

# 基于深度 Q 学习和连续小波变换的旋转机械故障诊断方法

陈仁祥<sup>1</sup>, 周君<sup>1</sup>, 胡小林<sup>2</sup>, 韩兴波<sup>1</sup>, 朱孙科<sup>1</sup>, 张晓<sup>1</sup>

(1. 重庆交通大学交通工程应用机器人重庆市工程实验室, 重庆 400074;

2. 重庆工业大数据创新中心有限公司, 重庆 400056)

**摘要:** 针对旋转机械故障诊断中深度神经网络特征学习能力强、决策能力弱的问题, 利用卷积神经网络拟合强化学习中的 Q 函数, 通过 Q-learning 算法学习策略实现故障诊断, 提出了基于深度 Q 学习和连续小波变换的旋转机械故障诊断方法。对振动信号进行连续小波变换得到时间尺度矩阵, 构建出环境状态空间, 实现智能体与环境间的交互; 用 CNN 拟合 Q-learning 中的 Q 函数得到深度 Q 网络, 将环境返回的状态输入到深度 Q 网络中学习故障数据具体的状态特征表示, 并据此表征学习策略, 智能体采用  $\epsilon$ -贪婪方式决策出动作, 利用奖励发生器对动作进行评价; 通过智能体与环境间不断交互学习以最大化 Q 函数值, 得到最优策略实现故障诊断。这种方式融合了深度学习的感知能力和强化学习的决策能力, 从而有效提高了诊断能力。通过不同工况及不同样本量下齿轮箱故障诊断实验证明了所提方法的有效性。

**关键词:** 故障诊断; 旋转机械; 连续小波变换; 深度 Q 学习

**中图分类号:** TH165.3; TH133 **文献标志码:** A **文章编号:** 1004-4523(2021)05-1092-09

**DOI:** 10.16385/j.cnki.issn.1004-4523.2021.05.024

## 引言

旋转机械故障诊断本质是模式识别, 其中重要的步骤是故障特征提取与故障类型识别<sup>[1]</sup>。近年来, 由于深度神经网络强大的特征学习能力, 使其在旋转机械故障诊断领域得到快速发展<sup>[2-3]</sup>。

利用深度学习进行故障诊断时, 常见的模式是通过建立不同的深度神经网络从训练样本中学习故障特征进而实现故障诊断。如: 胡莛庆等<sup>[4]</sup>对行星齿轮箱故障信号进行经验模态分解, 得到内禀模式函数 (Intrinsic Mode Function, IMF) 后利用深度卷积神经网络融合特征信息明显的 IMF 进行故障诊断。李巍华等<sup>[5]</sup>利用深度置信网络直接从原始信号逐层学习轴承故障特征进行轴承故障诊断。Cao 等<sup>[6]</sup>将长短期记忆网络 (Long Short Term Memory Networks, LSTM) 用于风机齿轮箱故障的模式识别。Chen 等<sup>[7]</sup>对行星齿轮箱信号进行离散小波变换得到二维时频信息, 再输入卷积神经网络 (Convolution Neural Network, CNN) 进行分类识别。

Zeng 等<sup>[8]</sup>对信号进行 S 变换提取时频特征信息, 提出了基于 S 变换和 CNN 的齿轮箱故障诊断。以上方法都取得了不错的效果, 但是大部分神经网络直接在最后一层使用 Softmax 分类器进行分类, 且训练时更注重深度神经网络的特征学习能力, 而忽略了决策能力在神经网络中的作用, 这会影响诊断结果的准确性和可靠性, 特别是在样本量不足时。

深度 Q 学习 (Deep Q Network, DQN) 融合了深度学习的感知能力和强化学习的决策能力, 已在视频游戏中得到广泛应用<sup>[9-11]</sup>。DQN 通过智能体与环境间的互动, 利用深度 Q 网络来提取数据的抽象表征, 同时据此表征进行自我激励的强化学习, 优化解决策略<sup>[12]</sup>。DQN 不仅具有深度学习特征能力强的特点, 还具备强化学习决策能力强的优势, 使其在故障诊断中具有巨大潜力。

同时, 由于旋转机械结构复杂, 工况多变, 信号传递过程中易耦合, 采集到的振动信号往往具有非平稳非线性的特点, 利用时频分析方法获得其二维时频信息更有利于深度 Q 网络学习特征和作出决策。常见的时频分析方法中, S 变换对高频带分辨

**收稿日期:** 2020-03-04; **修订日期:** 2020-06-06

**基金项目:** 国家自然科学基金资助项目 (51975079); 国家重点研发计划 (2018YFB1306601); 重庆市教委科学技术研究项目 (KJQN201900721); 重庆市自然科学基金资助项目 (cstc2016jcyjA0467); 重庆市北碚区科学技术局技术创新与应用示范项目 (2020-6)

率模糊,连续小波变换(Continuous Wavelet Transfer, CWT)具有很强的弱信号检测能力,相对于离散小波和二进制小波具有不易发生漏检和泄露的特点<sup>[13]</sup>,故采用 CWT 对原始信号进行变换获得时间尺度谱。

综上所述,提出基于深度 Q 学习和连续小波变换的旋转机械故障诊断方法。首先利用 CWT 对非线性非平稳信号进行预处理,挖掘数据的二维时间尺度特征信息,以时间尺度矩阵构建出环境状态空间。然后,用 CNN 拟合 Q-learning 中的 Q 函数,将环境返回的状态输入到深度 Q 网络中学习故障数据具体的状态特征表示,并据此表征学习策略。再通过智能体与环境间不断交互学习以最大化 Q 函数值,得到最优策略,从而实现旋转机械故障诊断。通过不同工况和不同样本量下齿轮箱故障诊断实验证明了所提方法的可行性和有效性。

