

DOI: 10.19650/j.cnki.cjsi.J2108164

基于深度宽卷积 Q 网络的行星齿轮箱故障智能诊断方法*

王 辉¹, 徐佳文¹, 严如强^{1,2}

(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:针对行星齿轮箱故障诊断常依赖较强的专业知识,诊断模型通用性差的问题,基于深度强化学习,提出一种深度宽卷积 Q 网络的行星齿轮箱故障智能诊断方法。首先将行星齿轮箱的故障诊断分解为序贯决策问题,采用分类马尔科夫决策过程进行描述,并建立故障诊断模拟环境;其次设计深度宽卷积神经网络作为深度 Q 网络模型中的动作值网络,增强对环境状态的感知能力;最后模型通过与环境间的不断交互,并依据环境反馈的奖励,自主学习最佳诊断策略,从而完成行星齿轮箱的状态辨识。试验及案例结果表明:该方法能够在多个工况下均可有效、准确地实现行星齿轮箱的智能诊断,诊断准确率均超过 99%,增强了诊断模型的泛化性和通用性。

关键词:行星齿轮箱;故障诊断;马尔科夫决策过程;卷积神经网络;深度 Q 网络

中图分类号: TH165+.3 TH17 **文献标识码:** A **国家标准学科分类代码:** 510.40

Intelligent fault diagnosis for the planetary gearbox based on the deep wide convolution Q network

Wang Hui¹, Xu Jiawen¹, Yan Ruqiang^{1,2}

(1. School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China;

2. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The fault diagnosis of the planetary gearbox often relies on strong professional knowledge, and the universality of the diagnosis model is poor. Based on deep reinforcement learning, an intelligent fault diagnosis method of the planetary gearbox using the deep wide convolution Q network is proposed. Firstly, fault diagnosis of the planetary gearbox is resolved into a sequential decision problem, which is described by the classification Markov decision process. The fault diagnosis simulation environment is established. Secondly, a deep wide convolutional neural network is designed as an action-value network in the deep Q network model to enhance the perception ability of the environmental state. Finally, the model learns the best diagnostic policy autonomously by interacting with the environment and according to the reward of the environment. In this way, the state identification of the planetary gearbox can be achieved. Experiment and case results show that this method can effectively and accurately realize the intelligent diagnosis of the planetary gearbox under multiple working conditions. The diagnostic accuracy is more than 99%, which enhances the generalization and universality of the diagnosis model.

Keywords: planetary gearbox; fault diagnosis; Markov decision process; convolutional neural network; deep Q network

0 引 言

行星齿轮箱作为机械传动系统的重要部件,具有质量轻、体积小、效率高、传动比大以及载荷强等优点,已广泛应用于风机、直升机、船舶等重大装备中^[1-2]。然而,复

杂的工作环境常使其齿轮、轴承、齿圈等关键部件发生磨损或疲劳裂纹故障,影响装备的安全性^[3]。因此,研究行星齿轮箱故障诊断方法对保障机械装备的安全可靠运行,降低维护成本具有重要意义。

近年来,随着人工智能的不断兴起,基于数据驱动的智能诊断成为了研究热点^[1,4]。传统智能诊断方法多利

收稿日期:2021-06-28 Received Date: 2021-06-28

* 基金项目:国家重点研发计划(2018YFB1702400)项目资助

用时、频域特征统计^[5]、快速傅里叶变换^[6]、小波变换^[7]、经验模态分解(empirical mode decomposition, EMD)^[4-5]等提取特征,再结合支持向量机(support vector machine, SVM)^[3,8]、多层感知器(multilayer perceptron, MLP)^[9]、极限学习机^[10]等实现故障分类。它们虽取得不错的效果,但特征提取需要人工干预且严重依赖专家知识和经验,所提指标往往仅适应特定的任务,且浅层模型难以挖掘抽象的特征映射,影响诊断的通用性。尤其机械健康状态监测已进入“大数据”时代^[11],上述传统方法更难以满足机械设备的智能诊断需求。

目前,深度学习(deep learning, DL)^[12]由于在自动特征提取和识别方面的独特优势,已成为智能诊断中炙手可热的“利器”^[11,13]。赵光权等^[14]、李巍华等^[15]提出基于深度置信网络(deep belief network, DBN)的滚动轴承故障特征提取及诊断,增强了诊断的通用性和适应性;Sun等^[16]提出一种基于稀疏自编码的深度神经网络(deep neural network, DNN)实现感应电机故障诊断,取得较好的效果。卷积神经网络(convolutional neural network, CNN)作为深度学习的重要分支,在模式识别等领域获得巨大的成功,逐渐被应用于故障诊断领域^[17]。Zhao等^[18]将动态加权小波系数与深度残差卷积网络(deep residual convolutional network, DRCN)相结合实现行星齿轮箱故障诊断;胡芑庆等^[4]提出一种基于EMD和深度CNN的行星齿轮箱故障诊断方法;Chen等^[19]将时、频域统计特征作为CNN的输入实现齿轮箱智能诊断;陈仁祥等^[20]提出一种基于离散小波变换与CNN的滚动轴承智能诊断方法;曲建岭等^[21]提出一种基于自适应一维(one-dimensional, 1D)CNN的“端到端”滚动轴承智能诊断方法;Zhang等^[22]提出第一层具有宽卷积核的深度CNN(deep CNN with wide first-layer kernels, WDCNN)模型用于滚动轴承的智能诊断。以上方法在智能诊断中取得显著的效果,极大提升了诊断的准确性,但大多模型仍需人工干预,缺乏信号与故障模式的直接映射,影响模型的适应性;信号与故障模式间存在较强的反馈,模型需要被动学习具体的故障模式,无法探索并主动解决新问题。这些因素均不同程度地影响故障诊断的泛化性、通用性和智能性。

深度强化学习是人工智能的巨大突破,能够赋予智能体强大的感知能力和自主决策能力,具备实现通用人工智能的潜力^[23]。深度Q网络(deep Q network, DQN)^[23]作为首个深度强化学习算法,自提出以来,已在游戏、推荐系统、控制等领域取得惊人的结果^[24]。近年来,经过一些学者的探索,其逐渐可用于分类任务中。Wiering等^[25]提出一种分类马尔可夫决策过程(classification Markov process decision, CMDP),可将分类问题定义为序贯决策任务,在此框架下的MLP模型优于

标准MLP;Lin等^[26]提出不平衡CMDP,并采用DQN解决图像的不平衡分类。在故障诊断领域,Ding等^[27]利用堆栈自编码提取特征,结合DQN实现滚动轴承健康状态辨识,提升诊断的泛化性,但完全忽略了状态间的联系;康守强等^[28]提出一种改进的DQN模型应用于轴承不平衡故障诊断,取得不错效果;Wang等^[29]提出基于时频表征与DQN的行星齿轮箱智能诊断,增强诊断的泛化性。但文献[28-29]中模型输入均为时频图像,不仅需要一定的专业知识,且可能会破坏原始信号的相关性,丢失部分故障信息。

因此,本文立足于深度强化学习,提出一种基于深度宽卷积Q网络(deep wide convolution Q network, DWCQN)的行星齿轮箱故障智能诊断方法。利用CMDP将故障诊断描述为一个序贯决策任务,建立故障诊断模拟环境;提出一种DWCQN模型,通过与环境不断地交互自主学习诊断策略,实现“端到端”智能诊断;试验及案例分析验证了所提方法的有效性,相较于SVM, DNN, WDCNN等,结果表明所提方法在多个工况下均有超过99%的准确率和更好的稳健性,增强了行星齿轮箱智能诊断的泛化性和通用性。

1 理论背景

1.1 分类马尔科夫决策过程

深度强化学习主要解决序贯决策任务,为此本文定义一个CMDP,由元素 $\{S, A, R, P, \gamma\}$ 构成,将故障诊断分解为简化的序贯决策问题,其中 S 表示状态集, A 表示动作集, R 表示奖励策略, P 表示状态转移概率, γ 表示折扣因子^[26,29],总体过程如图1所示。

