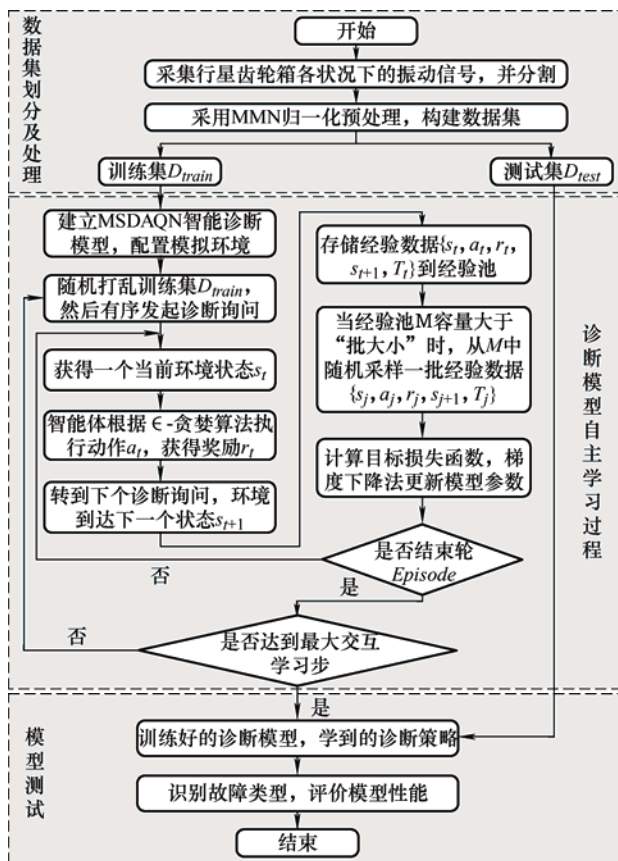


图 5 诊断模型学习及测试流程

步骤 1: 通过振动传感器采集 PGB 运行在多工况下各健康状态的振动数据, 并对其无重叠分割。

步骤 2: 采用最小-最大归一化(Min-max normalization, MMN)对分割后的振动数据归一化预处理, 归一化范围为  $[-1, 1]$ , 构建训练集和测试集。

步骤 3: 建立 MSDAQN 智能诊断模型, 构建诊断智能体交互的模拟环境, 设定学习轮次、奖励函数、学习率等参数。

步骤 4: 在模拟环境下, 智能体通过与环境交互自主学习最佳诊断策略, 学习过程中通过梯度下降法更新网络参数。

步骤 5: 在所学诊断策略下, 将测试样本逐个输入诊断模型识别, 通过评价指标分析模型性能。

### 2.4 所提方法参数设定

基于上述诊断流程, 在建立的 MSDAQN 网络结构中(图 3), 所有 MP 层的池化尺寸均为 2, 步幅为 2, “valid” 填充; 第一层卷积采用 16 个  $64 \times 1$  的宽卷积核, 卷积步幅为 8; 第二层为四分支的 MSFL 层, 每分支均有 16 个卷积核, 卷积核大小分别为  $3 \times 1$ 、 $5 \times 1$ 、 $7 \times 1$ 、 $11 \times 1$ ; 第三、四、五层均为 CAP 组合层, 卷积核数量依次为 32、64 和 128, 卷积核大小均为  $3 \times 1$ ; 模型中所有卷积填充方式均为 “same”; ACA 模块有相同的卷积结构, 式(17)中参数  $\lambda$  和  $b$  分别为 2 和 1; 最后 GAP 得 128 个全局特征输入 FC 层, 输出  $Q$  值。

此外, 在所提 MSDAQN 方法中, 状态为  $1 \times 2048$  的振动信号; 经验回放池  $D$  的存储量是 100 000; 在模型学习中,  $\epsilon$  贪婪算法<sup>[15]</sup>用于智能体对环境的 “探索”,  $\epsilon$  值采用线性退火策略从 1 至 0.01 间线性变化, 其步幅为 100 000; 智能体与环境交互次数为 120 000 步; RAdam (Rectified Adam, RAdam)<sup>[23]</sup> 优化器用于更新 Eval-Net 网络, 学习率为 0.001, “批大小” 为 128; 目标网络 Target-Net 的延迟更新步  $N$  为 10 000。