## 1 连续小波变换和深度 Q 学习算法

### 1.1 连续小波变换原理

设函数  $\psi \in L^2(\mathbf{R}) \cap L^1(\mathbf{R})$  并且  $\hat{\psi}(0) = 0$ , 由  $\psi$  经伸缩和平移可以得到一簇函数

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

式中  $a, b \in \mathbf{R}, a \neq 0$ , 称  $\psi_{a,b}$  为连续小波,  $a$  为伸缩因子,  $b$  为平移因子,  $\psi$  为母小波。  $a$  用于改变连续小波的形状,  $b$  用于改变连续小波的位移。

对于任意函数  $f(t) \in L^2(\mathbf{R})$ , 则其连续小波变换定义为

$$W_f(a,b) = \langle f, \psi_{a,b} \rangle = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (2)$$

式中  $\langle f, \psi_{a,b} \rangle$  表示两函数内积,  $\overline{\psi(t)}$  表示  $\psi(t)$  的复共轭。

通过尺度为  $a$ 、平移为  $b$  的子小波和信号内积可以得到每个小波变换系数  $W_f(a,b)$ , 当信号与该子波越相似, 则系数值越大, 特征成分将在时间尺度相平面上某处集结为高幅值的能量块, 反之则能量发散。通过不同的尺度变换, 最终得到系数矩阵  $W_f(a,b)$  能够刻画出信号的二维时间尺度特征。Morlet 小波与旋转机械发生故障时产生的冲击信号形状相似, 故选其为母小波。

### 1.2 深度 Q 网络

强化学习包含四个元素: 智能体、环境状态、动作、奖励。其交互学习过程如图 1 所示。

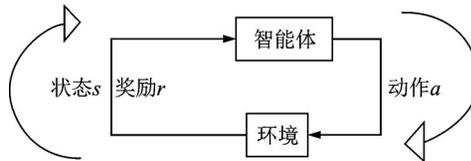


图 1 交互学习示意图

Fig. 1 The diagram of interactive learning

智能体通过不停的交互利用从环境中得到的反馈信号, 调整策略  $\pi$  来最大化长期累积奖励, 定义累积奖励  $G_t$  如下

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (3)$$

式中  $\gamma$  为折扣系数, 通常使用状态-动作值函数来描述每个状态-动作对的好坏, 状态-动作值函数也常被称为 Q 函数, 公式如下

$$q_{\pi}(s,a) = E_{\pi} \left[ \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s, A_t = a \right] \quad (4)$$

强化学习中基于值函数的策略学习算法包括 SARSA<sup>[14]</sup>、Q-learning<sup>[15]</sup> 等, 本文选用 Q-learning 算法, 其迭代公式如下

$$Q(s,a) \leftarrow Q(s,a) + \delta [r + \gamma \max_{a'} Q(s',a') - Q(s,a)] \quad (5)$$

式中  $Q(s,a)$  为 Q 函数;  $s, a, r$  分别为当前状态、动作和奖励;  $Q(s,a)$  取得最大值的  $s', a'$  即为下一个状态和动作。折扣系数  $\gamma$  介于 0-1 之间, 其数值越小表明智能体在学习过程中越在意当前的奖励,  $\delta$  为 Q 学习率, 用于权衡 Q 函数学习的快慢。

CNN 利用多层特征提取层, 自适应提取状态特征形成更加抽象的特征表达, 具有拟合任何函数的特点。因此, 可用 CNN 拟合表示 Q 函数得到深度 Q 网络, 结构如图 2 所示。

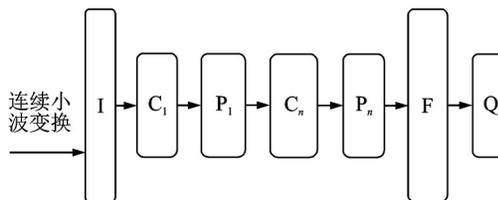


图 2 深度 Q 网络模型结构

Fig. 2 Deep Q network model structure

其中 I 层为输入层,  $C_1$  层为第一层卷积层,  $P_1$  层为对应池化层, 经过多次卷积与池化, 最后将最后一层扩展成全连接层 F, Q 层为输出层, 每一层的具体介绍如下:

1) 卷积层。多个卷积核分别与输入图像进行卷积, 加上对应的偏置值, 一个卷积核对应一个偏置值, 加上偏置值之后通过激活函数就可以得到一系

列特征图,其过程用数学描述如下

$$X_j^l = f\left(\sum_{i \in M_j} X_j^{l-1} * W_{ij}^l + b_j^l\right) \quad (6)$$

式中  $X_j^l$  为  $l$  层第  $j$  个元素,  $M_j$  为  $l-1$  层特征图的第  $j$  个卷积区,  $X_j^{l-1}$  为前一层特征图上第  $j$  个元素,  $W_{ij}^l$  为  $l$  层的权重矩阵,  $b_j^l$  为  $l$  层的偏置矩阵,  $f(\cdot)$  为激活函数,使用 Relu 函数如下

$$f(x) = \max(0, \lg(1 + e^x)) \quad (7)$$

2) 池化层。通过对图像进行降维处理,除去冗余信息,保留数据的重要特征,同时减少了计算量,提高特征学习的效率,公式如下

$$X_j^l = f(\beta_j^l \text{down}(X_j^{l-1}) + b_j^l) \quad (8)$$

式中  $\text{down}(\cdot)$  为下采样函数,  $\beta_j^l$  表示第  $l$  层第  $j$  个特征图的权值,  $b_j^l$  表示第  $l$  层第  $j$  个特征图的偏值,经过池化层  $l$  后的特征图和卷积层  $l-1$  的特征图在数量上保持不变,大小改变。