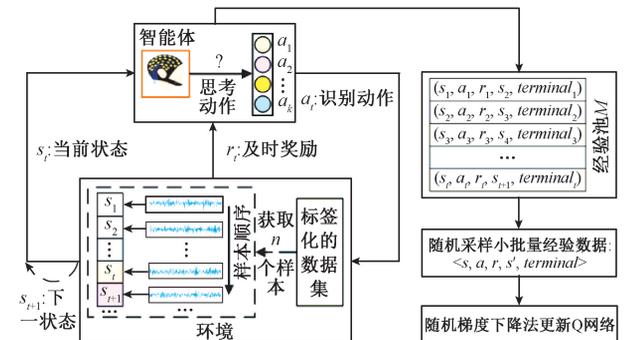


图1 CMDP的总体流程

Fig. 1 Overall flow of CMDP

假设训练数据集为 $D = \{(x_1, l_1), (x_2, l_2), \dots, (x_N, l_N)\}$, D 中样本有序,其中 x_i 表示第 i 个样本, l_i 表示样本 x_i 的标签, N 为样本总量。CMDP及相关变量定义如下。

1) 状态集 S : 包含环境的所有状态, $S = \{s_1, s_2, \dots, s_N\}$, 其中状态 s 取决于 D 中样本 x, t 时间步的状态 s_t 对应于样本 x_t 。环境初始状态 s_1 对应于样本 x_1 , 每次开始新的情节 *Episode* 时, 随机打乱训练集 D 。

2) 动作集 A : 智能体可执行的所有动作, 每个动作与 D 中标签相关, 动作 a_t 是智能猜测的某类标签, $A = \{0, 1, \dots, K-1\}$, 其中 K 是类别数。

3) 奖励策略 $R: S \times A \rightarrow R$, 奖励 r_t 表示智能体在环境状态 s_t 时执行动作 a_t 后来自环境的反馈, 用于衡量智能体动作的正确与否。为了指导智能体学习最佳识别策略, 本文中奖励函数定义为:

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

4) 转移概率 $P: S \times A \rightarrow S$, 当诊断开始后, 状态转移概率 $p(s_{t+1} | s_t, a_t)$ 在 CMDP 中是确定性的, 表示在当前状态 s_t 执行动作 a_t 后会依据 D 中样本顺序转到下一状态 s_{t+1} 。

5) 折扣因子 $\gamma: \gamma \in [0, 1]$, 用于平衡当前奖励和未来奖励, 当 $\gamma=0$ 表示智能体只在乎当前及时奖励 r_t , 当 $\gamma=1$ 表示智能体更在乎长期回报。

6) 情节 *Episode*: 从初始状态 s_1 到终状态 s_T 的转移轨迹, $Episode = \{s_1, a_1, r_1, s_2, a_2, r_2, \dots, s_T, a_T, r_T\}$, 其中 T 是终止时间步, 当 $T=N$ 时, 结束一个 *Episode*。

7) 策略 π_θ : 映射函数 $\pi: S \rightarrow A, \pi_\theta(s_t)$ 指导智能体在状态 s_t 时执行动作 a_t , 策略 π_θ 在 CMDP 中可看作是带参数 θ 的诊断智能体。

通过上述说明, 故障诊断问题可看作是为了寻找最优的策略 $\pi_\theta^*: S \rightarrow A$, 使得智能体在 CMDP 中获得最大的累计奖励, 从而可采用 DQN 算法进行实现。

1.2 深度 Q 网络理论

DQN^[23] 是由 DeepMind 团队在 Nature 上提出的首个深度强化学习算法, 其将 Q-learning^[30] 强化学习算法和 DNN^[12] 进行有效结合, 在 Atari 2600 视频游戏中实现了超越人类玩家的水平, 且具有较强的适应性和通用性。在 DQN 中, 智能体通过与环境互动, 依据环境的反馈学习最优策略 π^* , 具体如下。

在交互过程中, 智能体学到的策略 π 可通过在状态 s_t 下执行动作 a_t 的概率表示:

$$\pi(a | s) = P(a_t = a, s_t = s) \quad (2)$$

智能体的目标是在策略 π 的指导下获得最大的累积奖励 G_t , 通过累积奖励 G_t 衡量所学策略 π 的优劣。

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k} \quad (3)$$

在时间步 t 时, 通过计算累计奖励 G_t 的期望, 即动作值函数 $Q^\pi(s, a)$ 评价智能体在状态 s_t 选择动作 a_t 的好坏程度, 可通过式(4)计算:

$$Q^\pi(s, a) = \mathbb{E}_\pi[G_t | s_t = s, a_t = a; \pi] \quad (4)$$

根据 Bellman 方程^[31], 式(4)可简化为:

$$Q^\pi(s, a) = \mathbb{E}_\pi[r_t + \gamma Q^\pi(s_{t+1}, a_{t+1}) | s_t = s, a_t = a] \quad (5)$$

通常, 最优的动作值函数 $Q^*(s, a)$ 决定了智能体在 CMDP 中最优的可能表现, 其由式(6)计算:

$$Q^*(s, a) = \mathbb{E}_\pi[r_t + \gamma \max_a Q^*(s_{t+1}, a_{t+1}) | s_t = s, a_t = a] \quad (6)$$

最优策略 π^* 可由上述最优动作值函数 $Q^*(s, a)$ 确定, 其过程可表示为:

$$\pi^*(a | s) \rightarrow Q^*(s, a) = \max_a Q^*(s, a) \quad (7)$$

因此, 根据贪婪策略, 通过最大化 $Q^*(s, a)$ 可得最优策略 π^* , 定义为:

$$\pi^*(a | s) = \begin{cases} 1, & a = \operatorname{argmax}_{a \in A} Q(s, a) \\ 0, & \text{其他} \end{cases} \quad (8)$$

不同于 Q-learning 算法采用 Q 表进行记录动作值函数 $Q^\pi(s, a)$, DQN 采用 DNN 实现动作值函数 $Q^\pi(s, a)$ 的估计, 能够更好地适应高维状态空间。此外, DQN 算法中采用经验回放机制保证智能学习的稳定性。如图 1 所示, 将智能与环境交互的经验数据 $e = \{s, a, r, s', terminal\}$ 存储在经验回放池 M 中, 每次从 M 中随机抽取“小批量”数据 B , 根据均方误差 (mean squared error, MSE) 损失函数 $L(\theta)$, 执行梯度下降法更新 Q 网络:

$$L(\theta) = \mathbb{E}[(y - Q(s, a; \theta))^2] \quad (9)$$

式中: θ 表示 DNN 中的模型参数; y 是 Q 值的目标估计, 其可由式(10)获得:

$$y = r + \gamma \max_{a'} Q(s', a'; \theta^-) \quad (10)$$

式中: s', a' 分别表示下一个状态和动作。之后根据 $L(\theta)$ 的梯度对深度 Q 网络进行参数更新, 其可表达为:

$$\nabla_\theta L(\theta) = \mathbb{E}_{e \sim U(M)}[(y - Q(s, a; \theta)) \nabla_\theta Q(s, a; \theta)] \quad (11)$$

式中: $e \sim U(M)$ 表示从 M 中均匀地采样经验数据。

经过上述过程, 通过最小化损失函数 $L(\theta)$, 即可获得最佳动作值函数 $Q^*(s, a)$, 从而完成最佳策略 π^* 的学习。为了使累积奖励 G_t 最大化, DQN 算法还采用了 ϵ -贪婪算法^[23] 权衡学习过程中的“探索”和“利用”。

通过上述 DQN 算法, 即可完成本文所述 CMDP 中最佳诊断策略 $\pi_\theta^*: S \rightarrow A$ 的学习。由于本文中的环境状态为一维的振动信号, 因此 1D CNN 将被用于 DQN 模型的动作值函数估计。