## 3 试验验证与分析

为了验证所建诊断模型的有效性, 在 Keras2.1.6 框架内以 Tensorflow1.8.0 为后端程序化诊断模型, 在环境配置为 Ubuntu16 系统、Intel®Xeon(R)E5-2603 CPU 处理器和 NVIDIA GeForce GTX 1080Ti GPU 显卡的电脑端运行。通过 DDS 故障诊断综合试验台模拟行星齿轮箱的健康状况, 采集振动信号, 进行试验验证与分析。

### 3.1 试验平台

本文采用的 DDS 故障诊断综合试验台如图 6 所示, 其主要由驱动电机、两级行星齿轮箱、平行齿轮箱、磁粉制动器和控制系统等组成。其中, PCB 608A11 振动传感器安装在行星齿轮箱径向输入端, 采样频率为 5 120 Hz。试验模拟行星齿轮箱工作在

多种转速-负载下的 9 种健康状况,包括 1 类健康状态(HEA),4 类齿轮故障:切齿故障(CTF)、缺齿故障(MTF)、齿根裂纹故障(RCF)和表面磨损故障(SWF),和 4 类滚动轴承故障:滚动体故障(BWF)、复合磨损故障(CWF)、内圈故障(IRF)和外圈故障(ORF),详情如文献[19]所述。

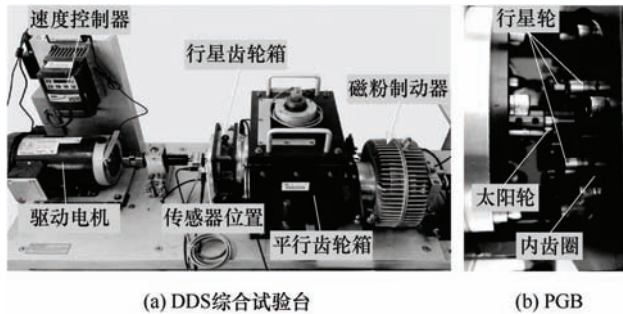


图 6 DDS 故障诊断综合试验台及行星齿轮箱

### 3.2 数据说明

试验采集了行星齿轮箱在四种转速-负载工况下前述 9 种健康状况的振动数据,分别为转速 1 200 r/min-负载 0 N·m,转速 1 800 r/min-负载 3.66 N·m,转速 1 800 r/min-负载 10.98 N·m,和转速 2 400 r/min-负载 0 N·m。每种工况下每类健康状况的振动信号无重叠分割成 2 000 个样本,单个样本长度为 1×2 048,每个工况下均有 18 000 样本。本文所用多工况数据来源于上述四种工况,取每个工况下每类健康状况 500 个样本,共有 18 000 个样本。

### 3.3 结果比较与分析

#### 3.3.1 对比方法

本文将所提方法与多个现有方法进行比较,具体为:1) SVM:利用主成分分析(Principal component analysis, PCA)将振动信号降维至 40 个特征,利用 SVM 进行识别,其采用高斯径向基核函数(RBF),核系数为 0.01;文献[24]、[11]中的深度学习方法,如:2) LetNet-5<sup>[24]</sup>;3) 第一层有宽卷积核的 WDCNN<sup>[11, 24]</sup>;4) 根据文献[19]中定义的模拟环境,设计深度宽卷积 Q 网络(Deep wide convolution Q network, DWCQN)的 DRL 模型,其结构为五个卷积池化交替组合,卷积数依次为 16、16、32、32、64,池化尺寸和步幅均为 2,第一层卷积核宽度为 64,其余层均为 3,最后包含节点 100 的全连接层和输出层;5) MSCNN:其结构同本文网络一致,但训练方式为标准的监督式;6) 单尺度深度注意 Q 网络(Deep attention Q network, DAQN):其结构移除图 3 中多尺度模块中的分支 2、3、4,仅保留分支 1,其余参数设置同所提方法一致;7) MSDAQN':网络结构同本文方法一致,但智能体交互环境依据文献

