



基于柔性残差神经网络的滚动轴承智能故障诊断方法

陈闯 李先锋 史建涛 岳冬冬

Intelligent fault diagnosis method for rolling bearings based on flexible residual neural network

CHEN Chuang, LI Xianfeng, SHI Jiantao, YUE Dongdong

引用本文:

陈闯, 李先锋, 史建涛, 岳冬冬. 基于柔性残差神经网络的滚动轴承智能故障诊断方法[J]. 北科大: 工程科学学报, 2025, 47(3): 480–488. doi: 10.13374/j.issn2095–9389.2024.06.24.006

CHEN Chuang, LI Xianfeng, SHI Jiantao, YUE Dongdong. Intelligent fault diagnosis method for rolling bearings based on flexible residual neural network[J]. *Chinese Journal of Engineering*, 2025, 47(3): 480–488. doi: 10.13374/j.issn2095–9389.2024.06.24.006

在线阅读 View online: <https://doi.org/10.13374/j.issn2095–9389.2024.06.24.006>

您可能感兴趣的其他文章

Articles you may be interested in

一种基于轻量级神经网络的高铁轮对轴承故障诊断方法

Fault diagnosis of high-speed train wheelset bearing based on a lightweight neural network

工程科学学报. 2021, 43(11): 1482 <https://doi.org/10.13374/j.issn2095–9389.2020.12.09.001>

基于KPCA–MTCN的锂离子电池故障诊断方法

Lithium-ion battery fault diagnosis method based on KPCA–MTCN

工程科学学报. 2024, 46(12): 2297 <https://doi.org/10.13374/j.issn2095–9389.2024.01.30.001>

基于深度学习的高超声速飞行器执行器零样本故障辨识

Deep learning based actuator fault identification for hypersonic vehicles: A zero-shot case

工程科学学报. 2024, 46(9): 1613 <https://doi.org/10.13374/j.issn2095–9389.2023.09.27.001>

基于WOA–VMD与PSO–SVM的锂离子电池内短路故障诊断方法

Research on internal short-circuit fault diagnosis methods for lithium-ion batteries based on WOA–VMD and PSO–SVM

工程科学学报. 2023, 45(12): 2162 <https://doi.org/10.13374/j.issn2095–9389.2022.10.04.004>

自适应动模式分解和GA–SVM在行星轴承故障分类中的应用

Adaptive dynamic mode decomposition and GA–SVM with application to fault classification of planetary bearing

工程科学学报. 2023, 45(9): 1559 <https://doi.org/10.13374/j.issn2095–9389.2022.07.01.001>

基于3D卷积神经网络的膏体屈服应力预测

Prediction of paste yield stress based on three-dimensional convolutional neural networks

工程科学学报. 2024, 46(8): 1337 <https://doi.org/10.13374/j.issn2095–9389.2023.10.11.005>

基于柔性残差神经网络的滚动轴承智能故障诊断方法

陈 闯^{1,2)}, 李先锋¹⁾, 史建涛¹⁾✉, 岳冬冬¹⁾

1) 南京工业大学电气工程与控制科学学院, 南京 211816 2) 上海交通大学电子信息与电气工程学院, 上海 200240

✉通信作者, E-mail: sjt11@tsinghua.org.cn

摘 要 滚动轴承作为旋转机械的重要组成部分, 其正常运行直接影响机器的使用寿命和运行状态. 为了提高滚动轴承故障诊断的准确性, 本文提出一种基于动态减法平均优化器(DSABO)和平行注意力模块(PAM)的柔性残差神经网络(FResNet), 用于滚动轴承故障诊断. 具体而言, 首先设计一种基于卷积神经网络的柔性残差模块来构建 FResNet. 该模块允许在 DSABO 迭代时更改卷积层数、卷积核数和跳跃连接数, 从而增强网络故障特征提取能力并减少网络退化. 其次, 设计具有卷积层的 PAM 来融合通道注意力和空间注意力输出权重, 通过与滚动轴承运行数据结合, 实现数据特征增强. 于是, DSABO、PAM 和 FResNet 的集成形成了一个有效的滚动轴承故障诊断模型, 命名为 DSABO-PAM-FResNet. 最后, 利用美国凯斯西储大学滚动轴承故障数据集验证所提 DSABO-PAM-FResNet 模型的可行性和有效性. 结果显示, 在信噪比为−6 dB 环境下所提模型对滚动轴承故障诊断的准确率为 97.18%, 证明所提模型具有较好的抗噪能力; 在 0.75 kW、1.5 kW 和 2.25 kW 不同负载条件下, 所提模型对滚动轴承故障诊断的平均准确率为 98.2%, 证明所提模型具有良好的变工况诊断适应能力. 与其他智能故障诊断方法的对比结果表明, 所提 DSABO-PAM-FResNet 模型的诊断精度更高, 为滚动轴承故障诊断提供了一种新的有效智能方法.

关键词 轴承故障诊断; 柔性残差神经网络; 动态减法平均优化器; 平行注意力模块; 噪声干扰

分类号 TP183

Intelligent fault diagnosis method for rolling bearings based on flexible residual neural network

CHEN Chuang^{1,2)}, LI Xianfeng¹⁾, SHI Jiantao¹⁾✉, YUE Dongdong¹⁾

1) College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, China

2) School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

✉Corresponding author, E-mail: sjt11@tsinghua.org.cn

ABSTRACT Rolling bearings play a crucial role in rotating machinery, and their efficient operation is vital for the machine's longevity and performance. In numerous real-world situations, diagnosing faults in rolling bearings presents significant challenges. Signals obtained from industrial applications often contain unavoidable noise, complicating analysis. Additionally, the intricate working conditions in actual operations can greatly influence bearing signal characteristics. Consequently, traditional diagnostic techniques struggle to effectively handle the effects of varying loads and noise. To improve the accuracy of fault diagnosis for rolling bearings in noisy and variable working conditions, a new approach using a flexible residual neural network (FResNet) is introduced. This network is built on a dynamic subtraction average-based optimizer (DSABO) and a parallel attention module (PAM). The core of FResNet is a flexible residual module based on convolutional neural networks, which allows for adjustments in the number of convolutional layers,

