

基于 GADF-CWT-GCNN 的 滚动轴承故障诊断方法研究

张小丽, 罗鑫, 李敏, 梁旺, 王芳珍

(长安大学 道路施工技术与装备教育部重点实验室, 陕西 西安 710064)

摘 要:针对滚动轴承故障诊断在小样本环境下引起的模型泛化能力差、诊断精度低的问题,提出一种基于格拉姆角分场(GADF)和连续小波变化(continuous wavelet transform, CWT)与并行二维组归一化卷积神经网络(parallel convolutional neural network, P2D-GCNN)的滚动轴承故障诊断方法。对采集的数据进行预处理,采用格拉姆角场和连续小波变换将一维振动信号转换成二维图像作为模型输入,再选用数据增强技术扩充样本子图,满足网络输入要求,并将其导入搭建的组归一化卷积神经网络中进行诊断检测。结果表明:文中数据处理方法与搭建模型在小样本环境下泛化能力远高于 SVM 和 1D-CNN 等其他网络模型。为进一步验证模型在小样本数据下的识别能力,取数据集的 70%, 40% 和 20% 样本量进行多次实验,所对应的训练准确率及测试准确率分为 99.38%, 99.02%, 99.47%, 98.29%, 99.05%, 97.08%。结果证明,文中模型在小样本环境下对轴承故障诊断具有很高的准确率。

关 键 词:滚动轴承;故障诊断;格拉姆角分场(GADF);小波变换(CWT);并行二维卷积神经网络(P2D-GCNN)

中图分类号:TH133.33

文献标志码:A

文章编号:1000-2758(2024)05-0866-09

滚动轴承是工业生产中最常见的机械零件,它对机械的正常运作有着重要的作用,轴承一旦发生故障,极易造成重大的安全事故^[1]。因此对滚动轴承的有效故障诊断十分必要。传统的故障诊断方法有模态分析^[2]、小波变化^[3]等,这些方法虽然对故障诊断有着一定的效果,但太过于依赖人工经验诊断,存在特征提取弱、泛化能力不强等问题。

近年来,研究热点集中在深度学习网络和信号处理技术的结合来实现故障诊断。常见的神经网络有卷积神经网络^[4]、迁移神经网络^[5]等。信号处理技术有傅里叶变换^[6]、格拉姆角场变换^[7]、小波变换^[8]等。Chen 等^[9]对齿轮箱进行离散小波变化将一维信号转化为二维图像,再输入到卷积神经网络中进行故障分类识别。Zeng 等^[10]提出用 S 变换来

获取时频信息,然后基于 S 变换和卷积神经网络完成分类识别。虽然以上方法效果不错,但大部分均只是对信号进行时域或时频域分析,其信号处理效果不够理想,存在数据信息缺失,这使得故障诊断的准确性和稳定性较低。

轴承的工作环境复杂,工况多变,采集的轴承信号也大多有非线性特点。所以同时进行时域和时频域方法获取二维信息会更有利于深度学习的数据特征提取及训练。时域分析方法中,格拉姆角分场既能保留信号的完整信息,也保持着信号对时间的依赖性;时频分析方法中,连续小波变换具有很强的弱信号检测能力,相对于傅里叶变换不存在信息漏检和窗口固定不变的情况。综上所述,提出基于连续小波变换与格拉姆角分场变化和卷积神经网络的轴承故障诊断方法,通过不同模型对比和不同数据集规模下轴承故障诊断实验,验证了本文方法的有效性和可行性。

收稿日期:2023-09-05

基金项目:陕西省自然科学基金基础研究计划(2021JM-169, 2023-JC-YB-477)资助

作者简介:张小丽(1983—),博士研究生

通信作者:张小丽(1983—) e-mail: Lilyzhang@chd.edu.cn

1 相关原理介绍

1.1 格拉姆角分场变换

格拉姆角场变换包括格拉姆角和场 (GASF) 和格拉姆角分场 (GADF), 但格拉姆角分场转化处理的图像更为清晰, 具体变换步骤如下。

第 1 步 假设时间序列是 $\mathbf{X} = \{x^1, x^2, x^3, \dots, x^n\}$, 其中 $x^i (i = 1, 2, 3, \dots, n)$ 为第 i 个采样信号, n 为采样点数, 通过最大归一化和最小归一化, 将时间序列数据缩放到 $[0, 1]$, 表达式如 (1) 式所示

$$x_{\text{norm}}^i = \frac{x^i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

第 2 步 当处理的时间序列过长时, 生成的图像尺寸过大会给故障诊断过程带来负担。为了不丢失时间序列信息的前提下减少时间序列的长度, 一般采用分段聚合近似^[11]对缩放后的时间序列 $\mathbf{X}_{\text{norm}}$ 进行降维, 表达式如 (2) 式所示

$$x_{\text{norm}}^j = \frac{1}{m} \sum_{i=m(j-1)+1}^{m \times j} x_{\text{norm}}^i, j = 1, 2, \dots, w \quad (2)$$