3) 全连接层。经过多个卷积层和池化层,网络提取到数据的特征,通过全连接层将前一层的特征图连接,最后连接分流层,与全连接层连接前,需要将前一层的特征图展成一维特征向量。

$$y^k = f(\omega^k x^{k-1} + b^k) \quad (9)$$

式中  $y^k$  为全连接层的输出层,  $x^{k-1}$  是展开的一维特征向量,  $f(\cdot)$  为激活函数,  $k$  为网络层数序号,  $\omega^k$  为权重系数,  $b^k$  为偏置值。

4) 输出层。通过前面的卷积、池化等操作,卷积神经网络在输出层得到故障数据状态下每类动作  $Q$  值,即

$$Q(s, a; \theta) \approx Q(s, a) \quad (10)$$

式中  $Q(s, a; \theta)$  为卷积神经网络的  $Q$  值,  $\theta$  为卷积神经网络的权重参数。

### 1.3 深度 Q 学习算法

通过卷积神经网络拟合  $Q$  函数,同时结合 Q-learning 算法得到的深度 Q 学习算法如表 1 所示,其中  $Q$  函数的迭代公式如下

$$y_i = Q(s, a; \theta_i) + \delta(g(s, a, s') + \gamma \max_a Q(s', a'; \theta_i^-) - Q(s, a; \theta_i)) \quad (11)$$

式中  $y_i$  为第  $i$  次迭代目标值的优化结果;  $g(s, a, s')$  为奖励发生器,  $\theta_i, \theta_i^-$  为深度 Q 网络、目标网络的权重参数。目标网络与深度 Q 网络结构相同,每迭代  $C$  次将 DQN 的参数  $\theta_i$  复制给目标网络的参数  $\theta_i^-$ 。

$\epsilon$ -贪婪策略和奖励发生器  $g(s, a, s')$  公式如下

$$a_i = \begin{cases} a, & \text{if random number} < \epsilon \\ \max_a Q(s, a; \theta), & \text{other} \end{cases} \quad (12)$$

式中  $0 < \epsilon < 1$ , 用于权衡智能体在学习过程中对环境进行探索和利用的程度。

表 1 深度 Q 学习算法

Tab. 1 The learning algorithm of DQN

输入数据状态空间 $S$ , 动作空间 $A$ , 折扣率 $\gamma$ , 学习率 $\eta$ , 贪心率 $\epsilon$
初始化经验池 $D$ , 容量为 $N$
初始化 Q 网络的参数 $\theta$
初始化目标 Q 网络的参数 $\theta^-$
初始化状态 $s$
设置训练轮数 $epi$
For $episode=1, epi$
在状态 $s$ , 采用 $\epsilon$ 贪婪策略得到动作 $a_i$
智能体执行动作 $a$ , 数据环境返回奖励值 $r$ 和对应故障类型的下一个数据状态 $s'$ 。在状态 $s'$ 采用 $\epsilon$ 贪婪策略得到动作 $a_i'$
将 $e_i=(s, a, r, s')$ 存储在 $D$ 中
从 $D$ 中随机选取 $k$ 个经验样本 $(s_j, a_j, r_j, s_j')$ , $j=1, 2, \dots, k$ , $k < N$
$y_i = \begin{cases} r_j, & \text{for terminal} \\ Q(s, a; \theta_i) + \delta(g(s, a, s') + \gamma Q(s', a'; \theta_i^-) - Q(s, a; \theta_i)), & \text{for other} \end{cases}$
对损失函数 $L(\theta)$ 执行梯度下降算法, 更新神经网络参数 $\theta$ 。
每 $C$ 次更新目标网络参数 $\theta^- = \theta$
End
Until $Q(s, a; \theta_i)$ 收敛

$$g(s, a, s') = \begin{cases} r, & \text{right} \\ 0, & \text{wrong} \end{cases} \quad (13)$$

图 3 为智能体决策过程中的移动示意图, 每层具有  $Z$  个节点, 每个节点代表故障种类。每一次迭代, 智能体根据环境返回的数据状态使用公式 (12) 做出某个动作, 当实际决策做出的动作所对应的节点与标记故障类型一致时, 设置奖励值  $r$ 。否则奖励值为 0。

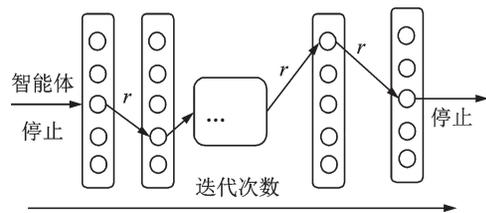


图 3 智能体移动示意图

Fig. 3 The schematic diagram of agent movement

通过智能体与环境交互作用产生大量经验样本, 然后随机抽取输入到深度 Q 网络和目标网络产生误差, 采用随机梯度下降方法调节深度 Q 网络的权重和偏值, 使误差达到最小值, 本文使用的损失函数为均方误差函数

$$L(\theta_i) = E_{s, a, r, s'} [(y_i - Q(s, a; \theta_i))^2] \quad (14)$$

对式 (12) 求一阶偏导数, 可以将得到的误差逐

层反向传递,进而更新卷积神经网络的参数:

$$\omega_{kj}^l = \omega_{kj}^l - \eta \frac{\partial L}{\partial \omega_{kj}^l} \quad (15)$$

$$b_j^l = b_j^l - \eta \frac{\partial L}{\partial b_j^l} \quad (16)$$

左侧为更新后的  $\omega_{kj}^l$  和  $b_j^l$ ,  $\eta$  为学习率,依次类推,逐层更新由  $l$  层到  $l-1$  层的权重和偏置值。

## 2 基于深度 Q 学习和连续小波变换的旋转机械故障诊断流程

所提故障诊断方法流程图如图 4 所示。智能体由深度 Q 网络、目标网络、以及回放记忆单元等构成,智能体通过卷积神经网络可以感知数据状态的抽象表征,同时根据环境反馈奖励值进行策略学习。