收稿日期: 2024-06-24

基金项目: 国家自然科学基金资助项目(62303217, 62373184, 62333010); 中国博士后科学基金资助项目(2024M751977); 江苏省自然科学基金资助项目(BK20211502); 江苏省高等学校自然科学基金资助项目(23KJB510006)

convolutional kernels, and skip connections during optimization. These design features improve the network's ability to extract fault features and prevent degradation. Second, a DSABO with a dynamic position update strategy is proposed for parameter optimization of the above FResNet with the flexible residual module. This optimizer helps the model avoid being trapped in local optima, strengthening the fault diagnosis performance of the network. Third, a PAM is integrated, featuring convolutional layers that combine channel and spatial attention. This integration enhances data feature extraction by aligning it with rolling bearing operation data. Together, DSABO, PAM, and FResNet create an effective rolling bearing fault diagnosis model known as DSABO-PAM-FResNet. Finally, the feasibility and effectiveness of the proposed DSABO-PAM-FResNet model are validated using the rolling bearing fault dataset from Case Western Reserve University in the United States. The ablation experiments reveal that the DSABO model consistently achieves accuracies above 97% across different noise environments. This performance surpasses that of models using grey wolf optimizer (GWO), butterfly optimization algorithm (BOA), and whale optimization algorithm (WOA), indicating the excellent search capabilities of the DSABO proposed in this paper. In noisy environments, the model incorporating the PAM module consistently achieves fault recognition accuracies above 97%. This performance exceeds that of models using the efficient channel attention module (ECAM) and spatial attention module (SAM), demonstrating PAM's excellent capability to highlight fault signals. In challenging environments with a signal-to-noise ratio of -6 dB, the proposed model achieves a fault diagnosis accuracy of 97.18%, proving its strong noise resistance. Under different load conditions of 0.75 kW, 1.5 kW, and 2.25 kW the proposed model maintains an average accuracy of 98.2% in environments with a -4 dB signal-to-noise ratio. This demonstrates the model's excellent adaptability to variable working conditions. Comparison results demonstrated that DSABO-PAM-FResNet outperforms other intelligent fault diagnosis methods in terms of diagnostic accuracy, providing a new and effective intelligent method for rolling bearing fault diagnosis.

KEY WORDS bearing fault diagnosis; flexible residual neural network; dynamic subtraction average based optimizer; parallel attention module; noise interference

作为旋转部件的关键组件之一,滚动轴承能够承载负荷,减少摩擦,确保设备平稳运行.滚动轴承不仅能提高系统效率,降低能耗,还能延长设备寿命^[1].在工业生产和交通运输等领域,滚动轴承的可靠性直接关系到生产安全和运行稳定性,因此保障滚动轴承正常运行是维持机械设备高效运转的基础^[2].然而,由于长时间运转、高速旋转和重负荷操作,滚动轴承往往面临摩擦、磨损、损坏等问题,致使机械系统故障风险增加^[3].因此,滚动轴承故障诊断在机械工程领域扮演着至关重要的角色,它可及时发现轴承故障,继而采取维护维修措施降低事故风险,保证了生产现场的安全性^[4].

当前深度学习展现了在大规模数据集中自动提取特征的卓越能力^[5],其强大适应性^[6]和出色表现^[7]使其在轴承故障诊断领域备受瞩目.作为深度学习的主要应用方法之一,卷积神经网络(Convolutional neural network, CNN)因其稀疏交互^[8]、参数共享^[9]、空间池化^[10]等能力,受到了大量学者的关注与青睐.相应地,众多基于CNN或其变种的故障诊断方法被提出. Xu等^[11]通过预训练离线卷积神经网络初始化在线卷积神经网络参数,利用预训练好的源域特征实现域适应,形成用于轴承故障诊断的在线传递卷积神经网络; Sinitsin等^[12]提出了一种基于混合CNN-MLP模型的混合诊断

方法,将传感器采集的信号转换为混合输入数据进行轴承故障诊断; Ruan等^[13]基于轴承加速度信号的物理特性,优化了CNN结构与参数,提高了轴承故障诊断准确率; Ayas等^[14]设计了一种新型深度残差学习网络以学习轴承时域图像与轴承故障类型之间的端到端的映射,获得了较好诊断结果; Zhang等^[15]利用压缩感知进行降噪和特征提取,稀疏处理设备振动数据,摆脱数据干扰和冗余,提高诊断精度; Kong等^[16]训练多个具有不同激活函数的深度自编码器以获得不同类型特征,对获得的特征进行评估并选择,将选择的特征输入CNN,从而获得较好识别效果.上述方法将轴承信号与CNN结合,在处理简单轴承信号时取得了较好诊断精度.然而,在实际工业中,轴承运行工况多变,且常常混杂噪声,使得传统方法难以同时应对工况和噪声两个方面的挑战.

针对以上问题,本文提出一种基于动态减法平均优化器(Dynamic subtraction average based optimizer, DSABO)和平行注意力模块(Parallel attention module, PAM)的柔性残差神经网络(Flexible Residual Neural Network, FResNet),用于滚动轴承故障诊断,命名为DSABO-PAM-FResNet.一方面,所提模型将DSABO与FResNet结合,在DSABO迭代中搜寻使FResNet获得最佳性能的网络结构,

提升网络训练效果. 另一方面, 在 FResNet 网络结构中嵌入 PAM, 关注有效故障特征, 抑制噪声干扰. 利用美国凯斯西储大学的滚动轴承故障数据集, 对所提模型在不同噪声干扰和变负载条件下进行验证. 结果表明, 本文模型在不同噪声和变负载条件下均具有较好的识别精度.

本文的主要贡献如下:

(1) 设计了一种基于 CNN 的柔性残差模块. 该模块允许在 DSABO 迭代时更改卷积层数、卷积核数和跳跃连接数, 从而增强网络故障特征提取能力并减少网络退化.