式中: x_{norm}^j 为降维后的时间序列; $m = \frac{n}{w}$ 是压缩比, $w \leq n$ 。

第 3 步 用极坐标表示压缩时间序列, 其中, θ_j 是角余弦的极坐标, r_j 是极直角, t_j 是时间戳, N 是正则化极坐标系统跨度的常数因子, 表达式如 (3) ~ (4) 式所示

$$\theta_j = \arccos x_{\text{norm}}^j \quad (3)$$

$$r_j = \frac{t_j}{N} \quad (4)$$

通过上述变换, 可以将时间序列转化为对角对称的特征图, $\cos \theta$ 在 $[0, \pi]$ 上是单调递减的, 对于给定的时间序列, 它在极坐标系统中有唯一的对应值, 其逆映射也是唯一的。因此, 极坐标相对于笛卡尔坐标可以保持绝对的时间关系, 更有利于时间序列的分析。

第 4 步 GADF 根据不同时间点之间的角度差异识别出不同时间点的时间相关性, 采用公式 (5) 在矩阵的几何结构中进行编码

$$G_s = \mathbf{X}_{\text{norm}}^T \cdot \mathbf{X}_{\text{norm}} - \sqrt{\mathbf{I} - \mathbf{X}_{\text{norm}}^2}^T \cdot \sqrt{\mathbf{I} - \mathbf{X}_{\text{norm}}^2} \quad (5)$$

式中: $\mathbf{I}$ 是单位行向量 $[1, 1, 1, \dots, 1]$, $\mathbf{X}_{\text{norm}}^T$ 是 $\mathbf{X}_{\text{norm}}$ 的转置向量, 对一维时间序列映射到极坐标系统后, 将每个时间步的时间序列作为一维度量空间。通过上

述变换, 对于长度为 m 的输入时间序列, GADF 变换的合成矩阵维数为 $m \times m$, 原始信息保持正对角线, 其他时间序列之间关系反映在矩阵的其他区域。

1.2 连续小波变换

连续小波变换 (continuous wavelet transform, CWT) 通过使用一个随频率变化的“时频”窗口, 可以将原始振动信号的时频特征显示在图像上。设 $\psi(t) \in L^2(\mathbf{R})$, $\psi(t)$ 是一个母小波, 对小波函数进行拉伸、平移获得连续小波函数, 表达式为

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), a > 0, b \in \mathbf{R} \quad (6)$$

式中: a 是伸缩因子; b 是平移因子。

给定平方可积信号 $f(t)$, 当 $f(t) \in L^2(\mathbf{R})$ 时, $f(t)$ 的连续小波变换表达式为

$$C_f(a, b) = \frac{1}{\sqrt{a}} \int f(t) \psi'\left(\frac{t-b}{a}\right) dt = \int f(t) \psi'_{a,b}(t) dt = \langle f(t), \psi_{a,b}(t) \rangle \quad (7)$$

当 $C_f(a, b)$ 进行逆变换时, 可以得出

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \psi_{a,b}(t) C_x(a, b) \frac{1}{a^2} da db \quad (8)$$

式中, $C_\psi = 2\pi \int_{-\infty}^{\infty} \frac{|\psi(w)|}{|w|} dw$ 。

当 a 值较大时, 适合提取信号中的低频特征, 反之, 适合提取信号中的高频特征。小波变换的关键是小波基函数的选择, 它决定了小波变换的有效性和效率。对振动信号进行小波变换得出时频图, 本文选取 cmor 小波作为 CWT 的小波基函数。

2 并行二维卷积神经网络

2.1 P2D-CNN 模型提出及特点

本文对故障轴承一维振动信号同时进行时域、时频的二维图像转换后输入到并行二维卷积神经网络中, 便于网络模型获取更多丰富的故障特征信息, 为提高网络识别率和诊断精度, 需要对并行二维卷积神经网络进行参数设置和模型优化。

卷积神经网络的基本结构由卷积层、激活层、池化层和全连接层组成。卷积层使用多个卷积核扫描输入数据, 提取出各种特征。激活层引入非线性, 增强网络的表达能力。池化层通过降低特征图的空间维度, 减少计算量。全连接层则将前面的特征图展平成一维向量, 并通过加权和偏置进行最终分类。

本文模型采用 Relu 作为激活函数,以解决梯度消失问题并引入非线性。池化层使用最大池化,通过选择局部区域中的最大值减少特征图尺寸,减轻计算负担并保留关键特征。损失函数选择了交叉熵损失,有效衡量分类任务中的预测准确性。优化方面,使用了 Adam 优化器,结合了动量法和自适应学习率调整,能加快收敛速度并提高训练效果。