[20]中建立。其中,LetNet-5、WDCNN 和 DWCQN 优化器均为 Adam,学习率均为 0.001,“批大小”为 128;LetNet-5、WDCNN、MSCNN 训练 epoch 均为 100, DWCQN 最大学习情节数为 2 000。

#### 3.3.2 试验结果比较

为综合评价各方法的诊断性能,采用总体准确率(Overall accuracy, OA)<sup>[19]</sup>作为准确性评价指标,采用 Kappa 系数<sup>[19]</sup>作为诊断一致性评价指标。表 1 列出了各方法在多工况下五折交叉验证的结果比较。

表 1 各方法的五折交叉验证结果比较

方法	平均总体准确率(%)		平均测试 Kappa (%)	平均训练时间/s
	训练	测试		
SVM	53.74±0.32	40.09±0.31	32.60±0.35	49.44
LetNet-5	98.28±3.43	65.84±1.03	61.57±1.16	119.11
WDCNN	97.65±1.24	82.22±4.31	79.99±4.84	289.67
DWCQN	99.13±0.31	97.93±0.34	97.67±0.38	47 673.44
MSCNN	99.99±0.02	98.86±0.52	98.72±0.59	267.82
DAQN	99.75±0.14	99.22±0.43	99.12±0.48	2 283.49
MSDAQN'	99.79±0.15	99.33±0.22	99.25±0.25	2 483.92
MSDAQN	99.97±0.02	99.69±0.14	99.64±0.16	2 560.84

由表 1 可得,所提方法在行星齿轮箱多工况智能诊断上具有较高的准确率,相较于 SVM, LetNet-5 和 WDCNN 方法性能提升显著,分别提升 59.60%, 33.85%, 17.47%,即使比先进的 DWCQN, MSCNN, DAQN, 和 MSDAQN'方法仍提升了 1.76%, 0.83%, 0.47%, 和 0.36%,且具有更好的诊断稳定性和泛化性。对比于 DWCQN 和 MSDAQN',从诊断结果亦能看出所设计的诊断模拟环境更具优势,交互过程更符合人类认知过程,智能体可以学习更好的诊断策略;相较于单尺度 DAQN,表明所提多尺度方法具有更强的特征学习能力和适应性;相较于相同网络结构的 MSCNN,所提 MSDAQN 深度强化学习诊断方法有更好的泛化性。虽然所提方法由于需要同环境交互,平均耗时 2 560.84 s,比监督学习方式耗时长,但仍比 DWCQN 模型更快,亦表明本文建立 CMDP 的优势。同时,模型一旦学习到诊断策略,测试时间相对较短,仅为 0.5 s。

此外,对比各方法的测试 Kappa 系数,所提方法的 Kappa 系数为 99.64%,相较其余 7 种方法分别提升 67.04%, 38.07%, 19.65%, 1.97%, 0.92%, 0.52%, 0.39%,表明该方法具有更好的诊断一致性,对多工况下九种分类别均有较高的准确率,其从图 7 中归一化混淆矩阵亦能看出。

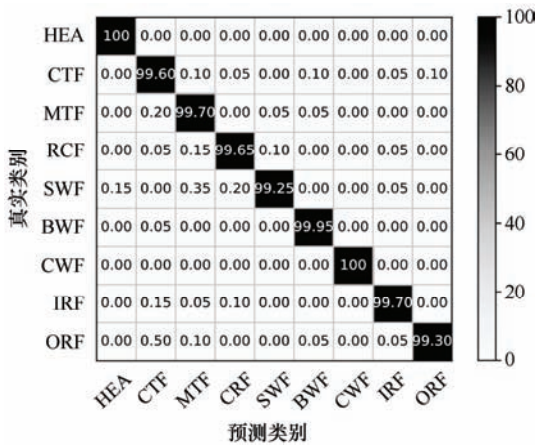


图 7 所提方法五折交叉验证的归一化混淆矩阵

综上，通过上述多工况试验验证了所提方法的有效性和优越性，表明其对多工况下行星齿轮箱智能诊断具有较强的适应性和通用性，相较于 SVM、WDCNN、MSCNN 等，所提 MSDAQN 有更高的诊断准确性和稳定性。