(2) 提出了基于动态位置更新策略的 DSABO, 用于配置上述 FResNet 的结构参数. 提出的 DSABO 具备跳出局部最优的能力, 从而强化 FResNet 的故障诊断性能.

(3) 设计了一种具有卷积层的 PAM. PAM 融合通道注意力和空间注意力输出权重, 通过与滚动轴承运行数据结合, 实现数据特征增强.

1 相关理论

1.1 卷积神经网络 (CNN)

CNN 是一种深度学习模型, 主要由卷积层、池化层和全连接层组成.

1.1.1 卷积层

卷积层是 CNN 最为重要的组成部分, 主要用于提取特征, 其计算公式为:

$$X_j^l = f \left(\sum_{i \in M_j} X_i^{l-1} \times w_{ij}^l + b_j^l \right) \quad (1)$$

其中, X_j^l 为第 l 层的第 j 个元素; M_j 为第 $l-1$ 层的第 j 个卷积区域; X_i^{l-1} 为对应卷积区域中的元素; w_{ij}^l 为对应卷积核的权重; b_j^l 为偏置项; $f(\cdot)$ 为激活函数.

1.1.2 池化层

池化层常用于减少非必要的特征数量, 同时防止过拟合. 在 CNN 中, 最大池化应用最为广泛, 其计算公式为:

$$P_k^l = \max(u_k^l) \quad (2)$$

其中, u_k^l 是 l 层的第 k 个区域, P_k^l 是该区域的最大池化输出.

1.1.3 全连接层

输入数据经过多个卷积层和池化层后, 提取的特征交由全连接层进行非线性整合, 再经 Softmax 分类, 作为网络输出. 对于任意一个输入的 z 维向量 Q , 其经由 Softmax 的输出 S 为:

$$S_j = \frac{e^{Q_j}}{\sum_{i=1}^z e^{Q_i}}, j = 1, \dots, z \quad (3)$$

其中, S_j 为 S 的第 j 个元素; Q_j 为 Q 的第 j 个元素; e 为自然常数.

1.2 减法平均优化器 (SABO)

减法平均优化器 (Subtraction average based optimizer, SABO) 的基本灵感来自数学概念, 包括平均值、种群个体位置的差异以及目标函数的差值. SABO 利用种群个体的减法平均值更新个体在搜索空间中的位置^[17]. 首先, 减法平均优化器定义了一个新的计算概念 “ $-_v$ ”, 定义如下:

$$A -_v B = \text{sign}(F(A) - F(B))(A - V * B) \quad (4)$$

其中, A 和 B 表示种群个体位置; $\text{sign}(\cdot)$ 为符号函数; $F(\cdot)$ 为适应度函数; V 表示一个维度等于决策变量个数 m 的随机向量, 其中元素在集合 $\{1, 2\}$ 中随机生成; $*$ 表示哈达玛积.

减法平均优化器利用所有种群个体的算术平均位置进行位置更新, 其位置更新的数学表达式如下:

$$DN_i^k = D_i^k + r_i * \frac{1}{N} \sum_{j=1}^N (D_i^k -_v D_j^k) \quad (5)$$

$$D_i^{k+1} = \begin{cases} DN_i^k, & F(DN_i^k) < F(D_i^k) \\ D_i^k, & F(DN_i^k) \geq F(D_i^k) \end{cases} \quad (6)$$

其中, D_i^k 表示第 k 次迭代中的第 i 个个体位置; DN_i^k 表示第 k 次迭代中第 i 个个体基于当前位置的理论更新位置; D_i^{k+1} 表示第 $k+1$ 次迭代中的第 i 个个体位置; r_i 表示维度为 m 的随机向量, 包含元素为 $[0, 1]$ 内的随机数且满足正态分布; N 为种群个体总数.

减法平均优化器使用所有种群个体的算术平均位置, 而不是仅仅使用最佳或最差种群个体的位置来更新所有种群个体的位置, 具有全局寻优能力强, 不易陷入局部最优等特点.

2 基于 DSABO-PAM-FResNet 的轴承故障诊断模型

2.1 故障诊断模型框架

在实践中, 滚动轴承的故障诊断面临着许多挑战. 首先, 对于从实际工业应用中获得的信号, 大量噪声是不可避免的. 其次, 现实中复杂多变的工况会极大影响轴承信号. 对此, 本文提出了一种基于 DSABO 和 PAM 的 FResNet (命名为 DSABO-PAM-FResNet), 用于滚动轴承故障诊断. 所提 DSABO-PAM-FResNet 模型框架如图 1 所示.

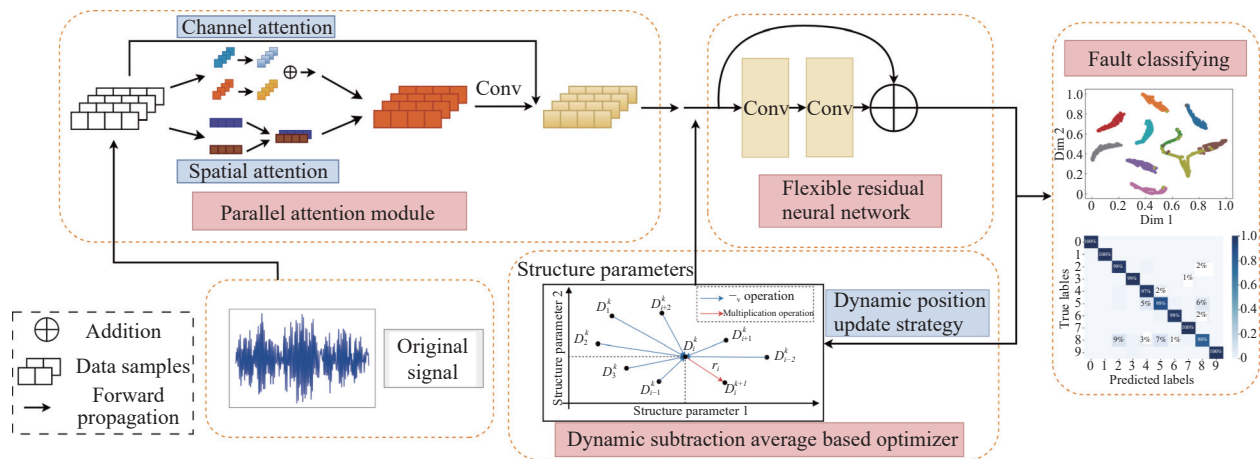


图1 DSABO-PAM-FResNet 故障诊断模型框架

Fig.1 DSABO-PAM-FResNet model framework for fault diagnosis

从图1中可以看出,基于DSABO-PAM-FResNet的滚动轴承故障诊断模型核心在于柔性残差神经网络、动态减法平均优化器和平行注意力模块三部分。

2.2 柔性残差神经网络 (FResNet)

