文章编号:1008-1542(2025)02-0129-12

基于 CWT-IDenseNet 的滚动轴承故障诊断方法

贾广飞[1],梁汉文[1],杨金秋[1],武　哲[1],韩雨欣[2]

(1.河北科技大学机械工程学院,河北石家庄　050018;
2.河北大学资源利用与环境保护研究中心,河北保定　071002)

摘　要:针对一维信号所含信息不全面和 DenseNet 网络在变工况下存在过拟合等问题,提出了基于连续小波变换时频图像和改进密集连接卷积网络(improved DenseNet, IDenseNet)的滚动轴承故障诊断方法 CWT-IDenseNet。首先,将一维振动信号通过 CWT 转为二维时频图像;其次,对 DenseNet 网络进行改进,将 DenseNet 第 1 个卷积块中的 ReLU 激活函数替换为 Swish 激活函数(Swish 激活函数更平滑);同时,在网络中引入基于风格的卷积神经网络重校准模块(style-based recalibration module,SRM)和空间与通道注意力机制模块(convolutional block attention module,CBAM),SRM 关注特征通道权重,CBAM 则从通道和空间 2 个维度增强特征表达能力,进而得到 IDenseNet;最后,将二维时频图像输入到 IDenseNet 模型中进行特征提取和故障诊断,通过模型的 Softmax 层输出故障诊断结果。结果表明,所提方法在恒定工况及变工况下的平均故障识别准确率均达到 97.80%,且在迁移学习模型中,平均故障识别准确率达到了 99.44%。CWT-IDenseNet 方法可以有效提高模型的泛化能力,在恒定工况及变工况下具有显著优势,对提高滚动轴承故障诊断的准确率和可靠性具有参考价值。

关键词:机械动力学与振动;滚动轴承故障诊断;连续小波变换;密集连接卷积网络;注意力机制

中图分类号:TH133.33　　**文献标识码**:A　　**DOI**:10.7535/hbkd.2025yx02002

Fault diagnosis method for rolling bearings based on CWT-IDenseNet

JIA Guangfei[1], LIANG Hanwen[1], YANG Jinqiu[1], WU Zhe[1], HAN Yuxin[2]

(1. School of Mechanical Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China;
2. Research Center of Resources Utilization and Environmental Conservation, Hebei University, Baoding, Hebei 071002, China)

Abstract: Aiming at the problems of incomplete information contained in one-dimensional signals and overfitting of the DenseNet under variable working conditions, a rolling bearing fault diagnosis method based on continuous wavelet transform (CWT) time-frequency images and an improved densely connected convolutional network (IDenseNet) was proposed. Firstly, the one-dimensional vibration signal was converted into two-dimensional time-frequency images by CWT. Then, the DenseNet network was turned into IDenseNet, the ReLU activation function in the first convolutional block of DenseNet was replaced by

收稿日期:2024-05-08;修回日期:2024-08-28;责任编辑:冯民
基金项目:国家自然科学基金(52206224);中央引导地方科技发展资金项目(226Z1906G);河北省教育厅科学研究项目(CXY2024038)
第一作者简介:贾广飞(1972—),男,河北献县人,副教授,博士,主要从事状态监测与智能诊断方面的研究。
E-mail:jiagf_11@163.com

贾广飞,梁汉文,杨金秋,等.基于 CWT-IDenseNet 的滚动轴承故障诊断方法[J].河北科技大学学报,2025,46(2):129-140.
JIA Guangfei, LIANG Hanwen, YANG Jinqiu, et al. Fault diagnosis method for rolling bearings based on CWT-IDenseNet[J]. Journal of Hebei University of Science and Technology,2025,46(2):129-140.

the Swish activation function (which was smoother), and the style-based recalibration module (SRM) and the convolutional block attention module (CBAM) were introduced into the DenseNet network. The SRM focused on the weight of feature channels, while CBAM enhanced the feature expression ability from the two dimensions of channel and space. Finally, the two-dimensional time-frequency image was input into the IDenseNet model for feature extraction and fault diagnosis, and the fault diagnosis results were output through the Softmax layer of the model. The results show that the average fault recognition accuracy of the proposed method reaches 97.80% under constant and variable conditions, and the average fault recognition accuracy reaches 99.44% in the transfer learning model. The CWT-IDenseNet method can effectively improve the generalization ability of the model, which has significant advantages under constant and variable conditions, providing reference for improving the accuracy and reliability of rolling bearing fault diagnosis.