主要步骤如下:

### 1) 采集数据与划分

合理布置加速度传感器的位置,采集到的时域数据划分方式如下:确定数据状态(即样本)个数  $N_s = l_t / l_s$ ,其中  $l_t$  为时域信号长度,  $l_s$  为样本长度。由此,可得到某个时域数据状态  $d_k^{td}$ ,下标  $k$  表示状态序列,上标 td 表示时域数据。

### 2) 环境状态空间的构建

将故障类型映射到动作空间集  $A = \{a_1=0, a_2=1, \dots, a_i=n-1\}$ ,  $n$  为故障类型数。可得到某故障类型的数据状态空间集

$$S_{a_i} = \{s_1, s_2, \dots, s_k\}, s_k = d_k^{cwt} \quad (17)$$

式中  $i < n, k \leq N_s, s_k$  为当前故障数据状态空间中的某个状态,包含对时域数据  $d_k^{td}$  进行 CWT 变换得到时间尺度特征  $d_k^{cwt}$ ,动作空间集  $A$  与数据状态空间集  $S$  的对应关系如下

$$S = \{S_{a_1}; A = a_1, S_{a_2}; A = a_2, \dots, S_{a_i}; A = a_i\} \quad (18)$$

### 3) 训练深度 Q 网络

智能体根据某故障类型的数据状态  $s_k$  决策出动作  $a_i$ ,即决策出故障类型;然后环境根据动作  $a_i$  利用公式(18)返回对应故障类型的下一数据状态空间中某个状态  $s_k'$ ,同时根据公式(13)对当前的决策结果进行评价,给予奖励值  $r$ 。回放记忆单元将每一次的交互数据  $(s_k, a_i, r, s_k')$  存储在回放记忆单元  $D$  中,然后随机抽取  $K$  个样本输入到深度 Q 网络和目标网络,采用随机梯度下降法来减少公式(14)中的误差值,同时更新深度 Q 网络的权重和偏置值参数以最大化 Q 函数值。

### 4) 故障诊断结果输出

将测试样本输入到深度 Q 网络,输出诊断结果。

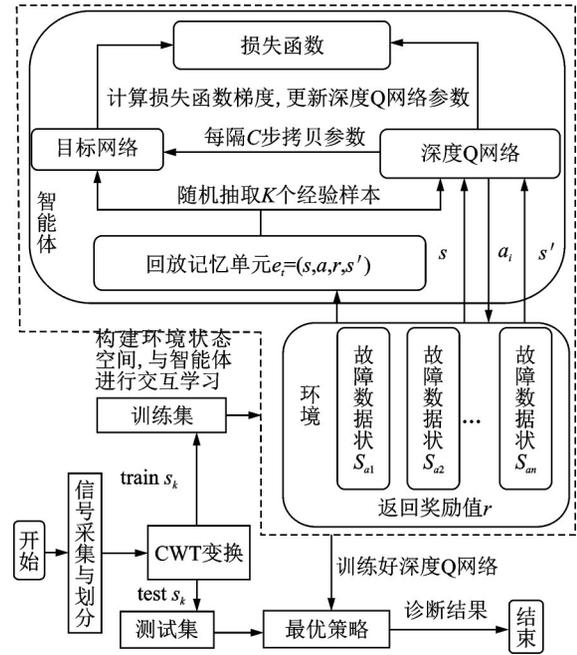


图 4 方法流程图

Fig. 4 The flow chart of method

## 3 实验验证

### 3.1 实验数据描述

通过动力传动故障诊断综合实验台<sup>[16]</sup>进行实验,该齿轮箱中齿轮和轴承均存在故障,实验台如图 5 所示,故障类型及描述如表 2 所示。

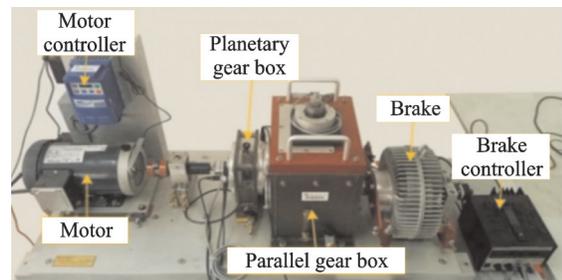


图 5 实验台

Fig. 5 The experimental set up

表 2 轴承和齿轮箱故障类型描述

Tab. 2 Bearing and gearbox fault types description

位置	类型	描述
齿轮箱	断齿	齿轮存在断齿故障
	缺齿	缺一个齿
	齿根	齿根存在裂纹
	齿面	齿轮表面磨损
轴承	滚动体	滚动体存在裂纹
	内圈	内圈出现裂纹
	外圈	外圈出现裂纹
	复合故障	内圈和外圈同时出现裂纹