1.3 一维卷积神经网络

CNN 是具有稀疏链接、权值共享特性的前馈性神经网络, 因其强大的特征表达能力, 已广泛应用于图像识别等诸多领域^[12, 32]。典型 CNN 主要包括 5 个部分: 输入层

(Input)、卷积层(convolutional layer, C)、池化层(pooling layer, P)、全连接层(fully-connected layer, FC)和输出层(output)。其中,卷积层和池化层通常交替组合进行逐层特征提取与降维,最后通过全连接完成特征表达,各层含义具体如下。

输入层(input):CNN的输入信号,本文将表征环境状态的振动信号 x_i 作为1D CNN模型的输入。

卷积层(C):通过卷积核对输入的特征矢量进行卷积操作,再经过激活函数提取非线性特征,卷积层运算可表达为:

$$\mathbf{x}_i^{h+1} = f\left(\sum_j \mathbf{x}_j^h * \mathbf{w}_{ij}^{h+1} + \mathbf{b}_i^{h+1}\right) \quad (12)$$

式中: \mathbf{x}_i^{h+1} 是第 $h+1$ 层的第 i 个输出特征图; \mathbf{x}_j^h 为第 h 层的第 j 个输出特征图; \mathbf{w}_{ij}^{h+1} 为连接第 j 个输入与第 i 个特征图的卷积核; \mathbf{b}_i^{h+1} 为第 $h+1$ 层偏置项;*为卷积操作; f 表示激活函数。激活函数通常采用整流线性单元(rectified linear units, ReLU)^[33],其可表达为:

$$\mathbf{v}_i^{h+1}(j) = f(\mathbf{z}_i^{h+1}(j)) = \max\{0, \mathbf{z}_i^{h+1}(j)\} \quad (13)$$

式中: $\mathbf{z}_i^{h+1}(j)$ 为卷积操作的输出值; $\mathbf{v}_i^{h+1}(j)$ 是 $\mathbf{z}_i^{h+1}(j)$ 的激活值。

池化层(P):通过降采样操作减少卷积层输出特征信息冗余,避免模型过拟合,通常采用最大池化(max pooling)操作,其可表示为:

$$\mathbf{P}_i^{h+1}(j) = \max_{(j-1)W+1 \leq m \leq jW} \{\mathbf{q}_i^h(m)\} \quad (14)$$

式中: $\mathbf{q}_i^h(m)$ 表示第 h 层第 i 个特征矢量中第 m 个神经元的值, $m \in [(j-1)W+1, jW]$, W 表示池化区域宽度; $\mathbf{P}_i^{h+1}(j)$ 为第 $h+1$ 层第 i 个特征矢量中第 j 个值。

全连接层(FC):经过多个卷积-池化组合操作后,将输出特征进行“展平”操作,得到一维有序特征矢量,再经FC操作对特征进一步整合,其可表示为:

$$\mathbf{f}\mathbf{c}_i^{k+1} = f\left(\sum_j \mathbf{f}\mathbf{c}_j^k \mathbf{w}_{ij}^{k+1} + \mathbf{b}_i^{k+1}\right) \quad (15)$$

式中: $\mathbf{f}\mathbf{c}_i^{k+1}$ 表示第 $k+1$ 层FC中第 i 个节点值; $\mathbf{f}\mathbf{c}_j^k$ 为第 k 层FC中第 j 个节点值; \mathbf{w}_{ij}^{k+1} 和 \mathbf{b}_i^{k+1} 分别表示相应的连接权值和偏置。

输出层(output):在分类任务中,通常经过最后一个FC后,采用Softmax分类器实现各类别的概率输出,其可表示为:

$$\mathbf{O} = f(\mathbf{f}_v \mathbf{w}_o + \mathbf{b}_o) \quad (16)$$

式中: \mathbf{O} 为分类任务中各类别的概率; \mathbf{f}_v 表示最后一个FC的特征矢量; \mathbf{w}_o 、 \mathbf{b}_o 分别为权值矩阵和偏差向量。

基于上述CNN原理,本文采用1D CNN作为DQN中的Q值网络,以对环境状态进行自动感知,完成深层特征提取。

2 基于深度宽卷积Q网络的智能故障识别

2.1 数据预处理

为了便于诊断模型的学习及测试,将行星齿轮箱在各状况下监测的长序列原始振动信号按照无重叠的方式划分多个样本,如图2所示。此外,分割后的每个样本进行最小-最大归一化(min-max normalization, MMN)到 $[-1, 1]$ 之间,它们被视为环境中的状态,MMN的数学表达为:

$$\hat{x}_i = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}} * (Nor_{\max} - Nor_{\min}) + Nor_{\min} \quad (17)$$

式中: $x_{i\min}$ 为样本 x_i 的最小值; $x_{i\max}$ 为样本 x_i 的最大值; Nor_{\min} 为-1; Nor_{\max} 为1; \hat{x}_i 为归一化后的样本, i 为样本序号, $i = 1, 2, \dots, n$, n 为样本量。

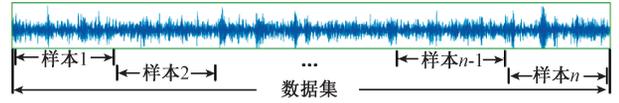


图2 数据划分方式

Fig. 2 Data segment way

2.2 Q网络结构设计

根据1.3节所述一维CNN原理,受WDCNN结构^[22]启发,CNN第1层具有宽的卷积核可增大感受视野,提升性能,设计了一种第1层具有宽卷积核的CNN作为DQN的动作值(Q值)网络,增强对状态信息的感知能力,网络结构如图3所示。该网络包括1个输入层Input,5个卷积层C1~C5,5个池化层P1~P5,1个“展平”层Flatten6,2个全连接层FC7和FC8。其中,输入层为振动信号样本状态,卷积层和池化层交替组合,卷积层C1的卷积核大小为 64×1 ,卷积层C2~C5的卷积核大小均为 3×1 ,所有卷积操作均采用“SAME”填充方式,所有池化层P1~P5均是最大池化“MaxPooling”操作,全连接层FC7的节点数为100,C1~C5和FC7层中的激活函数均是“ReLU”^[33]。

此外,不同于一般分类网络的输出层激活函数采用“Softmax”,本文中输出层FC8的输出是 K 维的动作值,其激活函数为线性激活“Linear”。

2.3 诊断流程

基于前述CMDP、CNN和DQN算法,所提基于DWCQN的行星齿轮箱智能诊断总体框架如图4所示,主要分为3大步骤。

1)数据采集及预处理:通过振动传感器采集行星齿轮箱振动信号,根据图2进行样本分割,通过式(17)进行MMN归一化,获得训练集 D_{train} 和测试集 D_{test} 。

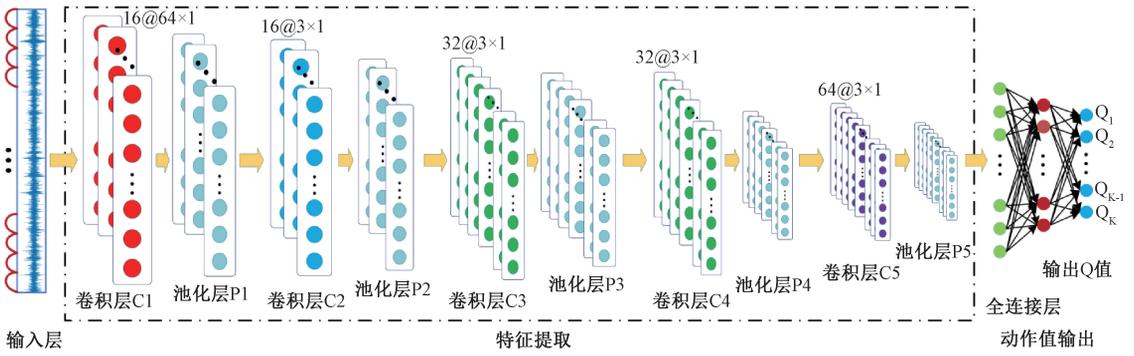


图3 设计的DWCQN模型结构

Fig. 3 Architecture of the designed DWCQN model

2) 建立诊断模拟器与自主学习阶段:在所定义的CMDP基础上,建立故障诊断模拟环境,如算法1所示。在此基础上,诊断智能体通过与环境进行交互,不断地试错,并利用经验更新DWCQN模型,从而自主完成诊断策略学习,自主学习的算法过程如算法2所示。在所提DWCQN模型中,当前网络(Eval-Net)和目标网络