Keywords: mechanical dynamics and vibration; fault diagnosis of rolling bearing; continuous wavelet transform; densely connected convolutional networks; attention mechanism

滚动轴承作为机械设备中重要的零部件之一,被广泛应用于火车、轮船、航空发动机等领域[1]。由于滚动轴承通常工作在复杂工况下,在长期运行后容易发生故障,甚至会导致安全事故的发生[2]。为避免滚动轴承发生故障后带来的危害,对其在发生故障前进行精准的故障诊断至关重要。

将一维振动信号输入到深度学习模型中,结合适当的神经网络结构进行处理和分析,是实现轴承故障诊断的常用方法。深度学习之所以在轴承故障诊断中被广泛应用,是因为它具有自动特征提取、适应性强、高准确率、实时性和可扩展性等优势。当前,许多学者基于振动信号对滚动轴承故障诊断进行了大量的研究[3-7],许多研究专注于一维时域信号,无法提供频域完整信息,导致处理效果不佳。相比之下,二维信号不仅包含时域和频域信息,还含有空间信息,能够提供更丰富和更全面的信息。赖奎等[8]提出了一种结合傅里叶变换、卷积神经网络(convolutional neural network,CNN)和注意力机制的方法,解决在复杂运行环境下配电终端采集模块故障识别困难的问题。刘鑫等[9]提出了一种基于灰度图像纹理分析的二维故障特征提取模型,可有效降低噪声干扰,简化计算。然而,傅里叶变换和灰度图等方法提取的特征不够丰富,无法捕捉到信号的所有重要信息,从而导致诊断准确度不高。

李振宇等[10]将原卷积神经网络里的激活函数 ReLU 替换为 Swish,提高了网络的识别准确率。ZHANG 等[11]将连续小波变换(continuous wavelet transform,CWT)与卷积神经网络相结合,采用 ADAM 优化器,提高了故障诊断的准确率并降低了处理时间。王焜等[12]提出了一种智能故障诊断神经网络模型,结合基于风格的卷积神经网络重校准模块(style-based recalibration module,SRM)和密集连接卷积网络(densely connected convolutional networks,DenseNet),解决了传统故障诊断中特征提取不足的问题。程建刚等[13]提出一种改进的 CNN 和空间与通道注意力机制(convolutional block attention module,CBAM)相结合的诊断方法,解决了传统机械故障诊断率低且人工参与程度高的问题。MA 等[14]提出一种将改进的生成对抗网络(generative adversarial network,GAN)和改进的密集连接卷积网络(IDenseNet)相结合的滚动轴承故障诊断方法,解决了由于样本数量有限导致的模型训练不足和故障类型诊断不及时的问题。上述方法在恒定工况下表现出较高的诊断准确率,但是在变工况下,这些方法的性能会受到影响,导致其诊断准确率下降。

为解决上述研究中一维信号所含信息不全面和 DenseNet 模型在变工况下过拟合的问题,提出了一种基于 CWT 时频图像和 IDenseNet 的滚动轴承故障诊断方法 CWT-IDenseNet,该方法替换了激活函数和引入注意力机制模块,通过不同工况的数据集进行实验验证。

1 改进的 DenseNet 网络

DenseNet 是一种深度神经网络,包含若干个 Dense Block 和过渡层(transition layer),每个 Dense Block 包含 L 层,每层输出与之前所有层的输出进行拼接。本文采用的 DenseNet-121 网络包含 121 层,其中有 4 个 Dense Block,每个 Block 分别包含 6、12、16 和 24 层,关于网络基本架构参见文献[15]。该网络在引入密集连接后,缓解了梯度消失问题,提高了信息传递效率,减轻了参数稀疏性[16]。

然而,DenseNet 存在过度拟合问题,并且在变工况下故障诊断效果不佳。为了解决这些问题,本文对 DenseNet 网络进行改进,提出 IDenseNet,具体改进包括:1)把第 1 个卷积块中的 ReLU 激活函数替换为 Swish 激活函数。Swish 激活函数是一种非线性激活函数,结合了 Sigmoid 的非线性和输入的原始信息。相