2.2 P2D-GCNN 模型提出及特点

在 CNN 网络训练过程中,网络参数被更新的同时数据分布随之更新,随着网络层数的加深,不可避免地影响了网络的收敛速度。因此,本文提出采用组归一化(group normalization, GN)算法进行数据的归一化处理,能加快网络训练,改进后的卷积神经网络称为并行二维组归一化卷积神经网络(P2D-GCNN),其结构如图 1 所示。

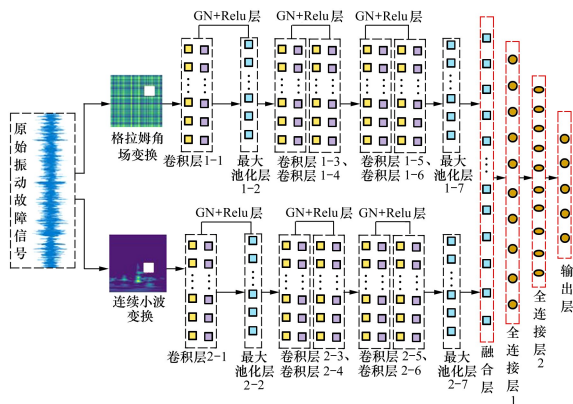


图 1 并行二维组归一化卷积神经网络模型结构图

组归一化算法即对数据进行归一化操作时,将通道 C 划分为若干 G 组,对每组分别计算均值和方差值完成归一化处理,最后再把归一化后的 G 组数据合并成 CHW,计算过程表达式为

$$S_i = \{k | k_N = i_N, \lfloor k_c/c \rfloor = \lfloor i_c/c \rfloor\} \quad (9)$$

式中: S_i 为平均值和方差的集合; G 是通道的分组数; C 是每一组的通道数; k_c 是关于 C 通道中 k 的子索引; i_c 是关于 C 通道中 i 的子索引,其中同一个通道索引的像素被一起归一化, $\lfloor k_c/c \rfloor = \lfloor i_c/c \rfloor$ 意味着子索引 k, i 是在相同通道中。

3 基于 P2D-GCNN 的滚动故障轴承诊断方法

本文通过对故障轴承一维振动信号同时进行格

拉姆角分场变换和连续小波变换,结合时域与时频域的互补故障特征信息挖掘到各特征信息内部关联性,结合卷积神经网络强大的自动特征提取能力,输入到 P2D-GCNN 网络中,提高网络识别率。

本文方法具体步骤为:

- 1) 从实验室获取不同故障状态的原始振动信号,用于后续模型的训练。
- 2) 对原始数据进行数据预处理,设置采样长度,利用格拉姆角分场变换和连续小波变换将一维振动信号转换成二维图像作为样本输入。
- 3) 利用数据增强技术,对时域样本图和时频样本图裁剪扩充样本数量,按照 7 : 1.5 : 1.5 比例划分训练集、验证集、测试集。
- 4) 搭建 P2D-GCNN 网络模型,初始化参数。
- 5) 将划分好的训练集输入到 P2D-GCNN 模型中训练,并使用优化算法迭代优化建模参数。如果达到最优值,则进行步骤 6); 否则,转到步骤 4),修改模型参数,直到达到最优参数,保留最优模型。
- 6) 将划分好的测试集样本输入到最优的 P2D-GCNN 模型中,最后由 softmax 输出分类结果和测试准确率。

4 故障诊断试验与分析

4.1 实验数据

在本文研究中,实验室使用的是滚轴故障模拟实验平台,该故障平台包括三相交流异步电机、柔性联轴节、变频控制器、 X 轴方向传感器、 Y 轴方向传感器、主轴、轴承座、不同故障种类的测试轴承、转子、加载装置、往复机构、皮带传送机构、DHDAS 动态信号采集分析系统和终端设备,见图 2。

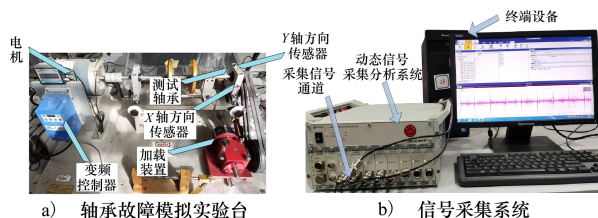


图 2 模拟实验台

在这个实验中,使用加速度传感器采集测试轴承的故障信号数据,变频器控制器被设置为固定的电机频率,并使用加载装置在 6 个工作条件下采集数据,电机负载设置为 0, 745.7, 1 481.4, 2 237.1,

2 982.8和 3 728.5 W,数据采集频率为 25 600 Hz。本文采用电机的运行频率为 30 Hz,选取加速度传

感器在 X 轴方向采集到的轴承故障数据,表 1 为实验室轴承故障统计信息。

表 1 实验轴承故障统计信息

标签	0	1	2	3	4
故障类型	正常轴承	滚动体故障	内圈故障	外圈故障	混合故障
工况/W	0~3 728.5	0~3 728.5	0~3 728.5	0~3 728.5	0~3 728.5