(Target-Net)采用相同的CNN网络结构。其中,每隔延迟C步,复制当前网络Eval-Net的参数给目标网络Target-Net,从而更新Target-Net网络。在学习过程中,每个情节Episode开始时从训练集 D_{train} 中随机抽取T个样本构成一轮诊断识别中的样本集 D_T ,并在结束一个情节后,随机打乱训练集 D_{train} 。

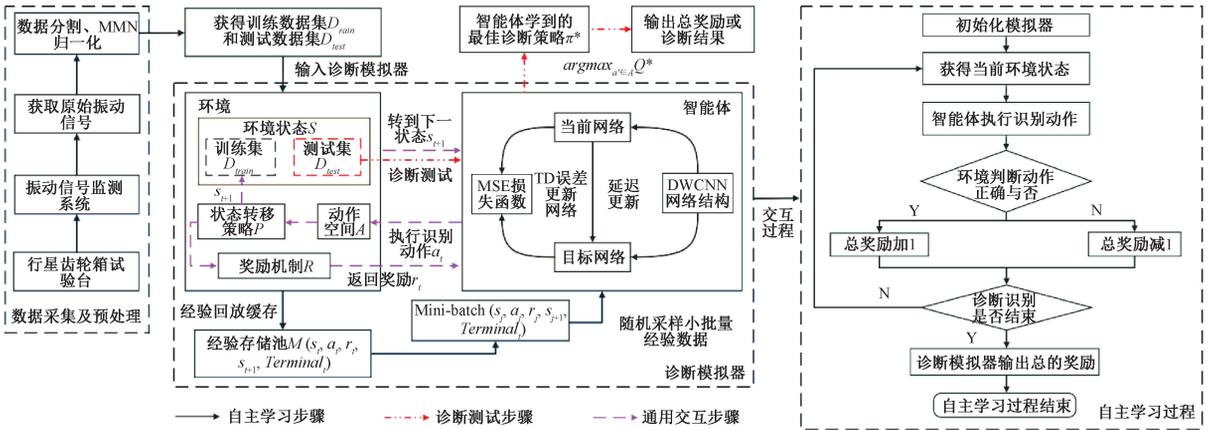


图4 所提基于DWCQN的行星齿轮箱智能诊断框架

Fig. 4 The intelligent diagnosis framework of the planetary gearbox based on DWCQN

算法1 环境仿真

初始化: $D_T = \{(x_1, l_1), (x_1, l_1), \dots, (x_T, l_T)\}$, $D_T \in D_{train}$

Class Environment(): #定义环境类

def reset(): #环境初始化

obs, __, __ = step(-1)

return obs

def step(a_t):

if $a_t = -1$ then

$s_t = x_1$

$r_t = 0$

$terminal_t = False$

else

$r_t = reward(a_t)$

$s_{t+1} = x_{t+1}$

while $t = T$ then

$terminal_t = True$

return $r_t, s_{t+1}, terminal_t$ #返回奖励,下一状态,终止标志

标志

def reward(a_t):

if $a_t = l_t$ then

$r_t = 1$

else

$r_t = -1$

return r_t

算法2 自主学习过程

输入: $D_{train} = \{(x_1, l_1), (x_1, l_1), \dots, (x_N, l_N)\}$; 情节 $Episode$ 最大量 E

初始化: 当前网络 Eval-Net 的权重 θ ; 目标网络 Target-Net 的权重 θ^- 批大小 $mini\text{-}batch$; 经验回放池 M ; 参数 ϵ, C, γ

训练:

for episode 1 to E **do**

reset () #环境初始化

count = 0

打乱训练集 D_{train}

for $t=1$ to T **do**

依据 ϵ -贪婪策略执行动作 $a_t = \pi_\theta(s_t)$: 以概率 ϵ 从动作空间 A 随机选择动作 a_t , 或执行 $a_t = \operatorname{argmax} Q(s_t, a; \theta)$

$r_t, s_{t+1}, terminal_t = \text{step}(a_t)$

存储经验 $e_t = (s_t, a_t, r_t, s_{t+1}, terminal_t)$ 到经验回放池 M 中

if M 的容量 $> mini\text{-}batch$ **then**

从 M 中随机采样 $mini\text{-}batch$ 数据 $B = (s_j, a_j,$

$r_j, s_{j+1}, terminal_j)$

根据式(10)计算 y_j :

$$y_j = \begin{cases} r_j, & terminal_j = \text{True} \\ r_j + \gamma \max_{a'} Q(s_{j+1}, a'; \theta^-), & terminal_j = \text{False} \end{cases}$$

根据式(11)执行梯度下降法最小化 MSE 损失函数式(9)

更新当前网络 Eval-Net 的模型参数 $\theta, count += 1$

根据式(18)更新 ϵ 的值

if $count \neq 0$ 且 $count \% C = 0$ **then**

更新目标网络 Target-Net 的模型参数 $\theta^- = \theta$

计算此情节中获得的总奖励

输出: 每个情节 $Episode$ 中的奖励变化, 保存诊断策略 π_θ^*

在学习过程的初期加大随机“探索”, 随着学习的深入, 逐渐降低“探索”概率, 增加“利用”, 以期更好学习诊断策略。因此, 本文中 ϵ -贪婪算法中的参数 ϵ 是动态变化的, 可通过式(18)表示:

$$\epsilon = \begin{cases} \epsilon \times \epsilon_{decay}, & \epsilon \geq \epsilon_{min} \\ \epsilon_{min}, & \text{其他} \end{cases} \quad (18)$$

式中: ϵ_{min} 表示 ϵ 的最小值; ϵ_{decay} 表示衰减系数。

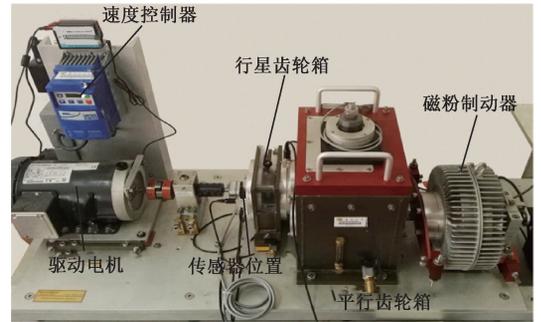
3) 诊断测试阶段: 智能体通过步骤 2) 完成自主学习后, 在所学最佳诊断策略 π_θ^* 的指导下在测试集 D_{test} 上完成行星齿轮箱的状态辨识, 并对所提方法进行评估分析。其中, 行星齿轮箱的状态辨识过程为: 将测试集 D_{test}

中状态逐次传送给当前网络 Eval-Net 计算动作值 $Q(s, a)$, 并依据贪婪策略选择最大动作值 $\max Q(s, a)$ 对应的动作 $a = \operatorname{argmax}_a Q(s, a)$, 进而根据标签信息评价动作的优劣, 分析模型性能。