比 ReLU，它在某些实验中表现更优，能更好地缓解梯度消失问题，提升模型的表达能力和泛化性能[17]。2）在第 1 个卷积层后引入 SRM 注意力机制模块。SRM 主要通过动态调整特征通道权重，选择性增强或抑制不同通道的响应，提升特征表达能力[18]，从而在一定程度上缓解过拟合现象。3）在最后一个密集连接块后引入 CBAM 注意力机制模块。CBAM 可以同时从通道和空间 2 个维度增强特征表示能力，使模型能够更加准确地捕捉图像中的重要信息，提高特征表达能力和泛化能力。图 1 为通道与空间注意力机制模块，CBAM 通过联合学习通道注意力和空间注意力，自适应地提取输入数据中最具有代表性的特征，并将其输入到后续的网络层中进行处理。这种联合注意力机制使得 CBAM 模块能够自适应地提取最具有代表性的通道和空间特征，从而增强模型的表示能力[19]。

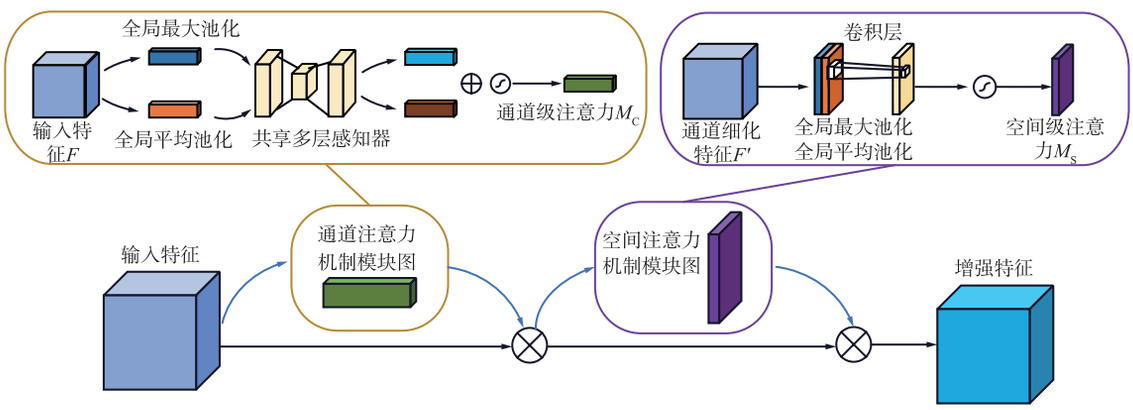


图 1　通道与空间注意力机制模块

Fig. 1　Channel attention and spatial attention module

综上所述，本文提出的 IDenseNet 通过将 ReLU 激活函数替换为 Swish 激活函数、引入 SRM 和 CBAM 注意力机制，有效提高了模型的表达能力和泛化性能，解决了 DenseNet 变工况下过拟合的问题。