选取电机工作频率为 30 Hz 时,在 6 种不同负载下实验台采集正常轴承、滚动体故障轴承、内圈故障轴承、外圈故障轴承和混合故障轴承的振动信号数据,图 3 为负载 1 491.4 W 下的原始测试轴承 5 种不同故障状态的振动信号实验曲线图。

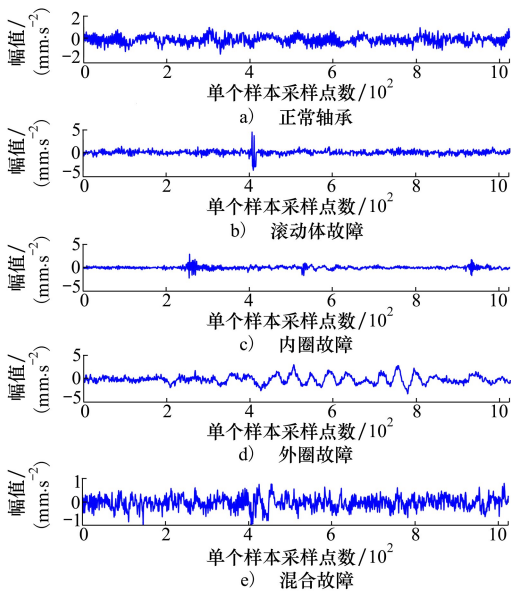


图 3 同负载下滚动体故障的原始信号图

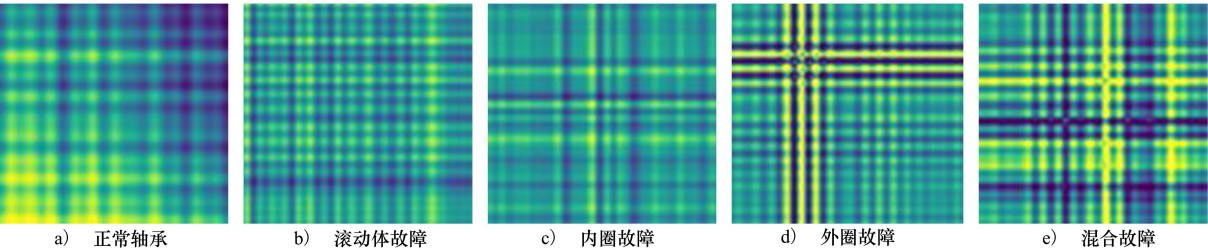


图 4 不同故障轴承的格拉姆角分场时域图像

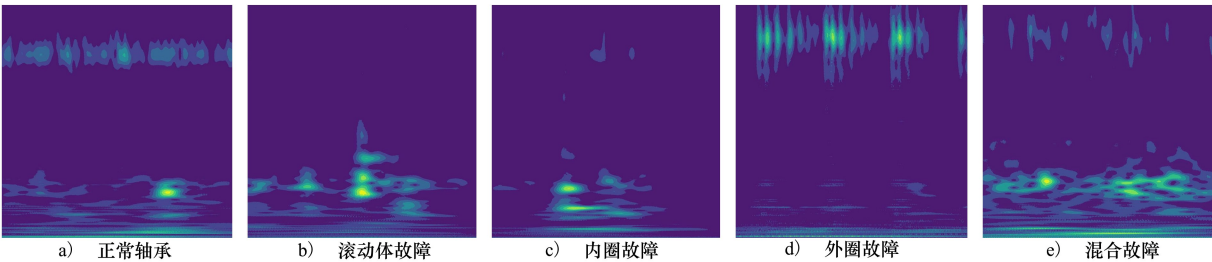


图 5 不同故障轴承的连续小波变换时域图像

进行上述样本扩充后,将裁剪后的 GADF 样本子图和 CWT 样本子图按照 7 : 1.5 : 1.5 的比例划分为训练集、测试集和验证集,表 2 为具体的样本划分结果,其中数据集 A 代表 0 W 下的样本、数据集

B~F 分别代表 745.7, 1 491.4, 2 237.1, 2 982.8, 3 278.5 W 下的样本集,这 6 个数据集的样本划分结果相同。

表 2 实验数据集 A 划分信息

故障状态	时域/时频转换	数据集 A			标签
		训练集	验证集	测试集	
正常	GADF/CWT	2 240/2 240	480/480	480/480	1
滚动体	GADF/CWT	2 240/2 240	480/480	480/480	2
内圈	GADF/CWT	2 240/2 240	480/480	480/480	3
外圈	GADF/CWT	2 240/2 240	480/480	480/480	4
混合	GADF/CWT	2 240/2 240	480/480	480/480	5