随着CNN网络层数增加,可能会导致网络退化问题,而残差神经网络(Residual network, ResNet)可将较浅的堆叠层输出直接传递到较深的堆叠层,从而减轻网络退化。ResNet思想是在传统CNN基础上添加跳跃连接,通过在网络中添加跨越层的直接连接,将输入信号与输出信号相加。这种直接连接允许网络学习残差部分,即输入与输出之间的差异,而不是直接学习整个映射^[18]。

ResNet有利于缓解网络退化、减轻恒等映射学习难度,但其结构、参数难以确定,影响了特征提取能力。例如,当卷积层数和卷积核数过少时,网络的特征提取能力不足;而当卷积层数和卷积核数过多时,网络退化风险则会增加。此外,当跳跃连接数过少时,ResNet可能无法有效缓解网络退化;相反,当跳跃连接数过多时,又会导致过拟合的风险提高。对此,本文设计一种结构和参数可变的ResNet,也即柔性残差神经网络。FResNet的基本思想是在优化算法的迭代中改变柔性残差模块包含的卷积层数、卷积核数和跳跃连接数,找寻并配置使FResNet达到最优训练效果的结构参数。设置跳跃连接的卷积层由跳跃连接数 c 与卷积层数 a 决定,定义为:

$$C_t = \left\lceil \frac{a}{c} \times t + 0.5 \right\rceil, t = 1, 2, \dots, c \quad (7)$$

其中, C_t 表示设置跳跃连接的卷积层; $\lceil \cdot \rceil$ 表示取整函数。

图2示出了带有跳跃连接数 c 、卷积层数 a 和卷积核数 s 的柔性残差模块结构。如图2所示,当跳跃连接数 c 、卷积层数 a 或卷积核数 s 发生改变时,柔性残差模块能够自动改变自身结构和参数。因此,基于柔性残差模块构建的FResNet能够随着优化算法的迭代配置结构参数,从而搜寻最优结构。示例性地,FResNet与SABO的整合流程如图3所示。

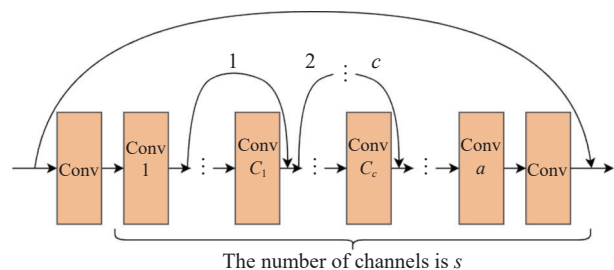


图2 柔性残差模块结构

Fig.2 Structure of the flexible residual module

首先,将柔性残差模块的结构参数作为优化算法的优化目标,将模型损失作为适应度函数。当优化算法迭代时,会将柔性残差模块需要的卷积层数、卷积核数、跳跃连接数等参数提供给FResNet,依据这些参数,建立包含特定结构的FResNet。最后依据模型损失判断结构优劣并继续优化参数,直到达到预期目标。FResNet既保留了ResNet减少网络退化的能力,还有利于克服ResNet由于配置不当而导致训练过拟合、准确率难以提升的问题,从而提升网络可靠性和泛化性能。

2.3 动态减法平均优化器 (DSABO)

SABO基于所有种群个体的算术平均位置更新位置,具有较好的全局搜寻能力。然而,在算法迭代后期,由于SABO的局部搜索能力较弱,使得

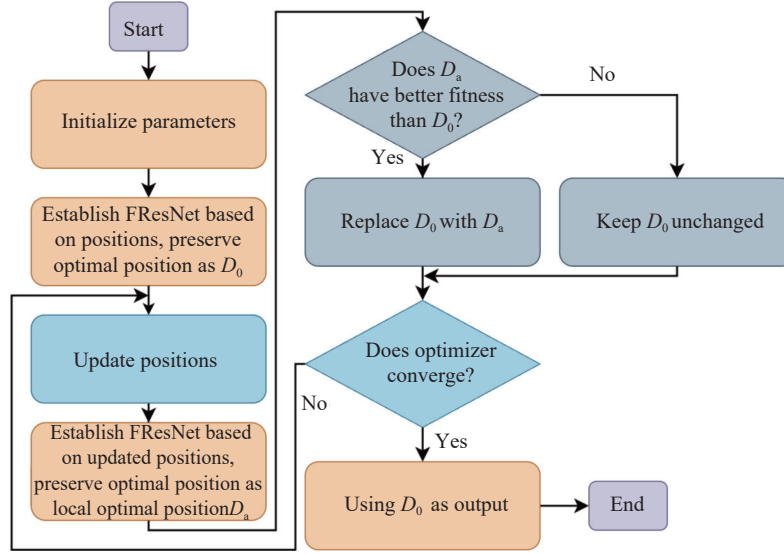


图3 SABO-FResNet 的整合流程

Fig.3 SABO and FResNet integration process

搜索结果难以快速接近理论最优值, 导致 SABO 收敛速度慢、收敛效果不好. 对此, 本文提出一种基于动态位置更新策略的动态减法平均优化器 (DSABO). 设当前迭代次数为 k , 总迭代次数为 K , 当 $\frac{k}{K} \leq 0.5$ 时, 位置更新策略的数学表达式如下:

$$\mathbf{DA}_i^k = \frac{k}{K} \times \mathbf{D}_0 + \left(1 - \frac{k}{K}\right) \frac{\mathbf{D}_i^k}{2} + r_i * \frac{1}{N} \sum_{j=1}^N (\mathbf{D}_i^k - \mathbf{D}_j^k) \quad (8)$$

$$\mathbf{D}_i^{k+1} = \begin{cases} \mathbf{DA}_i^k, & F(\mathbf{DA}_i^k) < F(\mathbf{D}_i^k) \\ \mathbf{D}_i^k, & F(\mathbf{DA}_i^k) \geq F(\mathbf{D}_i^k) \end{cases} \quad (9)$$

其中, $\mathbf{D}_0$ 表示全局最优位置; $\mathbf{DA}_i^k$ 表示基于第 k 次迭代中第 i 个个体基于当前位置和全局最优位置的理论更新位置.