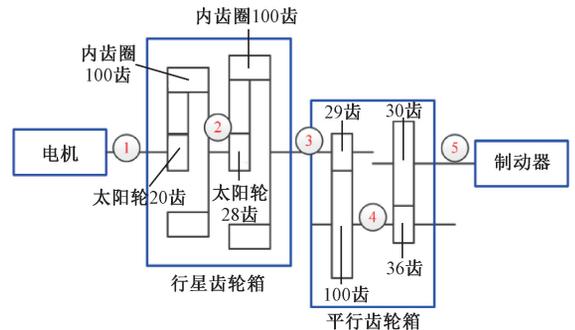
3 实验及结果分析

3.1 实验平台

本文在 SpectraQuest 动力传动综合实验台 (drive-train dynamics simulator, DDS) 上模拟不同转速-负载工况下行星齿轮箱的典型故障, 通过加速度传感器采集相应的振动信号, 以验证所提方法的有效性。该 DDS 综合实验台, 如图 5(a) 所示, 主要由两级行星齿轮箱、两级平行齿轮箱、驱动电机、振动传感器、速度控制器和制动器等组成, 其结构及关键参数如图 5(b) 所示。其中, 振动传感器的型号为 PCB 608A11, 垂直安装在一级行星齿轮箱输入端以采集行星齿轮箱的运行数据, 采样频率为 5 120 Hz。实验共模型了行星齿轮箱工作在 4 种转速-负载工况条件下, 具体包括电机转速-负载分别为 1 200 r/min-0 N·m、1 800 r/min-3.66 N·m、1 800 r/min-10.98 N·m 和 2 400 r/min-10.98 N·m。在上述每种工况下植行星齿轮箱的 9 种健康状况 ($K=9$), 包含完全健康状态, 4 种齿轮



(a) DDS综合实验台
(a) DDS test rig



(b) 结构示意图
(b) Structure diagram

图5 行星齿轮箱故障诊断综合试验台及结构

Fig.5 Planetary gearbox test rig and structure

故障:切齿故障、缺齿故障、齿根裂纹故障和表面磨损故障,4 种轴承故障:滚动体磨损故障、复合磨损故障、内圈故障以及外圈故障。实验中各种植入故障的失效图如图 6 所示。

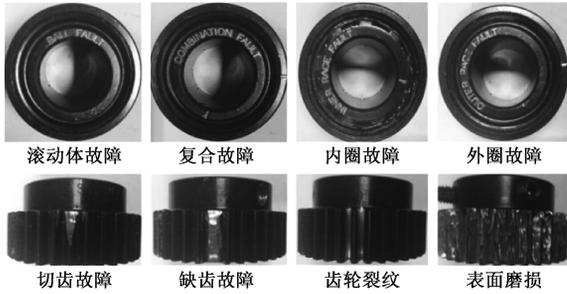


图 6 行星齿轮箱中典型故障模式部件

Fig. 6 Typical faulty parts of the planetary gearbox

3.2 数据说明

上述 4 种工况分别记为工况 A, B, C 和 D, 每种工况下获取的每类行星齿轮箱状态的原始振动信号长度为 4 096 000, 按照图 2 所示分割成许多长度为 $1 \times 2\ 048$ 的样本, 每类行星齿轮箱状态的样本量 n 为 2 000, 因此, 每种工况下行星齿轮箱 9 类健康状况的样本总量 N 为 18 000, 详情如表 1 所示。此外, 在诊断过程中, 每个样本在输入诊断模型前根据式 (17) 所示 MMN 进行归一化预处理。图 7 所示一组行星齿轮箱在工况 B 下 9 类状态的样本振动信号及其频谱。从图 7 可看出, 难以从时域振动信号及其频谱中直接区分 9 类行星齿轮箱的健康状况, 尤其对于复杂的齿轮故障, 其时域振动信号及频谱颇为相似。

表 1 试验中行星齿轮箱的 9 种健康状况
Table 1 Nine health conditions of the planetary gearbox in test

动作编号	健康状况的描述	动作标签	样本长度	样本数量
0	健康状态	HEA	$1 \times 2\ 048$	2 000
1	齿轮切齿故障	CTF	$1 \times 2\ 048$	2 000
2	齿轮缺齿故障	MTF	$1 \times 2\ 048$	2 000
3	齿轮根部裂纹故障	RCF	$1 \times 2\ 048$	2 000
4	齿轮表面磨损故障	SWF	$1 \times 2\ 048$	2 000
5	滚动轴承滚动体故障	BWF	$1 \times 2\ 048$	2 000
6	滚动轴承复合磨损故障	CWF	$1 \times 2\ 048$	2 000
7	滚动轴承内圈故障	IRF	$1 \times 2\ 048$	2 000
8	滚动轴承外圈故障	ORF	$1 \times 2\ 048$	2 000

3.3 DWCQN 模型详细参数

根据 2.2 节中设计准则, 所建 DWCQN 网络的详细参数如表 2 所示。模型输入层为 $2\ 048 \times 1$ 的振动信号,

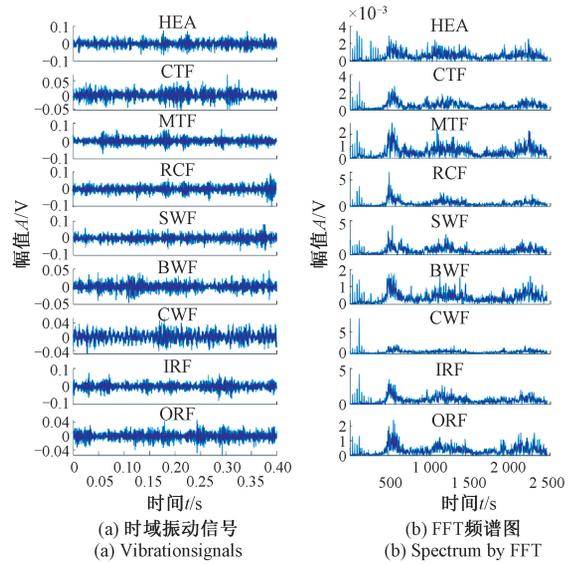


图 7 9 种状态下的时域振动信号和频谱

Fig. 7 Vibration signal and its spectrum under nine modes

所建 DWCQN 采用 Adam 优化器更新参数, 学习率设置为 0.001。本文式 (18) 中参数 ϵ_{\min} 为 0.01, ϵ_{decay} 为 0.995, ϵ 初始值为 1, 最大情节 E 设置为 2 000; 终止步长 T 设置为 500; “批大小” $mini\text{-}batch$ 为 128; 延迟更新步长 $C = 5$; 经验回放池 M 的最大存储量为 20 000。鉴于样本间的相关性较小, 智能体更在乎当前的及时奖励^[29], 因此折扣因子 γ 设置为 0.1。

表 2 详细的 DWCQN 网络结构参数

Table 2 Detailed parameters of the designed DWCQN

层名称	结构参数	输出大小	填充
Input		$2\ 048 \times 1$	
C1	CW=64; CH=1; CC=1; CN=16; Strides=8	256×16	是
P1	P_size=2; P_stride=2	128×16	否
C2	CW=3; CH=1; CC=16; CN=16; Strides=1	128×16	是
P2	P_size=2; P_stride=2	64×16	否
C3	CW=3; CH=1; CC=16; CN=32; Strides=1	64×32	是
P3	P_size=2; P_stride=2	32×32	否
C4	CW=3; CH=1; CC=32; CN=32; Strides=1	32×32	是
P4	P_size=2; P_stride=2	16×32	否
C5	CW=3; CH=1; CC=32; CN=64; Strides=1	16×64	是
P5	P_size=2; P_stride=2	8×64	否
Flatten6		512	
FC7	Dense=100	100	
FC8	Dense=K	K	

注: * CW、CH、CC、CN 分别表示卷积核的宽度、高度、通道数和卷积核数量; Strides 表示卷积核步幅; P_sizes 是池化大小; P_stride 是池化步幅。

此外,所设计的模型均采用 python 3.6、Keras 2.1.6、TensorFlow 1.8.0 进行程序化,在 Ubuntu 16.04 操作系统上运行,CPU 和 GPU 具体配置如下: Intel © Xeon (R) CPU E5-2603 v4@ 1.70 GHz×12, NVIDIA GeForce GTX 1080Ti 11 GB。