4.3 实验处理结果分析

本文仿真实验采用的硬件环境是:英特尔酷睿 i5-8250U 处理器,显卡芯片是 NVIDIA Geforce MX150, windows11 系统,python 编程语言,学习框架为 tensorflow2.1 版本。P2D-GCNN 模型在 tensorflow 中搭建完成后设置原始结构参数,将实验室各数据集输入到搭建好的 P2D-GCNN 模型中进行多次训练,确

定网络模型超参数。初始学习率为 0.001,激活函数选用 Relu 函数,选用 Adam 自适应优化器更新网络参数以获取参数最优值,dropout 值选取 0.5,选用 GN 归一算法。本次实验将各类数据集输入搭建好的 P2D-GCNN 网络模型中进行训练、测试,模型经过 50 轮训练后结果如图 6 所示。

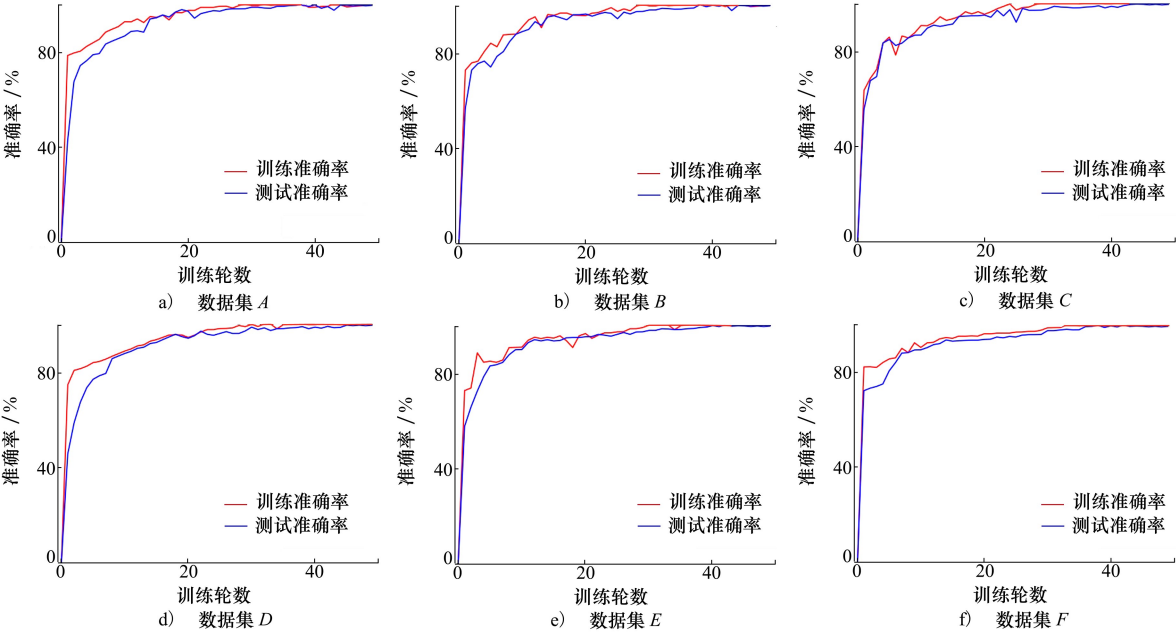


图 6 各数据集准确率与模型训练轮数的关系图

从图中可以看出,每个数据集的训练准确率和测试准确率随着训练轮数的增加而不断提升,当训练轮数到达 10,模型对数据集的训练准确率和验证准确率都开始大幅度提升,当模型训练了 20 轮,大部分数据集的训练准确率都在 90% 以上,当模

型训练到 50 轮时,6 个数据集的训练准确率和验证准确率都达到 99% 以上,接近 100%,这表明本文模型在轴承故障诊断中具有优异的故障识别能力。

4.4 不同数据集规模下模型的性能

在神经网络模型中,当输入的图片数量和尺寸

不同时,网络模型对故障诊断的效果也有所不同。当数据样本太少时,会因为网络模型中参数过多导致模型出现过拟合,因此网络模型在更少数据集情况下保持良好的诊断性能非常关键。为验证 P2D-GCNN 模型在小样本环境下的故障诊断能力,选取总数据集 50 个样本的 70%,40%和 20%,即 35 个样本、20 个样本和 10 个样本。为了满足网络输入数量的要求,对选取的 3 种小样本进行不同尺寸的裁剪。

第 1 种小样本环境是将 35 张 1 024×1 024 的样本图裁剪成 2 240 张 128×128 大小的子图;第 2 种小样本环境是将 20 张 1 024×1 024 的样本图裁剪成