当 $\frac{k}{K} > 0.5$ 时, 位置更新策略的数学表达式如下:

$$\mathbf{DB}_i^k = \mathbf{D}_0 + r_i * \frac{1}{N} \sum_{j=1}^N (\mathbf{D}_i^k - \mathbf{D}_j^k) \quad (10)$$

$$\mathbf{D}_i^{k+1} = \begin{cases} \mathbf{DN}_i^k, & \text{if } F(\mathbf{DN}_i^k) \text{ is } \min\{F(\mathbf{DN}_i^k), F(\mathbf{DB}_i^k), F(\mathbf{D}_i^k)\} \\ \mathbf{DB}_i^k, & \text{if } F(\mathbf{DB}_i^k) \text{ is } \min\{F(\mathbf{DN}_i^k), F(\mathbf{DB}_i^k), F(\mathbf{D}_i^k)\} \\ \mathbf{D}_i^k, & \text{if } F(\mathbf{D}_i^k) \text{ is } \min\{F(\mathbf{DN}_i^k), F(\mathbf{DB}_i^k), F(\mathbf{D}_i^k)\} \end{cases} \quad (11)$$

其中, $\mathbf{DB}_i^k$ 表示基于第 k 次迭代中第 i 个个体基于全局最优位置的理论更新位置; $\mathbf{DN}_i^k$ 由式 (5) 给出; $\min\{\cdot\}$ 表示取最小值函数.

在算法迭代初期 ($\frac{k}{K} \leq 0.5$), 种群个体的位置

更新受到当前位置和全局最优位置的共同影响: 当算法开始迭代时, 种群个体的位置更新受到当前位置的影响较大, 保证了算法的全局寻优能力; 随着算法迭代进行, 全局最优位置对位置更新的影响开始增大, 从而增强算法的局部寻优能力. 在算法迭代后期 ($\frac{k}{K} > 0.5$), 种群个体的位置更新由当前位置、基于当前位置的理论更新位置和基于全局最优位置的理论更新位置择优决定, 增加了种群多样性, 提高了优化算法的收敛速度和寻优能力.

2.4 平行注意力模块 (PAM)

在神经网络中, 注意力机制允许模型在处理输入数据时动态关注输入的不同部分, 而不是像传统神经网络一样对整个输入一视同仁, 这种机制将有助于提高模型性能. 受此启发, 本文设计了一种平行注意力模块 (PAM), 从通道和空间两个维度对输入数据提取依赖关系.

如图 4 所示, 输入数据经通道注意力子模块, 首先对输入数据在每一个通道上分别进行平均池化和最大池化, 从而得到长度与通道数相同的两组特征数据, 再将两组特征数据分别经由两个不同的卷积核数量为 1 的卷积层, 将两个卷积层的输出相加, 作为通道注意力子模块的输出权重.

同时, 如图 5 所示, 输入数据经空间注意力子模块, 首先对输入数据的每个区域分别进行最大池化和平均池化, 将分别得到的两组数据叠加成通道为 2 的一组数据, 经过一层卷积核为 1 的卷积层, 得到卷积层输出, 作为空间注意力子模块的输出权重.

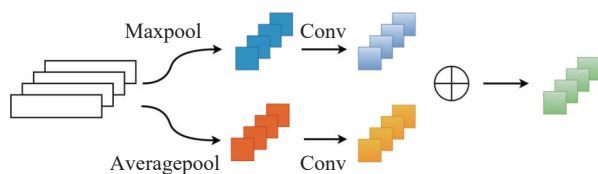


图4 PAM内部的通道注意力子模块

Fig.4 Channel attention submodule in the PAM

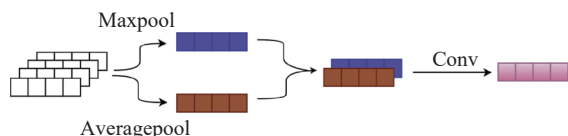


图 5 PAM 内部的空间注意力子模块

Fig.5 Spatial attention submodule in the PAM

最后将通道注意力子模块和空间注意力子模块的两组输出权重经由卷积层输出, 经由 sigmoid 激活函数转化到 $[0, 1]$ 的范围内, 与输入数据融合得到输出数据. PAM 流程如图 6 所示.

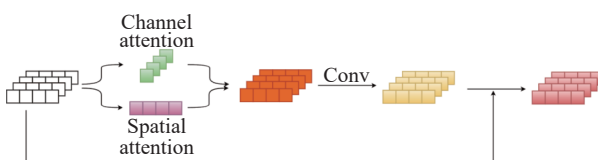


图 6 PAM 流程

Fig.6 PAM process

经由 PAM, 对输入数据分别实现了获取有用特征的通道和增强网络对数据中关键区域的空间感受能力。通过将 PAM 嵌入神经网络中, 逐层强化网络训练性能, 提高网络对特征的提取能力, 提高了网络的泛化性和适应性。

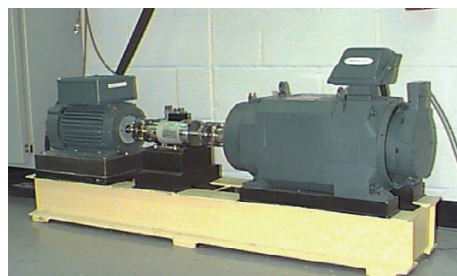


图 7 CWRU 数据集轴承故障试验台

Fig.7 Bearing fault test bed used in the CWRU dataset

验证选用型号 SKF6205 的深沟球轴承, 轴承故障为电火花加工单点故障. 故障直径分别为 0.18、0.36、0.54 mm, 轴承转速为 $1797 \text{ r} \cdot \text{min}^{-1}$, 在 12 kHz 采样频率下分别采集驱动端轴承的内圈、外圈和滚动体故障, 外加一种正常状态, 总共 10 种类别, 每种类别有 0、0.75、1.5、2.25 kW 四种负载状况.