表2中,齿轮故障类型包括断齿、缺齿、齿根裂纹、齿面磨损及正常共计5种。轴承故障包括滚动

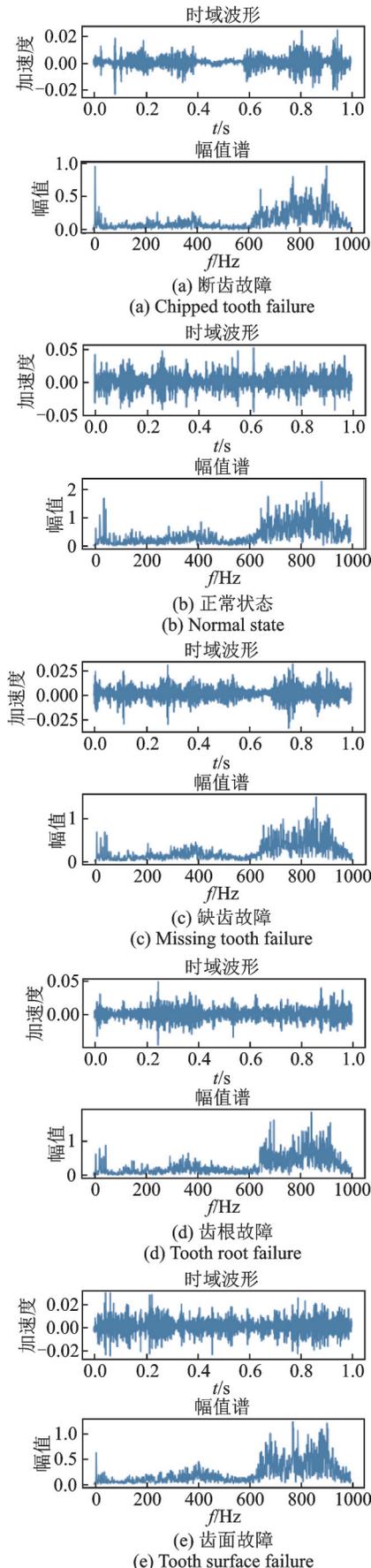


图6 时域波形图和幅值谱

Fig. 6 Time domain waveform and amplitude spectrum

体、内圈、外圈、复合故障及正常5种。实验中在两种工况下采集到振动加速度信号,两种工况分别为:1. 电机转频 20 Hz, 负载电压 0 V; 2. 电机转频 30 Hz, 负载电压 2 V。由于篇幅有限,具体采集过程可参考文献[17]的数据采集实验部分。

图6为齿轮箱中不同故障齿轮的时域波形图和幅值谱,不同故障的特征并不明显,如正常状态和齿面故障的波形极为相似,人工诊断极易出错。图7为对应的时间尺度图,可观察到不同故障类型的能量分布存在显著差异,如图中断齿故障和正常状态的能量分布具有差异。因此,信号通过CWT处理更能表达出不同故障之间的特征信息。

### 3.2 DQN 参数确定

首先进行齿轮故障诊断实验。根据参考文献[7]和经过多次实验,所使用的网络结构如下:两层卷积层,每层卷积层卷积核的个数为  $C_1=32, C_2=32, C_3=64$ ,卷积核大小均为  $5 \times 5$ ,且每层卷积层后连接  $2 \times 2$  池化层,池化方式采用平均池化方式;第一层全连接层神经元个数为  $F_1$ ,第二层隐藏层神经元个数为 1024;输出层神经元个数为 5,即故障类型个数。每个工况下的每类故障样本量为 150,训练集:测试集=7:3。

对于DQN参数设置,目前主要通过经验和采用控制变量法进行调参后确定。因篇幅有限,以Q函数学习率  $\delta$  和神经网络学习率  $\eta$  的调参过程为例进行说明。表3为不同  $\delta, \eta$  数值下得出的诊断结果。观察表3,不同参数下智能体学习到的策略存在一定差异,从而得到不同的诊断精度,当  $\eta=0.06, \delta=1$  时,在工况 20 Hz-0 V, 30 Hz-2 V 时分别为 99.76%, 94.51%, 同时达到最大值。因此选定参数  $\eta=0.06, \delta=1$ ,其他参数的设定过程类似。通过不断调参分析后参数设置如表4所示。

### 3.3 实验验证及分析

为验证所提方法的优越性,与文献[8]S变换+CNN方法、文献[18]STFT变换+CNN方法、CNN和CWT+CNN诊断方法进行了对比。

表5为对应的卷积神经网络结构参数设置情况,其中卷积层C包含卷积核个数和卷积核大小两个参数,池化层P包含层数和池化大小两个参数,全连接层F只有神经元数目一个参数,所有输入层I和输出层O的参数设置一样,输入维度为  $34 \times 34 \times 1$ ,输出维度大小为故障类型个数。

图8(a)显示了在训练过程中智能体获得奖励

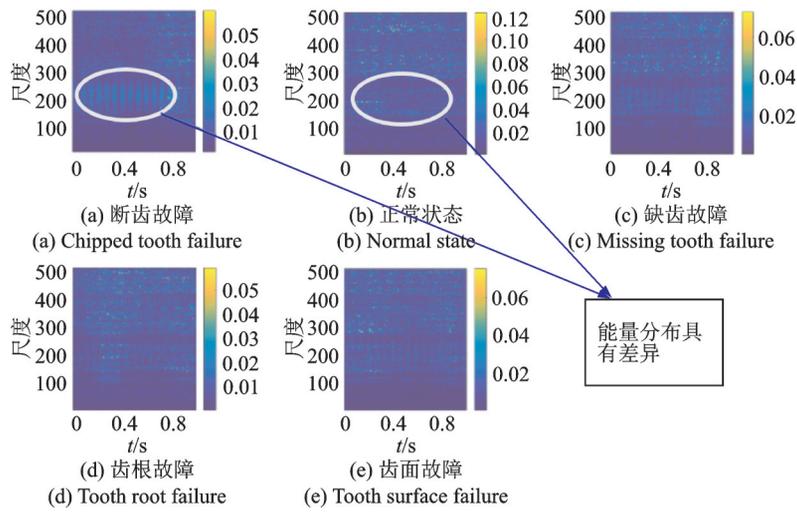


图7 时间尺度图

Fig. 7 The time scale graph

表 3 不同  $\delta$  值和  $\eta$  值下的诊断正确率

Tab. 3 Diagnostic accuracy under different  $\delta$  and  $\eta$

	$\delta=0.1$	$\delta=1$
$\eta=0.06$	20 Hz-0 V:99.65%	20 Hz-0 V:99.76%
	30 Hz-2 V:92.18%	30 Hz-2 V:94.51%
$\eta=0.1$	20 Hz-0 V:99.10%	20 Hz-0 V:99.47%
	30 Hz-2 V:93.80%	30 Hz-2 V:91.74%