5 120 张 64×64 的子图;第 3 种小样本环境是将 10 张 1 024×1 024 的样本图裁剪成 10 240 张 32×32 的子图,将这些子图按照 7 : 1.5 : 1.5 的比例分为训练集、测试集和验证集。与一维卷积神经网络模型(1D-CNN)^[12]、双输入 LeNet-5^[13]网络模型和多输入 PMCNN 模型^[14]这 3 种算法进行比较分析,1D-CNN和 PMCNN 模型的输入要求是一维数据输入,为满足神经网络训练的数据输入需求,将一维样本数据重复使用。其中,1D-CNN 模型进行一系列参数对比实验,确定性能最佳的 CNN 模型作为对比模型。为了避免实验偶然性对诊断结果的影响,实验进行 30 次并取平均值,结果如表 3 所示。

表 3 不同样本的训练集、测试集平均准确率

实验方法	35 个样本		20 个样本		10 个样本	
	训练准确率/%	测试准确率/%	训练准确率/%	测试准确率/%	训练准确率/%	测试准确率/%
PMCNN	96.29	94.77	93.01	91.86	88.48	86.66
1D-CNN	96.76	95.47	92.36	88.24	87.48	78.24
双输入 LeNet-5	96.96	94.58	94.78	93.45	90.67	89.58
P2D-GCNN	99.38	99.02	99.47	98.29	99.05	97.08

从表 3 可知,在 3 种不同的少样本数据规模下,本文所提出 P2D-GCNN 模型的平均识别精度最高。当数据规模为数据集的 70%时,本模型的训练识别精度为 99.38%,测试精度为 99.02%,明显高于 1D-CNN、PMCNN 和双输入 LeNet-5 模型。当数据量减少为总数据集的 40%时,其他 3 种方法的故障识别率明显下降。特别是当数据量为总数据集的 20%时,1D-CNN 的测试识别精度下降到 78.24%,是所有方法中效果最不佳的,比 70%数据集的测试精度下降了约 17%。而 PMCNN 和双输入 LeNet-5 也各

自下降了 8%和 5%左右,P2D-GCNN 模型仅仅下降了不到 3%左右,测试准确率能达到 97.08%。这足以表明当样本数量大幅度减少时,P2D-GCNN 模型仍能具有很高的故障识别率。

4.5 不同模型的识别性能比较分析

为了进一步说明 P2D-GCNN 模型对故障轴承的高识别性能,本节引入混淆矩阵对表 3 中 4 种方法在 10 个样本下的测试识别效果进行量化分析,如图 7 所示。

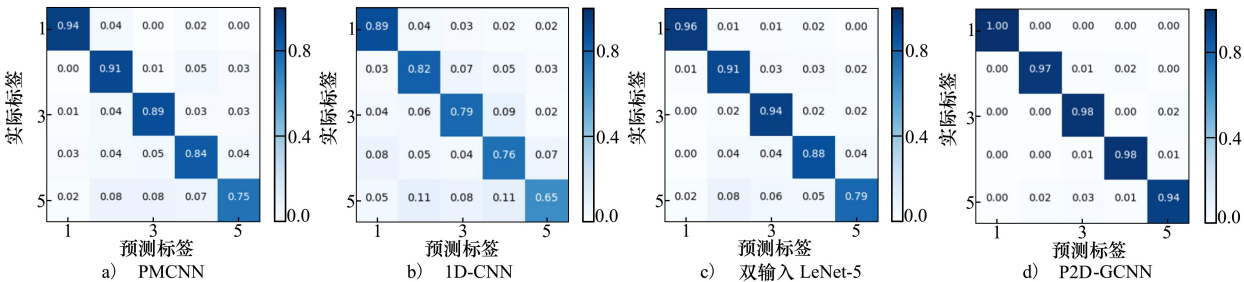


图 7 10 个样本下 4 种方法的混淆矩阵结果

从图中可以看出,使用 1D-CNN 模型的分类效果最差,部分内圈、外圈和混合故障被识别成其他故障种类。使用 PMCNN 模型的分类效果比 1D-CNN

稍微好点,接近 94%的正常轴承能被很好识别,但混合轴承故障的识别效果很差,超过 25%的混合故障被错误识别成其他 3 类故障。使用双输入 LeNet-

5 网络的分类效果较好,超过 90%的滚动体、内圈故障和正常轴承都能被正确识别,只有一部分外圈故障和混合故障的识别效果不理想。而本文的 P2D-GCNN 网络模型大部分故障都能很好地被正确分类,其中正常轴承分类结果完全准确为 100%,是 4 种网络里分类效果最好的,混合故障的轴承分类也高达 94%。