3.4 结果对比及分析

1) 评价策略及指标

为了充分利用样本,实验中使用了五折交叉验证策略进行评价诊断模型的性能。首先将每种工况下总样本量 N 随机均匀分为 5 个子数据集,每个子数据集样本量为 $N/5$ 。然后,依次每一个子数据集作为测试集,余下 4 个子数据集作为训练集,进行模型性能评估。其中每次测试的总体准确率(overall accuracy, OA)作为每次诊断的定量分析指标,如式(19)所示。

$$OA = \frac{\sum_{i=1}^K x_{test,correct}^{(i)}}{N_{test}} \quad (19)$$

式中: $x_{test,correct}^{(i)}$ 表示测试样本中第 i 类正确分类的样本个数; N_{test} 为测试样本量。最后将 5 次测试的平均 OA 作为模型的最终评价指标。

2) 对比方法

为了验证所提方法的有效性,本文对比于 3 种传统诊断方法和 4 种基于深度模型的诊断方法,具体为: SVM:人工提取 28 个时频域特征,包含均值、方差、偏度、峭度、峰峰值、裕度因子、波形因子等^[5,29],再采用 SVM 进行分类,其中 SVM 的核函数为高斯径向基函数(RBF),核系数为 0.01,惩罚系数为 1; PCA-SVM:利用主成分分析(principal component analysis, PCA)将 28 个特征降到 15 维,再采用 SVM 进行分类; MLP:其结构为 {层 1: 200 节点,层 2: 50 节点}; DNN:其结构为 {层 1: 500 节点,层 2: 200 节点,层 3: 100 节点,层 4: 50

节点};以及文献[22]与[34]中的诊断方法,如: Alexnet^[34]、LetNet5^[34]和 WDCNN^[22]。此处,MLP、DNN、Alexnet、LetNet5、WDCNN 和所提 DWCQN 模型的输入均为 2 048×1 的振动信号,模型训练的“批大小”均为 128,学习率均为 0.001。其中,MLP 和 DNN 的训练 epoch 为 20, Alexnet、LetNet5 与 WDCNN 的训练 epoch 为 100。

3) 试验结果比较

表 3 所示为各方法在 4 种工况下五折交叉验证的平均测试准确率。从表 3 中可以看出,在工况 A 下,所提 DWCQN 模型的平均测试准确率为 99.60%,相较于 SVM、PCA-SVM、MLP、DNN、Alexnet 和 LetNet5 模型分别提升了 22.40%、18.21%、30.81%、27.81%、13.43%、8.83%,即使对比于目前先进的 WDCNN 模型仍提升了 2.91%;在工况 B 下,所提 DWCQN 模型的平均测试准确率为 99.77%,相较于 SVM、PCA-SVM、MLP、DNN、Alexnet、LetNet5 和 WDCNN 诊断模型分别提升了 35%、30.09%、40.34%、35.51%、9.22%、16.24% 和 1.09%;在工况 C 下,所提 DWCQN 模型的五折交叉验证平均测试准确率为 99.13%,相较于上述 7 种诊断方法分别提升了 29%、23.56%、35.63%、36.54%、14.20%、20.00% 和 1.73%;在工况 D 下,所提 DWCQN 模型的五折交叉验证平均测试准确率为 99.82%,相较于 SVM、PCA-SVM、MLP、DNN、Alexnet、LetNet5 和 WDCNN 方法分别提升了 20.29%、12.65%、11.38%、7.86%、11.46%、11.25% 和 2.03%。在 4 种工况下,所提 DWCQN 方法的诊断准确率的平均值(average, AVG)达到了 99.58%,相较于上述 7 种方法分别提升了 26.67%、21.13%、29.55%、26.83%、12.08%、14.08% 和 1.94%。此外,本文所提 DWCQN 方法在五折交叉测试准确率的标准差在 4 种工况下分别为 0.25%、0.07%、0.36% 和 0.13%,相较于其他 7 种方法, DWCQN 模型在诊断过程中具有较好的诊断稳定性。

表 3 不同方法下五折交叉验证的总体准确率比较

Table 3 Testing OA comparisons of different methods of five-fold cross-validation

方法	4 种不同工况				AVG
	A	B	C	D	
SVM	77.20±0.34	64.77±0.43	70.13±0.96	79.53±0.72	72.91±0.61
PCA-SVM	81.39±0.29	69.68±0.47	75.57±0.83	87.17±0.41	78.45±0.50
MLP	68.79±2.25	59.40±1.69	63.50±0.98	88.44±0.58	70.03±1.38
DNN	71.79±1.90	64.26±2.25	62.99±1.28	91.96±2.11	72.75±1.89
Alexnet	86.17±4.49	90.55±2.23	84.93±2.29	88.36±1.73	87.50±2.69
LeNet5	90.77±2.56	83.53±4.08	79.13±4.28	88.57±1.28	85.50±3.05
WDCNN	96.69±2.52	98.68±1.36	97.40±1.33	97.79±1.97	97.64±1.80
DWCQN	99.60±0.25	99.77±0.07	99.13±0.36	99.82±0.13	99.58±0.20

表4所示为在工况A下8种方法的训练及测试时间对比。由于本文所提方法中诊断智能体是通过与环境进行交互,不断地试错,并依据环境反馈的奖励进行策略学习,所以模型学习过程较耗时。从表4可得,SVM、PCA-SVM, MLP、DNN、Alexnet、LetNet5 和 WDCNN 模型在工况A下的平均训练耗时分别为 10.93、6.56、16.98、23.21、171.21、154.66、280.55 s,而本文方法却高达 50 887.93 s。其中,PCA-SVM 比 SVM 耗时低,仅因特征降维后输入 SVM 模型的特征量减少了。然而当智能体学到最佳诊断策略后,其测试时间将同一般方法区别不大。在工况A下,前述7种方法的平均测试耗时分别为 3.62、2.35、0.27、0.34、0.51、0.55 和 0.68 s,本文方法的测试耗时也仅为 0.38 s,甚至比 WDCNN 更低,这是因为本文方法中所用 Q 网络结构更轻量。

综上,通过4种不同的工况试验,相较于 SVM、MLP、DNN、Alexnet 等方法,本文所提方法具有较好诊断通用性和泛化性,能够适应多种工况下行星齿轮箱的故障诊断需求。虽然模型在训练中需要与环境进行交互,训练时间较长,但当智能体学到最佳诊断策略后,面对新的诊断问题时同一般方法所需的测试时间相差无几,单个样本测试时间仅为 0.11 ms。

表4 各方法的平均时耗对比

Table 4 Average elapsed time of each method s

方法	平均训练时间	平均测试时间
SVM	10.93±0.05	3.62±0.018
PCA-SVM	6.56±0.03	2.35±0.01
MLP	16.98±0.46	0.27±0.01
DNN	23.21±0.60	0.34±0.03
Alexnet	171.21±3.15	0.51±0.02
LeNet5	154.66±3.76	0.55±0.03
WDCNN	280.55±2.83	0.68±0.02
DWCQN	50 887.93±386.40	0.38±0.01

4)t-SNE 可视化分析

为了观察智能体在深度强化学习框架下所学的最优诊断策略,直观地探究模型各层中的特征学习能力,将工况A第5折实验中测试样本作为当前网络 Eval-Net 模型的输入,利用 t 分布随机邻域嵌入(t-distributed stochastic neighbor embedding, t-SNE)^[35]将模型各层学习到的高维特征降至二维空间进行可视化,如图8所示。其中,图8(a)~(e)为各个卷积-池化组合操作后的输出;图8(f)为模型全连接层 FC1 的输出。