表 4 参数设置

Tab. 4 Parameter setting

$\delta$	$\gamma$	$\eta$	$\epsilon$	$epi$	$K$	$D$
1	0.001	0.06	0.7	$1.5 \times 10^3$	32	$3 \times 10^{10}$

表 5 网络结构参数设置

Tab. 5 Network structure parameter setting

方法	网络结构
CNN	I-C1(32,5×5)-P1(32,2×2)- C2(32,5×5)-P2(32,2×2)- C3(64,5×5)-F(1024)-O
ST+CNN	I-C1(6,5×5)-P1(6,2×2)- C2(12,3×3)-P2(12,2×2)-F(100)-O
STFT+CNN	I-C1(32,4×4)-P1(32,2×2)- C2(64,4×4)-P2(64,2×2)-F(100)-O
CWT+CNN	网络结构同 CNN 一样
所提方法	深度 Q 网络结构同 CNN 一样

值的情况,可以看到随着智能体不断学习,决策能力逐渐提高,从而得到更多的奖励值,最后达到平稳状态。图 8(b)显示了在训练过程中准确率的变化,由于决策能力逐渐提高,准确率逐渐提升,最终在 20 Hz-0 V 时最大达到 99.76%, 30 Hz-2 V 时达到 94.51%。

图 9 显示了工况为 20 Hz-0 V 时,训练过程中均方误差值和 Q 函数平均值的变化情况。图 9(a)显示了均方误差代价函数的变换情况,随着迭代次数

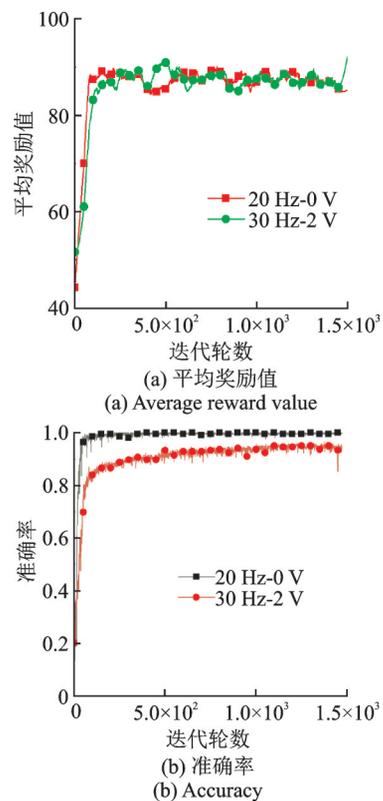


图 8 平均奖励值和准确率变化曲线

Fig. 8 The change curve of average reward value and accuracy rate

的增加,误差越来越小,最后趋于平缓。图 9(b)显示了 Q 函数值的变化情况:(1)在初始阶段,由于网络参数还不稳定,决策能力弱,Q 函数值波动较大。(2)在最后阶段,Q 函数曲线趋于平稳,表明深度 Q 网络具有收敛性,最终学习到最优策略,决策出故障类别。

表 6 为不同方法下齿轮箱的故障的识别结果。根据该表可看出:1)通过时频分析方法预处理原始数据后的准确率远高于直接利用 CNN 对原始数据

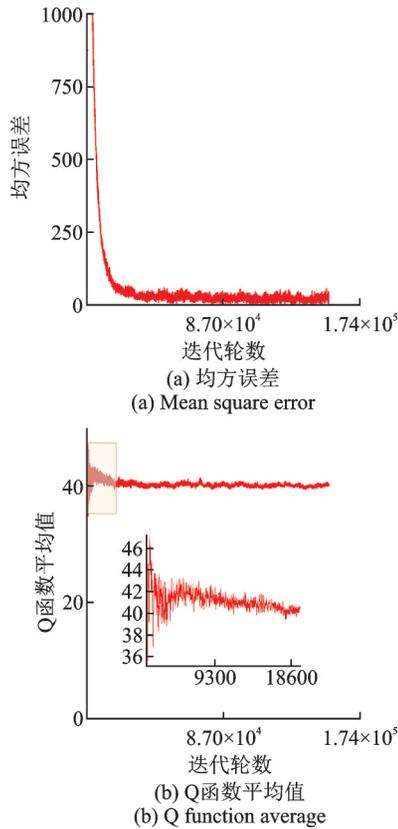


图9 均方误差和Q函数平均值

Fig. 9 Mean square error and Q function mean

进行学习的准确率;2)因ST, STFT对振动信号的时频特征展现不佳,使ST+CNN, STFT+CNN的准确率均低于CWT+CNN;3)CWT+CNN的准确率低于所提方法,但只是略低(平均准确率低2.21%),说明充分发挥CNN的特征学习能力可以获得较好的准确率。同时,也证明所提方法融合了深度学习特征学习能力与决策能力,可取得更好的结果,但优势并不突出。故后文分析不同样本量时各方法的表现,以进一步证明所提方法的优势。

为验证在不同样本量下所提方法的优势,将两种工况的每类故障样本量设为如表7所示,每种工况按比例7:3,5:5得到训练集、测试集,混合两种工况的训练集进行训练。诊断准确率与样本量的关系曲线如图10所示,根据该图,在样本量较少时,所提方法准确率明显高于其他方法,且不同样本量下准确率的波动小,验证了在混合两种不同工况样本下进行训练诊断时,结合特征提取能力和决策能力的深度Q网络具有更好的诊断优势。也说明,训练样本量不足会限制以CNN为代表的深度神经网络的特征学习能力,导致结果不佳。