2　CWT-IDenseNet 滚动轴承故障诊断方法

2.1　CWT-IDenseNet 故障诊断模型

图 2 为本文 CWT-IDenseNet 故障诊断网络模型。第一，利用 CWT 中的 Morlet 小波对一维信号进行

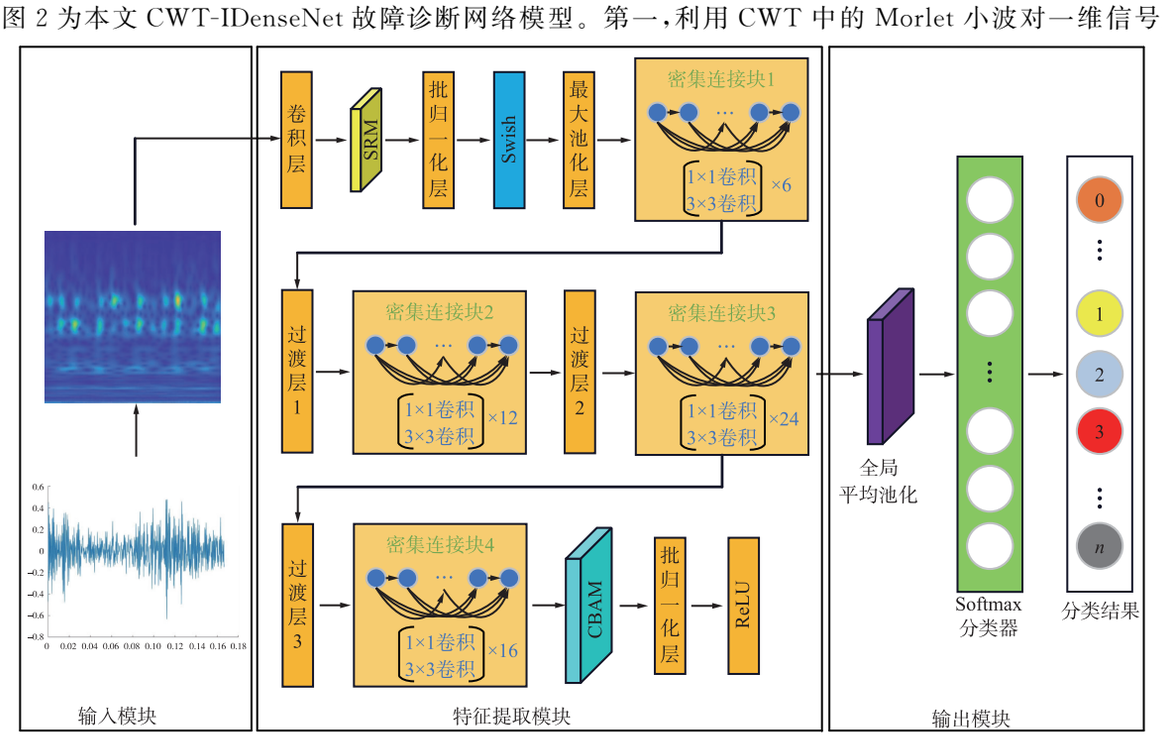


图 2　CWT-IDenseNet 故障诊断网络模型

Fig. 2　CWT-IDenseNet fault diagnosis network model

处理，生成二维时频图像。这种小波具有频率和时间平衡以及良好的局部化特性和带通特性等优点，在信号处理、图像处理和模式识别领域被广泛应用。第二，输入卷积核大小为7×7的卷积块，卷积块包含卷积层、SRM注意力机制模块、批归一化层、Swish激活函数和最大池化层。第三，把第1个卷积块中的ReLU激活函数替换为Swish激活函数，促进梯度流动，增强非线性建模能力。第四，在第1个卷积层后引入SRM注意力机制，SRM可以帮助模型更好地关注特征图中的重要区域特征和上下文信息，关注特征通道权重，提高特征表示能力和泛化能力，从而在一定程度上缓解过拟合现象。密集连接块和过渡层促进了特征的传播和重复利用，控制了信息流量，提高了网络的鲁棒性和稳定性。第五，在密集连接块4后引入CBAM注意力机制模块，CBAM可以同时从通道和空间2个维度增强特征表示能力，使模型能够更加准确地捕捉图像中的重要信息，提高特征表达能力和泛化能力。经过批归一化层和ReLU激活函数，加速了网络训练，提高了模型性能。第六，通过Softmax分类器进行故障诊断。

通过引入Swish激活函数、SRM和CBAM注意力机制模块，有效解决DenseNet网络的过拟合问题；同时，采用连续小波变换生成时频图，提高变工况下的故障诊断效果。

2.2 CWT-IDenseNet故障诊断方法流程

CWT提取的特征可以帮助识别图像中的结构和模式，而CWT图相比于灰度图、包络谱图和倒频谱图能够更准确地捕获信号在时间和频率上的局部变化，提供更精确的特征表示。连续小波变换过程如图3所示。