样本长度取为 1024 个数据点, 训练集包含 600 个样本, 验证集包含 200 个样本, 测试集包含 200 个样本. 数据集描述如表 1 所示.

3.2 DSABO 的性能分析

为验证本文提出的 DSABO 的性能优越性,选择灰狼优化算法 (Grey wolf optimizer, GWO)^[20]、蝴蝶优化算法 (Butterfly optimization algorithm, BOA)^[21] 和鲸鱼优化算法 (Whale optimization algorithm, WOA)^[22] 与 DSABO 进行对比。考虑到各个优化算法的公平性,种群大小和最大迭代次数等基本参数设置相同。另外,为避免偶然性的影响,结果取重复运行 10 次的平均值。图 8 描述了在信噪比 $-6 \sim 0$ dB 环境下不同优化算法的故障诊断结果。

从图 8 中可以看出, 不同噪声环境下(信噪比 $-6 \sim 0$ dB), 基于 DSABO 的诊断模型的准确率均在 97% 以上, 高于基于 GWO、BOA 和 WOA 的诊断模型. 因此, 本文构建的 DSABO-PAM-FResNet 模型具有更优性能.

3.3 PAM 的性能分析

为验证本文提出的 PAM 相比其他注意力模块在特征提取能力上的优越性, 基于信噪比 $-6 \sim 0$ dB

表1 CWRU 数据集描述

Table 1 Description of selected data from the CWRU dataset[illegible]

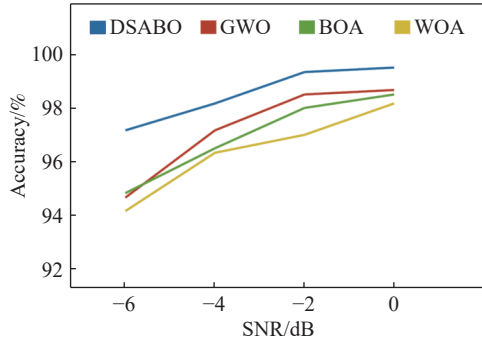


图 8 不同优化算法的故障诊断结果

Fig.8 Diagnosis results of DSABO, GWO, BOA, and WOA

环境, 使用 CWRU 数据集对本文模型进行测试. 对比注意力模块为高效通道注意力模块 (Efficient channel attention module, ECAM)^[23] 和空间注意力模块 (Spatial attention module, SAM)^[24]; 模型载体为 DSABO-FResNet. 取 10 次结果的平均值. 于是, 不同注意力模块的诊断结果如图 9 所示.

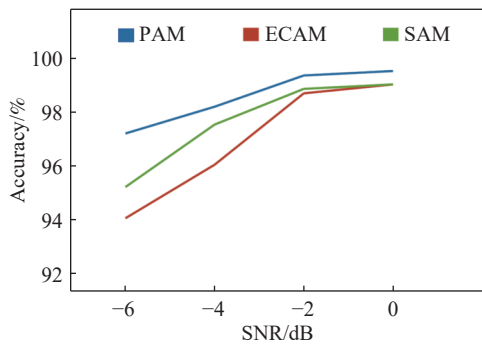


图 9 不同注意力模块的诊断结果

Fig.9 Diagnosis results of different attention modules

由图 9 可知, 不同噪声环境下 (信噪比 -6 ~ 0 dB) 基于 PAM 的诊断模型的准确率均在 97% 以上, 高于基于 ECAM 的诊断模型和基于 SAM 的诊断模型, 证明 PAM 模块更具故障特征提取能力.

3.4 强噪声环境下滚动轴承故障诊断

实际工业生产中, 滚动轴承的振动信号常受环境噪声影响. 为更准确评估所提模型性能, 将信噪比 -6 ~ 0 dB 的高斯白噪声添加到 CWRU 数据集中的原始振动信号. 验证中, 选择 JL-CNN^[25]、NCNN^[26]、Deep TL CNN^[27]、TCNN^[28] 用于对比分析:

(1) JL-CNN^[25]: 主要包括一个联合特征编码网络和两个基于注意力的编码器网络, 能够同时将故障诊断任务和信号去噪任务集成到 CNN 架构中;

(2) NCNN^[26]: 是一个结构复杂度低的稳定数据驱动模型, 采用批量归一化消除特征分布差异, 从而保证模型在不同工况下的泛化能力的前提;

(3) Deep TL CNN^[27]: 通过小波变换将原始传

感器数据转换为图像, 利用源域图片数据集进行预训练, 并在目标域图片数据集上微调模型参数;

(4) TCNN^[28]: 首先基于大型源任务数据集构建和预训练 CNN, 然后基于目标任务再次训练 CNN, 从而得到针对目标任务深度模型.

本文模型与以上四种先进模型进行对比验证, 所有验证均重复 10 次取平均值. 图 10 描述了不同噪声环境下的模型诊断结果.

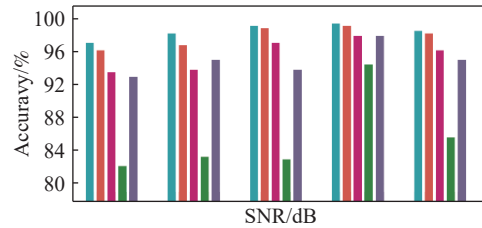


图 10 不同噪声环境下的模型诊断结果

Fig.10 Diagnosis results of models in different noise environments

由图 10 可知, 在信噪比 -6 ~ 0 dB 的噪声环境下, 本文提出的模型均有 97% 以上的准确率. 并且可以看出, 本文模型的平均准确率为 98.56%, 均高于其他模型. 特别地, 在信噪比为 -6 dB 的环境下本文模型仍有 97.18% 的准确率, 这说明本文模型有着较高的故障识别准确率和抗噪性能.