3.4 齿轮和轴承综合诊断实验验证

为了进一步验证所提方法的优越性,将齿轮箱

表6 齿轮故障诊断结果对比

Tab. 6 Comparison of gear fault diagnosis results

方法	20 Hz-0 V	30 Hz-2 V	平均
CNN	84.80%	48.08%	66.44%
ST+CNN	90.67%	79.11%	84.89%
STFT+CNN	95.14%	88.78%	91.96%
CWT+CNN	98.81%	91.05%	94.93%
所提方法	99.76%	94.51%	97.14%

表7 两种工况样本量

Tab. 7 Sample size of two working conditions

样本量序号	20 Hz-2 V	30 Hz-2 V	共计
0	40	10	50
1	45	15	60
2	50	20	70
3	55	25	80
4	60	30	90
5	65	35	100
6	70	40	110
7	75	45	120
8	80	50	130
9	85	55	140

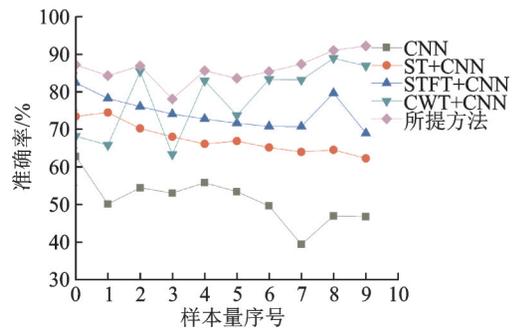


图10 不同样本量下齿轮故障诊断的准确率

Fig. 10 The accuracy of gear fault diagnosis under different sample size

中的齿轮故障与轴承故障综合进行实验验证。即故障类型包括:断齿、缺齿、齿根裂纹、齿面磨、滚动体、内圈、外圈、复合故障(内外圈均有)以及正常共计9类。使用所提方法进行诊断实验。同时CNN, ST+CNN, STFT+CNN, CWT+CNN模型均与上一节相同,实验结果如表8所示,图11为不同样本量下的实验结果。

观察表8和图11,在包含齿轮和轴承故障类型时,所提方法与前文的表现类似。在样本量充分时,所提方法表现最好,但优势不突出。而在样本量不足时,仅仅依靠深度学习的特征感知能力并不能得到满意的结果,而融合了深度学习的特征感知与强化学习的决策能力的所提方法可以获得满意结果,

进一步证明了所提方法的优势。

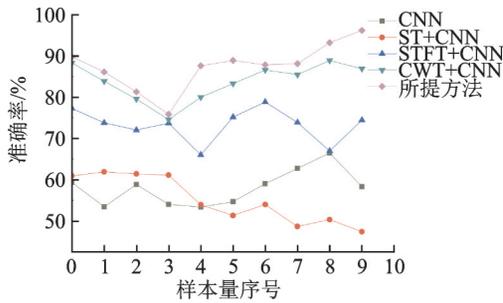


图 11 不同样本量下齿轮和轴承故障诊断的准确率  
Fig. 11 Accuracy of different numbers of gearbox and bearing failure samples

表 8 齿轮和轴承故障诊断

Tab. 8 Gear and bearing fault diagnosis

方法	20 Hz-0 V	30 Hz-2 V	平均
CNN	73.90%	53.81%	63.86%
ST+CNN	83.70%	80.25%	81.98%
STFT+CNN	86.44%	92.10%	89.27%
CWT+CNN	98.10%	91.78%	94.94%
所提方法	99.91%	96.79%	98.35%

从图 10 和 11 可以看到 CNN, ST+CNN, STFT+CNN 随着样本序号的增加准确率出现降低的情况,这是因为训练样本包含了两种工况,且两种工况的样本量配置不平衡,随着样本量的增加,由于不同的诊断模型具有不同的性能,因此有可能存在随着样本量的增加识别精度下降的情况。

在训练过程中,由于深度 Q 学习与 CWT+CNN 中卷积神经网络的结构参数一样,所以两者的网络模型的复杂度一样。但由于深度 Q 学习包含了两个网络模型(深度 Q 网络和目标网络),训练时耗时会增加。表 9 为两种方法的时间对比,可发现所提方法的训练时间明显大于 CWT+CNN,但是实际诊断过程中是使用训练好的模型进行诊断,所需的诊断时间非常少(不超过 1 s)。其中电脑配置如下:操作系统 WIN10;处理器为 inter(R) Core (TM) i5-4210M CPU@2.6 GHz;内存容量 8 GB;显卡为 NVIDIA GeForce GTX850M。

表 9 时间对比

Tab. 9 Time comparison

方法	训练时间/s		诊断时间/s	
	20 Hz-0 V	30 Hz-2 V	20 Hz-0 V	30 Hz-2 V
CWT+CNN	94	588	0.69	0.69
所提方法	1689	2379	0.65	0.65

## 4 结 论

本文针对故障诊断中深度神经网络感知能力强、决策能力弱的问题,提出了基于深度 Q 学习和连续小波变换的旋转机械故障诊断方法。利用连续小波变换挖掘信号的二维时间尺度特征信息,将时间尺度矩阵作为智能体的当前状态,输入到深度 Q 网络中提取状态特征;同时通过 Q-learning 算法进行策略学习,然后环境对智能体做出的动作进行评价。通过智能体与环境之间不停的交互“试错”,更新策略以提升决策能力,最终得到最优策略实现故障诊断。这种方法融合了深度学习的特征学习与强化学习的决策能力,在样本量有限时优势突出,进一步发展了基于深度学习的故障诊断理论与方法。

### 参考文献:

[1] 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.  
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.

[2] Lei Y, Jia F, Lin J, et al. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data[J]. IEEE Transactions on Industrial Electronics, 2016, 63(5): 3137-3147.