4 应用案例分析

为了进一步评估所提方法在实际应用中的可行性，本节采用某型直升机齿轮箱的振动数据进一步分析模型性能，并同多种方法比较，主要分两个案例，具体如下。

4.1 数据说明

案例 1：直升机齿轮箱包含健康和行星架存在 8.25 cm 裂纹故障两类状态，数据采于试验间，五倍降采样后的采样频率为 20 kHz。鉴于直升机齿轮箱故障植入成本巨大，这里将不同扭矩下上述两类健康状况也看作不同的状态，扭矩分别设置为 96.28 N·m、144.42 N·m、240.69 N·m、336.97 N·m、433.24 N·m 和 481.38 N·m。为此构建两个数据集 A 和 B，其中单工况数据集 A 包含 5 类健康状态 (H1-H5) 和 5 类行星架裂纹故障 (F1-F5)，健康和故

障数据均分别在 96.28 N·m、144.42 N·m、240.69 N·m、336.97 N·m 和 481.38 N·m 扭矩下采集；多工况数据集 B'包含 3 类健康状态 (H1-H3) 和 3 类行星架故障 (F1- F3)，其是在健康和故障两类前提下分别将扭矩为 96.28 N·m 和 144.42 N·m、240.69 N·m 和 336.97 N·m、433.24 N·m 和 481.38 N·m 下采集数据视为一类。数据集 A 和 B 详情见表 2。

案例 2：数据直接采集于直升机运行时，采样频率为 48 kHz。鉴于带“病”工作(行星架裂纹故障)，仅可在低扭矩情况下进行实验，扭矩分别设置 96.28 N·m 和 144.42 N·m 时采集直升机齿轮箱健康和行星架故障数据，依据案例 1 所述，构建两个数据集 C 和 D。其中，数据集 C 将 2 次不同时间段实验也看作不同的状态，共包含 4 类健康状态 (H1-H4) 和 4 类行星架故障 (F1-F4)；数据集 D 将 2 次不同时间段实验看作相同状态，因此其包含 2 类健康状态 (H1-H2) 和 2 类行星架故障 (F1-F2)。数据集 C 和 D 详情亦见表 2。

表 2 直升机齿轮箱数据集说明

案例	工况	类别数	单类样本量	样本长度	样本量
案例 1	A	10	1 750	1×2 048	17 500
	B	6	3 000	1×2 048	18 000
案例 2	C	8	585	1×2 048	4 680
	D	4	1 170	1×2 048	4 680

4.2 结果比较

表 3 列出了多种方法的对比结果，所有方法设置均如节 3 所述。从表 3 可知，所提 MSDAQN 方法具有较高的诊断准确率，仅在工况 A 下略低于 MSCNN 和 DAQN，但其仍高达 99.93%。在工况 A、B、D 下，该方法均有超过 99.85% 的 OA。工况 C 由于将不同时间段监测数据看作不同状态，较其他工况更难诊断，这从各方法的表现亦能看出，但所提方法仍有高达 97.52% 的诊断准确率，相较于 SVM，