3.5 变工况条件下滚动轴承故障诊断

实际工业生产中, 滚动轴承的负载条件经常发生变化, 因此模型需要具有在不同负载下诊断故障的能力. 基于表 1, 选择 CWRU 轴承故障数据集中 0.75、1.5 和 2.25 kW 负载条件的数据进行变工况轴承故障诊断. 训练集包含 600 个样本, 验证集包含 200 个样本, 测试集包含 200 个样本.

由于实际生产中存在环境噪声对轴承故障诊断的干扰, 因此在原始信号中加入信噪比为 -4 dB 的高斯白噪声, 以模拟真实的工业环境. 选择 DF-CNN^[29]、HCNN^[30]、STIM^[31] 和 TSFFCNN^[32] 与本文模型进行比较:

(1) DF-CNN^[29]: 将原始信号转换为二维图像, 并通过包含两个 dropout 层和两个全连接层的 CNN 提取特征, 从而消除专家经验对特征提取过程的影响;

(2) HCNN^[30]: 是一种基于 CNN 的深度学习方法, 该方法使用了基于 ReLU 激活函数和高斯窗函

数的局部对比度归一化;

(3) STIM^[31]: 将一维振动信号转换为二维灰度图像, 从灰色图像中提取故障特征并实现故障分类;

(4) TSFFCNN^[32]: 建立了 1D-CNN 和 2D-CNN

并行的双通道网络模型, 并采用粒子群算法优化支持向量机以获得更高的精度。

本文模型与以上四种先进模型进行对比验证。所有验证重复 10 次, 诊断结果取平均值。图 11 描述了变工况条件下的模型诊断效果。

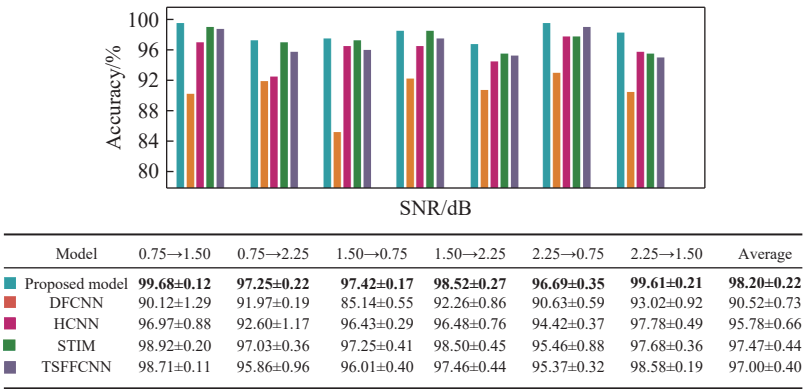


图 11 不同噪声环境下的模型诊断结果

Fig.11 Diagnosis results of models in different noise environments

在图 11 中, 0.75→1.5 表示将在负载 0.75 kW 下训练的模型迁移到负载为 1.5 kW 下进行故障诊断(训练集选择负载为 0.75 kW 的数据, 验证集和测试集选择负载为 1.5 kW 的数据)。从图 11 可以看出, 本文模型在数据集分布相似的情况下, 能够获得较好的识别效果, 例如 0.75→1.5 和 2.25→1.5 的准确率都在 99% 以上; 而在数据集分布差别较大时, 例如 2.25→0.75, 则识别效果略有下降, 但仍有 96% 以上的识别准确率。此外, 本文模型的平均准确率为 98.2%, 均高于其他对比模型, 说明本文模型有着较高的故障识别准确率和泛化能力。

4 结论

本文提出了一种基于动态减法平均优化器和并行注意力模块的柔性残差神经网络(DSABO-PAM-FResNet), 解决了传统轴承故障诊断网络特征提取不充分、诊断精度低的问题, 采用美国凯斯西储大学滚动轴承故障数据集对本文方法进行验证, 结论如下。

(1) 相比于 GWO、BOA 和 WOA, 本文提出的基于动态位置更新策略的 DSABO 具有更高的搜索精度。

(2) 相比于 ECAM 和 SAM, 本文提出的 PAM 能够进一步增强神经网络的特征提取能力, 进而提升故障诊断效果;

(3) 在不同噪声环境中, 本文模型准确率均高于其他智能诊断模型, 特别在信噪比-6 dB 环境下

仍有 97.18% 的准确率, 说明本文模型具有较好的抗噪能力。

(4) 在变工况条件下, 本文模型的平均准确率约为 98.2%, 准确率相比其他智能诊断模型平均高出 4.84%, 说明本文模型具有良好的变工况诊断适应能力。

需要指出的是, 本文提出的 DSABO-PAM-FResNet 模型在美国凯斯西储大学的滚动轴承故障数据集上具有较好的识别精度, 但在实际应用中仍然受到源域和目标域数据分布的限制。当源域和目标域的轴承数据集分布差异巨大时, 所提模型的轴承故障诊断精度将难以得到保证。因此, 未来的研究将集中于不同分布数据集上的迁移学习, 以确保滚动轴承在不同场景下的故障诊断效果。

参 考 文 献

[1] Chen C, Lu N Y, Jiang B, et al. Prediction interval estimation of aeroengine remaining useful life based on bidirectional long short-term memory network. *IEEE Trans Instrum Meas*, 2021, 70: 3527213

[2] Cai Z X, Dang Z, Lü Y, et al. Adaptive dynamic mode decomposition and GA-SVM with application to fault classification of planetary bearing. *Chin J Eng*, 2023, 45(9): 1559 (蔡志鑫, 党章, 吕勇, 等. 自适应动态模式分解和 GA-SVM 在行星轴承故障分类中的应用. *工程科学学报*, 2023, 45(9): 1559)

[3] Chen C, Shi J T, Shen M Q, et al. Pseudo-label guided sparse deep belief network learning method for fault diagnosis of radar critical components. *IEEE Trans Instrum Meas*, 2023, 72: 3510212