[3] Zhao Rui, Yan Ruqiang, Chen Zhenghua, et al. Deep learning and its applications to machine health monitoring[J]. Mechanical Systems and Signal Processing, 2019, 115: 213-237.

[4] 胡葛庆,陈徽鹏,程哲,等.基于经验模态分解和深度卷积神经网络的齿轮箱故障诊断方法[J].机械工程学报,2019,55(7):9-18.  
HU Niaoqing, CHEN Huipeng, CHENG Zhe, et al. Fault diagnosis for gearbox based on emd and deep convolutional neural networks[J]. Journal of Mechanical Engineering, 2019, 55(7): 9-18.

[5] 李巍华,单外平,曾雪琼.基于深度信念网络的轴承故障分类识别[J].振动工程学报,2016,29(2):340-347.  
LI Weihua, SHAN Waiping, ZENG Xueqiong. Bearing fault classification and recognition based on deep belief network[J]. Journal of Vibration Engineering, 2016, 29(2): 340-347.

- [6] Cao L, Zhang J, Wang J Y, et al. Intelligent fault diagnosis of wind turbine gearbox based on long short-term memory networks [C]. 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE). IEEE, 2019: 890-895.
- [7] Chen R, Huang X, Yang L, et al. Intelligent fault diagnosis method of gearboxes based on convolution neural network and discrete wavelet transform [J]. Computers in Industry, 2019, 106: 48-59.
- [8] Zeng X, Liao Y, Li W. Gearbox fault classification using S-transform and convolutional neural network [C]. International Conference on Sensing Technology. IEEE, 2016.
- [9] Lake B M, Salakhutdinov R, Tenenbaum J B. Human-level concept learning through probabilistic program induction [J]. Science, 2015, 350(6266): 1332-1338.
- [10] Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning [J]. 2013, arXiv: 1312.5602.
- [11] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518(7540): 529.
- [12] 刘全, 翟建伟, 章宗长, 等. 深度强化学习综述 [J]. 计算机学报, 2018, 41(1): 1-27.  
LIU Quan, ZHAI Jianwei, ZHANG Zong-zhang, et al. A survey on deep reinforcement learning [J]. Chinese Journal of Computers, 2018, 41(1): 1-27.
- [13] 林京, 屈梁生. 基于连续小波变换的信号检测技术与故障诊断 [J]. 机械工程学报, 2000, 36(12): 95-100.  
LIN Jing, QU Liangsheng. Feature detection and fault diagnosis based on continuous wavelet transform [J]. Journal of Mechanical Engineering, 2000, 36(12): 95-100.
- [14] Singh S, Jaakkola T, Littman M L, et al. Convergence results for single-step on-policy reinforcement-learning algorithms [J]. Machine Learning, 2000, 38(3): 287-308.
- [15] Watkins C J C H. Learning from delayed rewards [D]. Cambridge: University of Cambridge, 1989.
- [16] Shao Siyu. Mechanical-dataset [DS]. <https://github.com/cathysiyu/Mechanical-datasets>, 2018.1.17.
- [17] Shao Siyu, McAleer S, Yan Ruqiang, et al. Highly-accurate machine fault diagnosis using deep transfer learning [J]. IEEE Transactions on Industrial Informatics, 2019, 15(4): 2446-2455.
- [18] 李恒, 张氢, 秦仙蓉, 等. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法 [J]. 振动与冲击, 2018, 37(19): 132-139.  
LI Heng, ZHANG Qing, QIN Xianrong, et al. Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network [J]. Journal of Vibration and Shock, 2018, 37(19): 132-139.

## Fault diagnosis method of rotating machinery based on deep Q-learning and continuous wavelet transform

CHEN Ren-xiang<sup>1</sup>, ZHOU Jun<sup>1</sup>, HU Xiao-Lin<sup>2</sup>, HAN Xing-bo<sup>1</sup>, ZHU Sun-ke<sup>1</sup>, ZHANG Xiao<sup>1</sup>

(1.Chongqing Engineering Laboratory for Transportation Engineering Application Robot, Chongqing Jiaotong University, Chongqing 400074, China; 2.Chongqing Innovation Center of Industrial Big-Data Co. Ltd., Chongqing 400056, China)

**Abstract:** To solve the problems of strong neural network feature learning and weak decision-making ability in fault diagnosis of rotating machinery, a convolutional neural network (CNN) is used to fit the Q function in reinforcement learning, and the learning strategy is implemented by the Q-learning algorithm. For fault diagnosis, a fault diagnosis method for rotating machinery based on deep Q-learning and continuous wavelet transform is proposed. A continuous wavelet transform is performed on the vibration signal to obtain a time-scale matrix, and an environmental state space is constructed for the interaction between the agent and the environment. CNN is used to fit the Q function in Q-learning to obtain a deep Q network to convert the environment. The returned state is input to the deep Q network to learn the specific state feature representation of the fault data, and the learning strategy is characterized accordingly. The agent uses  $\epsilon$ -greedy mode to decide the action and reward generator to evaluate the action. The agent continuously interacts with the environment to maximize the Q function value and obtain the optimal strategy for fault diagnosis. This method combines the perceptual ability of deep learning and the decision-making ability of reinforcement learning, so as to effectively improve the diagnostic ability. The effectiveness of the proposed method is proved by the fault diagnosis experiments of gearbox under different working conditions and different sample sizes.

**Key words:** fault diagnosis; rotating machinery; continuous wavelet transform; deep Q-learning

作者简介: 陈仁祥(1983-),男,博士,教授,硕士生导师。电话:(023)62539903; E-mail: manlou.yue@126.com

通讯作者: 胡小林(1984-),男,硕士,讲师。电话:13594019266; E-mail: huxl0918@163.com