表 3 各方法五折交叉验证的测试结果比较

方法	平均总体准确率(%)				平均 Kappa 系数(%)			
	工况 A	工况 B	工况 C	工况 D	工况 A	工况 B	工况 C	工况 D
SVM	91.91±0.57	88.46±0.73	57.54±1.51	79.08±0.31	91.01±0.63	86.15±0.87	51.48±1.73	72.11±0.41
LetNet-5	97.88±0.36	97.33±0.31	62.07±1.28	86.39±1.51	97.64±0.40	96.80±0.38	56.66±1.46	81.85±2.01
WDCNN	98.92±2.03	92.49±8.05	80.39±4.19	98.63±1.74	98.80±2.26	90.99±9.66	77.58±4.78	98.18±2.32
DWCQN	99.65±0.40	99.18±0.51	94.72±0.51	99.38±0.61	99.61±0.44	99.02±0.62	93.97±0.58	99.17±0.82
MSCNN	99.98±0.02	99.84±0.10	94.51±1.66	99.23±1.12	99.98±0.02	99.81±0.13	93.72±1.89	98.97±1.50
DAQN	99.98±0.02	99.83±0.13	95.83±0.66	99.91±0.08	99.98±0.02	99.80±0.16	95.24±0.76	99.89±0.10
MSDAQN'	99.87±0.25	99.76±0.37	97.24±1.27	99.89±0.17	99.86±0.28	99.71±0.44	96.85±1.45	99.86±0.22
MSDAQN	99.93±0.07	99.89±0.08	97.52±0.46	99.98±0.04	99.92±0.08	99.87±0.10	97.17±0.52	99.97±0.06



LetNet-5, WDCNN, DWCQN, MSCNN, DAQN, 和 MSDAQN 分别提升了 39.98%, 35.45%, 17.13%, 2.80%, 3.01%, 1.69%, 和 0.28%。尤其, 对比于 DWCQN 和 MSDAQN 方法, 表明本文所建诊断模拟环境更有优势。此外, 实验中发现虽然该方法在工况 C 下的 OA 仅有 97.52%, 但其对健康和故障状态的识别可达近 100% 的准确率, 且有较好的诊断稳健性, 这从工况 D 中亦可推断出。

此外, 对比各方法的测试 Kappa 系数, 从中可得上述类似结论, 表明所提方法在实际案例中仍具有较好的诊断一致性。如在工况 C 下, 相较其余方法, 该方法的测试 Kappa 系数分别提高了 45.69%, 40.51%, 19.59%, 3.20%, 3.45%, 1.93%, 和 0.32%, 且具有较好的稳定性。综上, 通过上述多个直升机齿轮箱数据集进一步验证了所提方法在实际应用中可行性和潜在价值。

### 4.3 t-SNE 可视化分析

为了直观地分析该方法的特征学习能力, 在上

述直升机齿轮箱数据集下, 将所建诊断模型最后 GAP 层所学特征通过 t-SNE 技术<sup>[25]</sup>降维并可视化, 如图 8 所示。从图 8 可看出, 所提方法在四种情况下均具有较强的特征表达能力, 所学特征具有较大类间距和较小类内距, 特征聚类性较好。同时, 可以发现, 虽然个别类特征存在混叠现象, 但模型对健康状态和故障状态在四种工况下均达到近 100% 的可区分性。如在工况 C 下, 健康状态 H1 和 H4、H2 和 H3 存在局部类内混叠, 故障状态 F1 和 F4、F2 和 F3 存在局部类内混叠, 这也是该方法在工况 C 下具有相对低准确率的原因, 但从图 8c 中可看出健康状态和故障状态是完全可分的。进一步分析发现, 健康状态 H1 和 H4、故障状态 F1 和 F4 均是在扭矩设置为 96.28 N.m 时不同时间段采集的数据, 健康状态 H2 和 H3、故障状态 F2 和 F3 均是在扭矩为 144.42 N.m 下不同时刻采集的振动数据, 因此它们具有较大的类似性, 这也直观地验证了节 4.2 中的分析, 其亦可从工况 D 中表现(图 8 d)进一步得到验证。