- [4] Xiao Y M, Shao H D, Han S Y, et al. Novel joint transfer network for unsupervised bearing fault diagnosis from simulation domain to experimental domain. *IEEE/ASME Trans Mechatron*, 2022, 27(6): 5254
- [5] Ou Y, Guo Z Y, Luo D L, et al. Collaborative air combat maneuvering decision-making method based on graph convolutional deep reinforcement learning. *Chin J Eng*, 2024, 46(7): 1227
(欧洋, 郭正玉, 罗德林, 等. 基于图卷积深度强化学习的协同空战机动决策方法. 工程科学学报, 2024, 46(7): 1227)
- [6] Chen C, Shi J T, Shen M Q, et al. A predictive maintenance strategy using deep learning quantile regression and kernel density estimation for failure prediction. *IEEE Trans Instrum Meas*, 2023, 72: 3506512
- [7] Chen C, Li X F, Shi J T. Research progress on remaining useful life interval prediction of equipment based on deep learning. *Chin J Eng*, 2024, 46(4): 723
(陈闯, 李先锋, 史建涛. 基于深度学习的装备剩余寿命区间预测研究进展. 工程科学学报, 2024, 46(4): 723)
- [8] Liu Z M, Cheng H Y, Mao M F, et al. Prediction of paste yield stress based on three-dimensional convolutional neural networks. *Chin J Eng*, 2024, 46(8): 1337
(刘泽民, 程海勇, 毛明发, 等. 基于 3D 卷积神经网络的膏体屈服应力预测. 工程科学学报, 2024, 46(8): 1337)
- [9] Chen C, Tao G Y, Shi J T, et al. A lithium-ion battery degradation prediction model with uncertainty quantification for its predictive maintenance. *IEEE Trans Ind Electron*, 2024, 71(4): 3650
- [10] Qing L B, Wu M F, Liu G, et al. Deep ADMM network in wavelet domain for image restoration. *Adv Eng Sci*, 2022, 54(5): 257
(卿颢波, 吴梦凡, 刘刚, 等. 基于小波域 ADMM 深度网络的图像复原算法. 工程科学与技术, 2022, 54(5): 257)
- [11] Xu Q S, Zhu B, Huo H B, et al. Fault diagnosis of rolling bearing based on online transfer convolutional neural network. *Appl Acoust*, 2022, 192: 108703
- [12] Sinitsin V, Ibryaeva O, Sakovskaya V, et al. Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN-MLP model. *Mech Syst Signal Process*, 2022, 180: 109454
- [13] Ruan D W, Wang J, Yan J P, et al. CNN parameter design based on fault signal analysis and its application in bearing fault diagnosis. *Adv Eng Inform*, 2023, 55: 101877
- [14] Ayas S, Ayas M S. A novel bearing fault diagnosis method using deep residual learning network. *Multimed Tools Appl*, 2022, 81(16): 22407
- [15] Zhang Y, Liu W Y, Wang X, et al. A novel wind turbine fault diagnosis method based on compressed sensing and DTL-CNN. *Renew Energy*, 2022, 194: 249
- [16] Kong X G, Mao G, Wang Q B, et al. A multi-ensemble method based on deep auto-encoders for fault diagnosis of rolling bearings. *Measurement*, 2020, 151: 107132
- [17] Trojovský P, Dehghani M. Subtraction-average-based optimizer: A new swarm-inspired metaheuristic algorithm for solving optimization problems. *Biomimetics (Basel)*, 2023, 8(2): 149
- [18] Xu X Z, Li C, Zhang X L, et al. A dense ResNet model with RGB input mapping for cross-domain mechanical fault diagnosis. *IEEE Instrum Meas Mag*, 2023, 26(2): 40
- [19] Chen C, Li X F, Shi J T. Hierarchical gray wolf optimizer-tuned flexible residual neural network with parallel attention module for bearing fault diagnosis. *IEEE Sens J*, 2024, 24(12): 19626
- [20] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer. *Adv Eng Softw*, 2014, 69: 46
- [21] Arora S, Singh S. Butterfly optimization algorithm: A novel approach for global optimization. *Soft Comput*, 2019, 23(3): 715
- [22] Chakraborty S, Sharma S, Saha A K, et al. A novel improved whale optimization algorithm to solve numerical optimization and real-world applications. *Artif Intell Rev*, 2022, 55(6): 4605
- [23] Singh P, Sharma A. Attention-based convolutional denoising autoencoder for two-lead ECG denoising and arrhythmia classification. *IEEE Trans Instrum Meas*, 2022, 71: 4007710
- [24] Cao J W, Feng Y M, Zheng R Z, et al. Two-stream attention 3-D deep network-based childhood epilepsy syndrome classification. *IEEE Trans Instrum Meas*, 2022, 72: 2503412
- [25] Wang H, Liu Z L, Peng D D, et al. Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising. *ISA Trans*, 2022, 128: 470
- [26] Zhao B, Zhang X M, Li H, et al. Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. *Knowl Based Syst*, 2020, 199: 105971
- [27] Shao S Y, McAleer S, Yan R Q, et al. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans Ind Inform*, 2019, 15(4): 2446
- [28] Chen Z Y, Gryllias K, Li W H. Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network. *IEEE Trans Ind Inform*, 2020, 16(1): 339
- [29] Zhang J Q, Sun Y, Guo L, et al. A new bearing fault diagnosis method based on modified convolutional neural networks. *Chin J Aeronaut*, 2020, 33(2): 439
- [30] Lu C, Wang Z Y, Zhou B. Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification. *Adv Eng Inform*, 2017, 32: 139
- [31] Zhao J, Yang S P, Li Q, et al. A new bearing fault diagnosis method based on signal-to-image mapping and convolutional neural network. *Measurement*, 2021, 176: 109088
- [32] Xue F, Zhang W M, Xue F, et al. A novel intelligent fault diagnosis method of rolling bearing based on two-stream feature fusion convolutional neural network. *Measurement*, 2021, 176: 109226