4.6 单输入和双输入在不同网络模型的结果对比

在验证了并行二维卷积网络模型的有效性后,分别设计单输入卷积网络模型(P2D-GCNN 分支网络)和双输入并行网络模型(P2D-GCNN),对比不同单输入图像和双输入图像对卷积神经网络模型准确率的影响,选取实验室自采数据 2 237.1 W 下的 50 个样本数据集进行实验验证。遵循控制变量的设计方法,一共设计 5 个数据集,分别是:数据集 A(仅包括 GADF 时域图像)、数据集 B(仅包括 CWT 时频图像)、数据集 C(2 个 GADF 时域图像)、数据集 D(2 个 CWT 时频图像)和数据集 E(GADF 时域图像和 CWT 时频图像),具体结构如图 8 所示。将数据集 A~E 输入搭建好的 P2D-GCNN 网络模型中进行训练和测试,实验进行 30 次后取各数据集测试集的准确率,实验结果如图 9 所示。

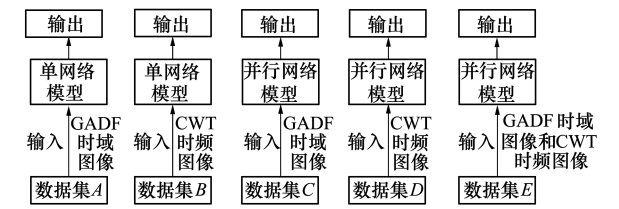


图 8 5 种不同网络模型结构简图

从图 9 可以看出,使用并行分支网络模型的测试准确率比单分支网络模型的测试准确率要高,说明多输入比单输入在网络模型中能获得更多特征信息,从而达到较高的识别分类效果。而 3 种不同双

输入的并行分支网络模型中,以 GADF 时域、CWT 时频图像输入的网络模型测试准确率是最高的,平均准确率达到 99.5%左右,在轴承故障诊断效果中具有较大的优势,该模型具有较好的稳定性和泛化能力。

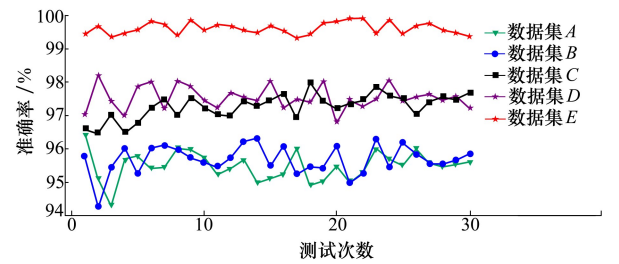


图 9 不同网络模型测试次数与准确率关系对比

5 结 论

提出一种基于 GADF-CWT-GCNN 的滚动轴承故障诊断方法,利用 GADF 和 CWT 分别对滚动轴承振动信号进行时域和时频变换,所得图像构建为数据集,将其输入 P2D-GCNN 网络中实现轴承的故障分类,实验结果表明:

1) 搭建的 P2D-GCNN 具有较好的性能,在 6 种不同负载条件下,数据集的训练准确率和验证准确率都能够达到 99% 以上。

2) 为了进一步验证模型在小样本数据下的识别能力,取数据集的 70%,40% 和 20% 样本量进行多次实验,结果证明,该模型在小样本环境下对轴承故障诊断具有很高的准确率。

3) 进行单输入和双输入在不同网络模型的实验,证明并行分支网络模型的测试准确率比单分支网络模型的测试准确率要高,多输入比单输入在网络模型中能获得更多特征信息,从而达到较高的识别分类效果。

参考文献:

[1] CUI L, WANG X, XU Y, et al. A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing[J]. Measurement, 2019, 135: 678-684

[2] LI Y, XU M, HUANG W, et al. An improved EMD method for fault diagnosis of rolling bearing[C] //2016 Prognostics and System Health Management Conference, Chengdu, 2016: 1-5

[3] LEE W, PARK C G. Double fault detection of cone-shaped redundant IMUs using wavelet transformation and EPSA[J]. Sensors, 2014, 14(2): 3428-3444