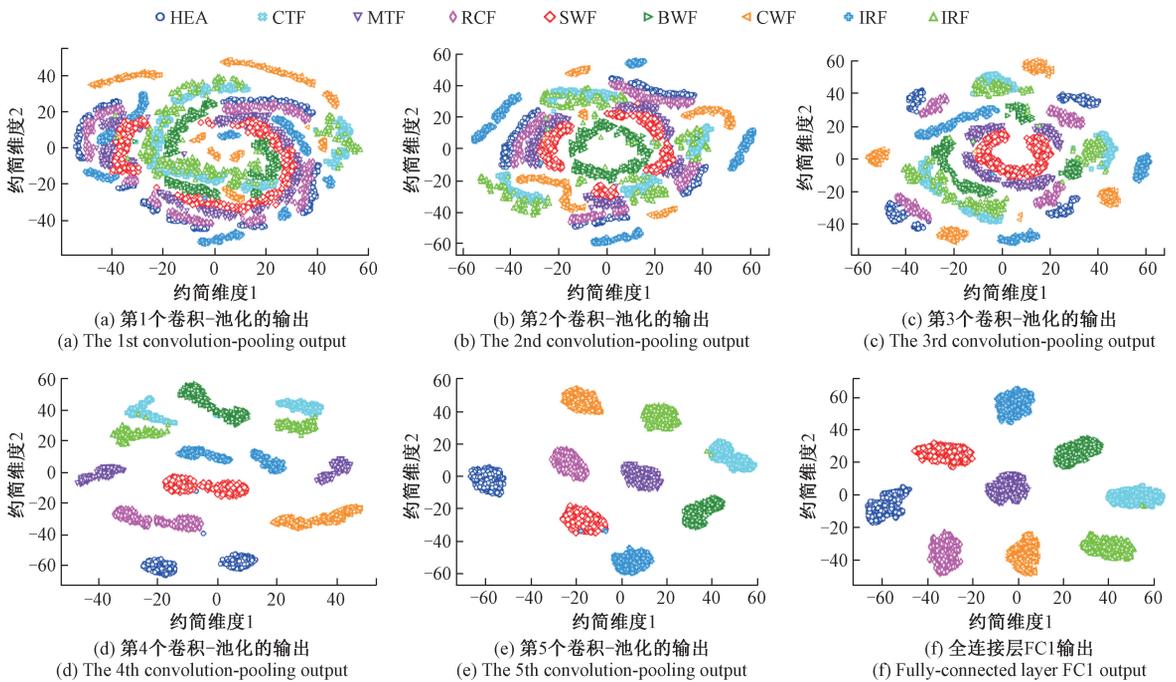


图8 工况A下第5折的测试样本输入 DWDCQN 后 t-SNE 特征可视化结果

Fig. 8 Feature visualization results of t-SNE after 5th-fold testing samples input to DWDCQN under working condition A

从图8(a)~(c)中可以看出前3个卷积-池化操作后,在模型所学策略下各状态特征具有一定的聚类性,但各状态的类内距较大、类间距小,状态间有较多的混叠;

从图8(d)~(e)中可得模型至第4个卷积-池化组合操作后,9类状态基本实现了明显的可分现象,但状态间的类内距仍旧相对较大,相同状态的特征分布具有一定的分

散性,而到了第5个卷积-池化操作后,各状态基本实现了完全可分,类间距变大,类内距变小;从图8(f)可得经过全连接层FC7后,进一步扩大了状态间的特征分布距离,缩小了类内距,使各状态的特征可分性更好。

总体上,从图8可看出本文方所提模型具有较强的自动特征学习能力,能够在所学诊断策略指导下实现对各状态的逐渐可分,提升诊断的准确性。

4 应用案例

4.1 数据说明

本节采用某型直升机齿轮箱的振动监测数据,进一步对所建诊断模型的有效性和泛化性进行实验验证。该直升机齿轮箱共有两大类健康状况,包含健康状态和行星架裂纹故障,其中行星架裂纹长度为8.25 cm。实验在某型直升机传动系统试验设备上执行,原始采样频率为100 kHz。鉴于直升机齿轮箱故障植入成本极其昂贵,这里将实验中直升机齿轮箱分别在扭矩设置为96.28、144.42、240.69、336.97和481.38 N·m下获得的振动数据看作5类不同的状况。因此,在上述设定下,此数据集共有10类直升机齿轮箱运行状况($K=10$),包含5类健康状态和5类故障状态。由于信号采样频率相对较大,为此通过5倍降采样对原始振动信号处理,降采样后的频率为20 kHz,此时每类状态下的振动信号长度为3 584 000,并按照图2进行数据分割,每个样本长度为 $1 \times 2 048$ 。因此,每类直升机齿轮箱的健康状况有1 750个样本,该数据集共有17 500个样本。

4.2 结果对比

表5所示为8种方法的五折交叉验证平均结果。由表5可知,DWCQN方法的平均训练准确率达到99.73%,相比于传统的SVM和PCA-SVM分别提升了3.67%、1.77%,比WDCNN提升了0.67%,这表明DWCQN模型在训练集上达到了较好的效果,诊断智能体能够学习到较好诊断策略。虽然,MLP、DNN、Alexnet和LeNet5在训练集上的准确率均超过了99.5%,且优于WDCNN和DWCQN模型,但这些模型在测试集上的准确率均低于98%,模型的泛化能力不足。而本文所提DWCQN模型在测试集上的准确率达到99.65%,表明所提方法具有较好的泛化性。其中,对比于SVM、PCA-SVM、MLP、DNN、Alexnet和LeNet5,所提DWCQN在测试集上的平均测试准确率分别提升了3.75%、1.78%、1.90%、1.85%、1.88%和1.77%,即使相较于当前较优的WDCNN仍提升了0.73%。此外,DWCQN模型在测试集上五折交叉验证的标准差为0.40%,表明诊断模型具有较强的诊断稳定性。

表5 各方法下五折交叉验证的测试准确率比较

Table 5 Testing accuracy of different methods of five-fold

方法	cross-validation	
	平均训练诊断率	平均测试准确率
SVM	96.06±0.08	95.90±0.28
PCA-SVM	97.96±0.03	97.87±0.06
MLP	100.00±0.00	97.75±0.31
DNN	99.93±0.15	97.80±0.80
Alexnet	99.80±0.36	97.77±0.55
LeNet5	100.00±0.00	97.88±0.36
WDCNN	99.06±1.81	98.92±2.03
DWCQN	99.73±0.30	99.65±0.40

综上所述,本文方法的有效性在直升机齿轮箱健康状态辨识中进一步得到实验验证。同时也表明了所提方法在不同的诊断任务中具有一定的通用性和较强的泛化性,亦有较大的实际应用潜在价值。

5 结论

本文提出一种基于深度宽卷积Q网络(DWCQN)的行星齿轮箱故障诊断方法,通过行星齿轮箱的振动监测数据,智能有效地实现了行星齿轮箱健康状况的辨识,增强了模型的通用性和泛化性。主要总结如下:

1) 该方法利用强化学习的自主决策能力和深度学习的特征学习能力,避免了传统诊断模型过度依赖专家知识,实现了行星齿轮箱的“端到端”故障诊断;

2) 诊断智能体在策略学习过程中,基于奖励的反馈机制降低了通常诊断方法中输入与故障模式间存在的强反馈关系,提升了诊断的智能性;

3) 多个工况试验和应用案例验证了所提方法的可行性和有效性,相较于SVM、PCA-SVM、MLP和DNN等方法,结果表明所提方法具有较高的诊断准确性、稳健性和较好的通用性。

参考文献

- [1] LEI Y, LIN J, ZUO M J, et al. Condition monitoring and fault diagnosis of planetary gearboxes: A review [J]. Measurement, 2014, 48: 292-305.
 - [2] 雷亚国,何正嘉,林京,等.行星齿轮箱故障诊断技术的研究进展[J].机械工程学报,2011,47(19): 59-67.
- LEI Y G, HE ZH J, LIN J, et al. Research advances of fault diagnosis technique for planetary gearboxes [J]. Journal of Mechanical Engineering, 2011, 47(19): 59-67.