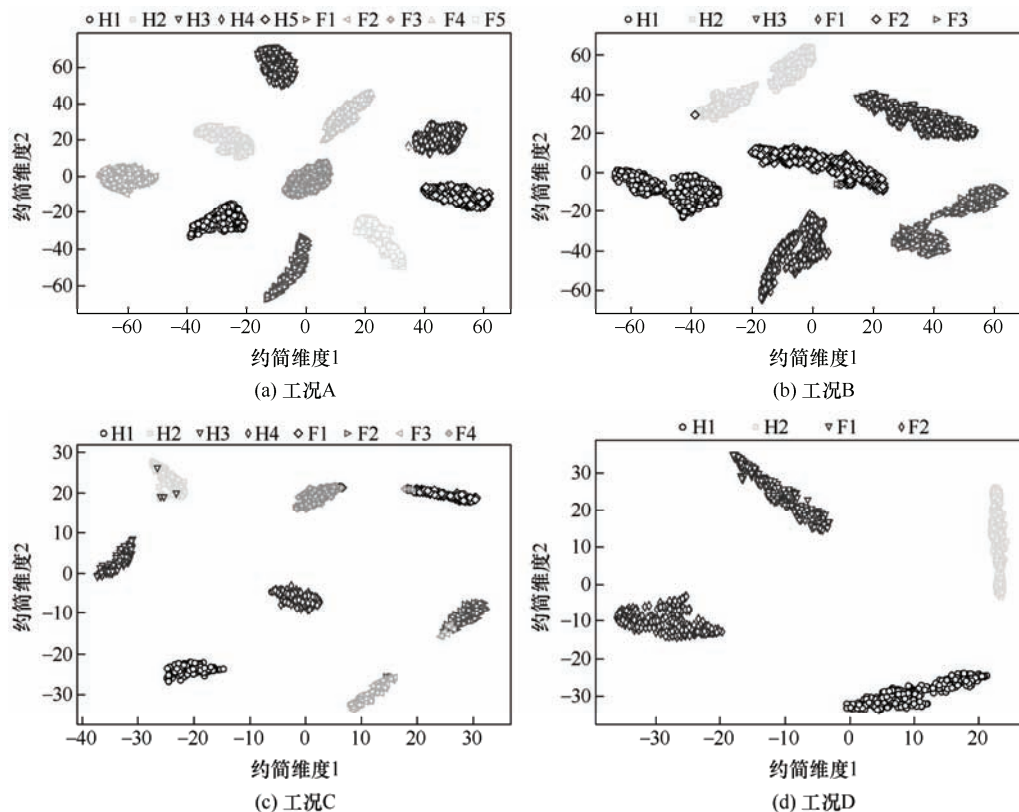


图 8 各工况下所提方法 t-SNE 特征可视化分析

## 5 结论

针对行星齿轮箱的多工况智能故障诊断需求, 本文立足 DRL, 提出一种基于多尺度深度注意 Q 网络(MSDAQN)的故障诊断方法。主要结

论如下。

(1) 该方法利用深度强化学习的自主决策能力和感知优势, 无需人工干预, 实现了行星齿轮箱的端到端智能诊断。

(2) 建立一种诊断环境自主交互规则, 利用奖励机制削弱输入与故障模式间的反馈关系, 从智能



体交互的经验中学习诊断策略, 增强了诊断的通用性。

(3) 提出融合多尺度 CNN 和自适应通道注意力的 Q 网络模型学习复杂故障特征、突出关键信息, 克服单尺度模型感知能力不足, 提升了工况适应性。

(4) 通过多工况试验和实际案例验证了该方法具有较高的诊断准确率, 较传统 SVM, 先进的 WDCNN、DAQN 等性能提升显著, 且诊断稳健性更好。

## 参 考 文 献

- [1] 陈雪峰, 李继猛, 程航, 等. 风力发电机状态监测和故障诊断技术的研究与进展[J]. 机械工程学报, 2011, 47(9): 45-52.
- [2] CHEN Xuefeng, LI Jimeng, CHENG Hang, et al. Research and application of condition monitoring and fault diagnosis technology in wind turbines[J]. Journal of Mechanical Engineering, 2011, 47(9): 45-52.
- [3] ZHAO M, ZHONG S, FU X, et al. Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2020, 68(3): 2587-2597.
- [4] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [5] 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战[J]. 机械工程学报, 2018, 54(5): 94-104.
- [6] LEI Yaguo, JIA Feng, KONG Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. Journal of Mechanical Engineering, 2018, 54(5): 94-104.
- [7] SHAO S Y, SUN W J, YAN R Q, et al. A deep learning approach for fault diagnosis of induction motors in manufacturing[J]. Chinese Journal of Mechanical Engineering, 2017, 30(6): 1347-1356.
- [8] SAUFI S R, AHMAD Z A B, LEONG M S, et al. Gearbox fault diagnosis using a deep learning model with limited data sample[J]. IEEE Transactions on Industrial Informatics, 2020, 16(10): 6263-6271.
- [9] 周兴康, 余建波. 基于深度一维残差卷积自编码网络的齿轮箱故障诊断[J]. 机械工程学报, 2020, 56(7): 96-108.
- [10] ZHOU Xingkang, YU Jianbo. Gearbox fault diagnosis based on one-dimension residual convolutional auto-encoder[J]. Journal of Mechanical Engineering, 2020, 56(7): 96-108.
- [11] JIA F, LEI Y, LIN J, et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data[J]. Mechanical Systems and Signal Processing, 2016, 72: 303-315.
- [12] WANG L H, ZHAO X P, WU J X, et al. Motor fault diagnosis based on short-time Fourier transform and convolutional neural network[J]. Chinese Journal of Mechanical Engineering, 2017, 30(6): 1357-1368.
- [13] ZHANG W, PENG G, LI C, et al. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals[J]. Sensors, 2017, 17(2): 425.
- [14] 叶壮, 余建波. 基于多通道加权卷积神经网络的齿轮箱振动信号特征提取[J]. 机械工程学报, 2021, 57(1): 110-120.