- [4] 吴春志, 江鹏程, 冯辅周, 等. 基于一维卷积神经网络的齿轮箱故障诊断[J]. 振动与冲击, 2018, 37(22): 51-56
WU Chunzhi, JIANG Pengcheng, FENG Fuzhou, et al. Gearbox fault diagnosis based on one-dimensional convolutional neural network[J]. Journal of Vibration and Shock, 2018, 37(22): 51-56 (in Chinese)
- [5] 雷亚国, 杨彬, 李乃鹏, 等. 跨设备的机械故障靶向迁移诊断方法[J]. 机械工程学报, 2022, 58(12): 1-9
LEI Yaguo, YANG Bin, LI Naipeng, et al. Cross-device targeted migration diagnosis method for mechanical faults[J]. Chinese Journal of Mechanical Engineering, 2022, 58(12): 1-9 (in Chinese)
- [6] 高博, 张忠, 魏龙, 等. 分数阶傅里叶变换改进算法在时频分析中的应用[J]. 强度与环境, 2021, 48(3): 8-15
GAO Bo, ZHANG Zhong, WEI Long, et al. Application of improved fractional Fourier transform algorithm in time-frequency analysis[J]. Strength & Environment, 2021, 48(3): 8-15 (in Chinese)
- [7] 骆家杭, 张旭, 汪靖翔. 基于格拉姆角场和多尺度 CNN 的轴承故障诊断[J]. 轴承, 2022(6): 73-78
LUO Jiahang, ZHANG Xu, WANG Jingxiang. Bearing fault diagnosis based on gram angle field and multi-scale CNN[J]. Bearing, 2022(6): 73-78 (in Chinese)
- [8] ANWARSHA A, NARENDIRANATH BABU T. Recent advancements of signal processing and artificial intelligence in the fault detection of rolling element bearings: a review[J]. Journal of Vibroengineering, 2022, 24(6): 1027-1055
- [9] CHEN R, HUANG X, YANG L, et al. Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform[J]. Computers in Industry, 2019, 106: 48-59
- [10] ZENG X, LIAO Y, LI W. Gearbox fault classification using S-transform and convolutional neural network[C]//2016 10th International Conference on Sensing Technology, 2016: 1-5
- [11] 李海林, 梁叶. 分段聚合近似和数值导数的动态时间弯曲方法[J]. 智能系统学报, 2016, 11(2): 249-256
LI Hailin, LIANG Ye. Dynamic time-bending method of piecewise aggregation approximation and numerical derivative[J]. Journal of Intelligent Systems, 2016, 11(2): 249-256 (in Chinese)
- [12] 王春雷, 路小娟. 一种基于深度学习的电机轴承故障诊断方法[J]. 兰州交通大学学报, 2020, 39(2): 43-50
WANG Chunlei, LU Xiaojuan. A method for motor bearing fault diagnosis based on deep learning[J]. Journal of Lanzhou Jiaotong University, 2020, 39(2): 43-50 (in Chinese)
- [13] 付忠广, 王诗云, 高玉才, 等. 双通道输入 LetNet-5 卷积神经网络旋转机械故障诊断模型研究[J]. 热力发电, 2023, 52(3): 81-87
FU Zhongguang, WANG Shiyun, GAO Yucai, et al. Research on fault diagnosis model of rotating machinery using LetNet-5 convolutional neural network with double channel input[J]. Thermal Power Generation, 2023, 52(3): 81-87 (in Chinese)
- [14] 周永庆. 基于一维并行多通道卷积神经网络的轴承故障研究[D]. 西安: 长安大学, 2021
ZHOU Yongqing. Bearing fault research based on one-dimensional parallel multi-channel convolutional neural network[D]. Xi'an: Chang'an University, 2021 (in Chinese)

A rolling bearing fault diagnosis method based on GADF-CWT-GCNN

ZHANG Xiaoli, LUO Xin, LI Min, LIANG Wang, WANG Fangzhen

(Key Laboratory of Road Construction Technology and Equipment for Ministry of Education,)
Chang'an University, Xi'an 710064, China

Abstract: Because of poor model generalization ability and low diagnostic accuracy caused by rolling bearing fault diagnosis in a small sample environment, a novel method based on the Gram angle division field (GADF), the continuous wavelet transform (CWT) and the parallel two-dimensional group normalization convolutional neural network (P2D-GCNN) for the fault diagnosis of rolling bearings is proposed. Firstly, collected data are preprocessed and one-dimensional vibration signals are converted into two-dimensional images by using the Gram angle division field and the continuous wavelet transform as the input of the model. Then the data enhancement technique is used to expand the sample sub-graph to meet the input requirements of the network. The sample sub-graph is imported into the convolutional neural network with the group normalization algorithm for diagnostic detection. The results show that the generalization ability of the data processing method and the model built in this paper in the small-sample environment is much higher than that of other network models such as the small vector machine and the 1D-CNN. In order to further verify the recognition ability of the model in the small sample environment, the sample sizes of 70%, 40% and 20% of the dataset are used to do experiments many times. Their corresponding training accuracy and test accuracy were 99.38%, 99.02%, 99.47%, 98.29%, 99.05% and 97.08% respectively, indicating that the model is highly accurate for the fault diagnosis of rolling bearings in the small sample environment.

Keywords: rolling bearings; fault diagnosis; Gram angular division field; continuous wavelet transform; parallel two-dimensional convolutional neural network

引用格式: 张小丽, 罗鑫, 李敏, 等. 基于 GADF-CWT-GCNN 的滚动轴承故障诊断方法研究[J]. 西北工业大学学报, 2024, 42(5): 866-874
ZHANG Xiaoli, LUO Xin, LI Min, et al. A rolling bearing fault diagnosis method based on GADF-CWT-GCNN[J]. Journal of Northwestern Polytechnical University, 2024, 42(5): 866-874 (in Chinese)