- [3] 沈飞, 陈超, 徐佳文, 等. 谱质心迁移在变工况轴承故障诊断的应用 [J]. 仪器仪表学报, 2019, 40(5): 102-111.
- SHENG F, CHEN CH, XU J W, et al. Application of spectral centroid transfer in bearing fault diagnosis under varying working conditions [J]. Chinese Journal of Scientific Instrument, 2019, 40(5): 102-111.
- [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行星齿轮箱故障诊断方法 [J]. 机械工程学报, 2019, 55(7): 9-18.
- HU N Q, CHEN H P, CHENG ZH, et al. Fault diagnosis for planetary gearbox based on EMD and deep convolutional neural networks [J]. Journal of Mechanical Engineering, 2019, 55(7): 9-18.
- [5] LEI Y, HE Z, ZI Y, et al. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs [J]. Mechanical Systems & Signal Processing, 2007, 21(5): 2280-2294.
- [6] WANG T Z, QI J, XU H, et al. Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter [J]. ISA Transactions, 2016, 60: 156-163.
- [7] XU Y, DENG Y, ZHAO J, et al. A novel rolling bearing fault diagnosis method based on empirical wavelet transform and spectral trend [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(6): 2891-2904.
- [8] 郑近德, 潘海洋, 张俊, 等. 基于多变量多尺度模糊熵的行星齿轮箱故障诊断 [J]. 振动与冲击, 2019, 38(6): 192-198.
- ZHENG J D, PAN H Y, ZHANG J, et al. Multivariate multiscale fuzzy entropy based planetary gearbox fault diagnosis [J]. Journal of Vibration and Shock, 2019, 38(6): 192-198.
- [9] SENANAYAKA J, HUYNH K V, ROBBERSMYR K G. Multiple classifiers and data fusion for robust diagnosis of gearbox mixed faults [J]. IEEE Transactions on Industrial Informatics, 2018, 15(8): 4569-4579.
- [10] 张淑清, 苑世钰, 姚玉永, 等. 基于 ICEEMD 及 AWOA 优化 ELM 的机械故障诊断方法 [J]. 仪器仪表学报, 2019, 40(11): 175-183.
- ZHANG SH Q, YUAN SH Y, YAO Y Y, et al. Machinery fault diagnosis method based on ICEEMD and AWOA optimized ELM [J]. Chinese Journal of Scientific Instrument, 2019, 40(11): 175-183.
- [11] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法 [J]. 机械工程学报, 2015, 51(21): 49-56.
- LEI Y G, JIA F, ZHOU X, et al. A deep learning-based method for machinery health monitoring with big data [J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56.
- [12] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [13] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战 [J]. 控制与决策, 2017, 32(8): 1345-1358.
- REN H, QU J F, CHAI Y, et al. Deep learning for fault diagnosis: The state of the art and challenge [J]. Control and Decision, 2017, 32(8): 1345-1358.
- [14] 赵光权, 葛强强, 刘小勇, 等. 基于 DBN 的故障特征提取及诊断方法研究 [J]. 仪器仪表学报, 2016, 37(9): 1946-1953.
- ZHAO G Q, GE Q Q, LIU X Y, et al. Fault feature extraction and diagnosis method based on deep belief network [J]. Chinese Journal of Scientific Instrument, 2016, 37(9): 1946-1953.
- [15] 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 W H, SHAN W P, ZENG X Q. Bearing fault identification based on deep belief network [J]. Journal of Vibration Engineering, 2016, 29(2): 340-347.
- [16] SUN W, SHAO S Y, ZHAO R, et al. A sparse auto-encoder-based deep neural network approach for induction motor faults classification [J]. Measurement, 2016, 89: 171-178.
- [17] 孔子迁, 邓蕾, 汤宝平, 等. 基于时频融合和注意力机制的深度学习行星齿轮箱故障诊断方法 [J]. 仪器仪表学报, 2019, 40(6): 224-230.
- KONG Z Q, DENG L, TANG B P, et al. Fault diagnosis of planetary gearbox based on deep learning with time-frequency fusion and attention mechanism [J]. Chinese Journal of Scientific Instrument, 2019, 40(6): 224-230.
- [18] ZHAO M, KANG M, TANG B P, et al. Deep residual networks with dynamically weighted wavelet coefficients for fault diagnosis of planetary gearboxes [J]. IEEE Transactions on Industrial Electronics, 2018, 65(5): 4290-4300.
- [19] CHEN Z Q, LI C, SANCHEZ R V. Gearbox fault identification and classification with convolutional neural networks [J]. Shock and Vibration, 2015, 2015: 1-10.
- [20] 陈仁祥, 黄鑫, 杨黎霞, 等. 基于卷积神经网络和离散小波变换的滚动轴承故障诊断 [J]. 振动工程学报, 2018, 31(5): 883-891.
- CHEN R X, HUANG X, YANG L X, et al. Rolling

- bearing fault identification based on convolution neural network and discrete wavelet transform [J]. *Journal of Vibration Engineering*, 2018, 31(5): 883-891.
- [21] 曲建岭, 余路, 袁涛, 等. 基于一维卷积神经网络的滚动轴承自适应故障诊断算法 [J]. *仪器仪表学报*, 2018, 39(7): 134-143.
- QU J L, YU L, YUAN T, et al. Adaptive fault diagnosis algorithm for rolling bearings based on one-dimensional convolutional neural network [J]. *Chinese Journal of Scientific Instrument*, 2018, 39(7): 134-143.
- [22] ZHANG W, PENG G L, LI C H, 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(3): 425.
- [23] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning [J]. *Nature*, 2015, 518(7540): 529-533.
- [24] 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 ZH Y, MU CH X, SUN CH Y. An overview on algorithms and applications of deep reinforcement learning [J]. *Chinese Journal of Intelligent Science and Technology*, 2020, 2(4): 314-326.
- [25] WIERING M A, VAN HASSELT H, PIETERSMA A D, et al. Reinforcement learning algorithms for solving classification problems [C]. 2011 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), 2011: 91-96.
- [26] LIN E, CHEN Q, QI X M. Deep reinforcement learning for imbalanced classification [J]. *Applied Intelligence*, 2020, 50: 2488-2502.
- [27] 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.
- [28] 康守强, 刘哲, 王玉静, 等. 基于改进 DQN 网络的滚动轴承故障诊断方法 [J]. *仪器仪表学报*, 2021, 42(3): 201-212.
- KANG SH Q, LIU ZH, WANG Y J, et al. A fault diagnosis method of rolling bearing based on the improved DQN network [J]. *Chinese Journal of Scientific Instrument*, 2021, 42(3): 201-212.
- [29] WANG H, XU J W, 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.
- [30] WATKINS C J C H, DAYAN P. Q-learning [J]. *Machine Learning*, 1992, 8(3-4): 279-292.
- [31] BARRON E N, ISHII H. The Bellman equation for minimizing the maximum cost [J]. *Nonlinear Analysis: Theory, Methods & Applications*, 1989, 13(9): 1067-1090.
- [32] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25: 1097-1105.
- [33] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]. *International Conference on Machine Learning*, 2010: 807-814.
- [34] 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.
- [35] LAURENS V D M, HINTON G. Visualizing data using t-SNE [J]. *Journal of Machine Learning Research*, 2008, 9(2605): 2579-2605.

作者简介



王辉, 2015 年于湖南工业大学获得学士学位, 2018 年于安徽大学获得硕士学位, 现为东南大学博士研究生, 主要研究方向为旋转机械的智能诊断。

E-mail: 230189693@seu.edu.cn

Wang Hui received his B. Sc. degree from Hunan University of Technology in 2015, and received his M. Sc. degree from Anhui University in 2018. He is currently a Ph. D. candidate at Southeast University. His main research interest is intelligent diagnosis of rotating machinery.



严如强(通信作者), 分别在 1997 年和 2002 年于中国科学技术大学获得学士学位和硕士学位, 2007 年于美国马萨诸塞大学阿默斯特分校获得博士学位, 现为西安交通大学教授、博士生导师, 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。

无线传感网络。

E-mail: ruqiang@seu.edu.cn

Yan Ruqiang (Corresponding author) received his B. Sc. and M. Sc. degrees both from University of Science and Technology of China in 1997 and 2002, and received his Ph. D. degree from University of Massachusetts, Amherst, USA in 2007. He is currently a professor and a Ph. D. advisor at Xi'an Jiaotong University. His main research interests include mechanical system condition monitoring and fault diagnosis, signal processing, and wireless sensor network.