- [15] YE Zhuang, YU Jianbo. Feature extraction of gearbox vibration signals based on multi-channels weighted convolutional neural network[J]. Journal of Mechanical Engineering, 2021, 57(1): 110-120.
- [16] LIU R, WANG F, YANG B, et al. Multiscale kernel based residual convolutional neural network for motor fault diagnosis under nonstationary conditions[J]. IEEE Transactions on Industrial Informatics, 2019, 16(6): 3797-3806.
- [17] 唐斯, 陈新楚, 郑松. 基于注意力与多尺度卷积神经网络的电机轴承故障诊断[J]. 电气技术, 2020, 21(11): 32-38.
- [18] TANG Si, CHEN Xinchu, ZHENG Song. Fault diagnosis method of motor bearing based on attention and multi-scale convolution neural network[J]. Electrical Engineering, 2020, 21(11): 32-38.
- [19] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [20] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484-489.
- [21] 刘朝阳, 穆朝絮, 孙长银. 深度强化学习算法与应用研究现状综述[J]. 智能科学与技术学报, 2020, 2(4): 314-326.

- LIU Zhaoyng, MU Chaoxu, SUN Changyin. An overview on algorithms and applications of deep reinforcement learning[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 314-326.
- [18] DING Y, MA L, MA J, et al. Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach[J]. Advanced Engineering Informatics, 2019, 42: 100977.
- [19] WANG H, XU J, SUN C, et al. Intelligent fault diagnosis for planetary gearbox using time-frequency representation and deep reinforcement learning[J]. IEEE/ASME Transactions on Mechatronics, 2021. doi: 10.1109/TMECH.2021.3076775.
- [20] LIN E, CHEN Q, QI X. Deep reinforcement learning for imbalanced classification[J]. Applied Intelligence, 2020, 50(8): 2488-2502.
- [21] SUTTON R S, BARTO A G. Reinforcement learning: An introduction[M]. MIT Press, 2018.
- [22] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks, 2020 IEEE[C]//CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 2020.
- [23] LIU L, JIANG H, HE P, et al. On the variance of the adaptive learning rate and beyond[J]. arXiv preprint arXiv:1908.03265, 2019.
- [24] WU P, NIE X, XIE G. Multi-sensor signal fusion for compound fault diagnosis method with strong generalization and anti-noise performance[J]. Measurement Science and Technology, 2020, 32(3): 035108.
- [25] LAURENS V D M, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 2605(9): 2579-2605.
- 
- 作者简介: 王辉, 男, 1992 年出生, 博士研究生。主要研究方向为旋转机械的智能诊断。  
E-mail: 230189693@seu.edu.cn
- 严如强(通信作者), 男, 1975 年出生, 博士, 教授, 博士研究生导师。主要研究方向为机械系统状态监测与故障诊断、信号处理、无线传感网络。  
E-mail: ruqiang@seu.edu.cn