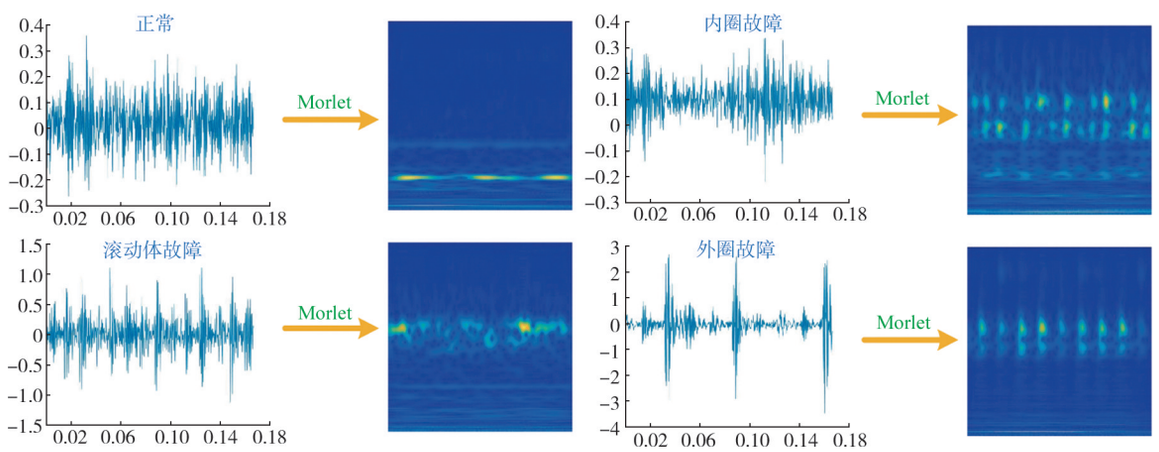


图3 Morlet小波变换生成时频图

Fig. 3 Morlet wavelet transform to generate time-frequency diagram

本文方法故障诊断流程如图4所示。具体的诊断步骤如下。

1）对原始振动信号进行分割，每个片段的长度设定为1 024，并采用间隔步长和重叠采样的方式进行数据增强，之后采用CWT中Morlet小波方法将一维信号转变为二维时频图像，生成图片的大小设置为224×224。

2）将得到的CWT二维时频图像像素大小按照第1步生成的224×224输入到网络中，采用随机翻转和随机旋转技术对数据进行增强，并进行归一化处理，将数据按照8∶2的比例划分。

3）使用IDenseNet模型对训练集进行训练，直至满足训练要求，并保存表现最佳的模型。

4）利用已训练好的模型对测试集进行分类识别，得出故障诊断结果。

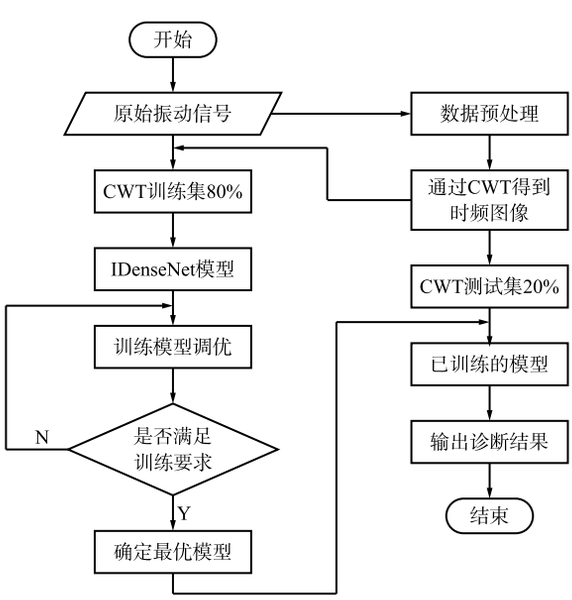


图4 CWT-IDenseNet故障诊断流程

Fig. 4 CWT-IDenseNet fault diagnosis process

3　实验验证与分析

3.1　实验说明及数据预处理

第 1 组采用凯斯西储大学轴承数据集[20]进行实验，选取电动机转速为 1 772、1 750 和 1 730 r/min，电动机载荷选取为 $A=0.746$ kW、$B=1.491$ kW 和 $C=2.237$ kW，驱动端轴承故障振动信号的数据采样频率为 12 kHz。故障包括内圈、外圈和滚动体 3 类，分别对应损伤尺寸为 0.18、0.36 和 0.54 mm。首先，在恒定工况载荷 $A=0.746$ kW 下验证本文方法的优越性，然后，利用 3 种不同工况载荷下的数据作为训练集和测试集，验证本文方法的泛化性能。振动信号是一维时间序列，每个样本选取连续的 1 024 个数据点，并使用 CWT 中的 Morlet 小波将其转换成大小为 224×224 的二维图像。每种状态生成 2 000 张二维图像，共有 10 种状态，因此共生成 20 000 张二维图像。根据 8∶2 的比例划分为训练集和测试集，具体实验数据集见表 1。

表 1　第 1 组实验数据集展示

Tab.1　Dataset display of the first experiment

故障尺类型	损伤尺寸/mm	故障标签	载荷 A		载荷 B		载荷 C	
			训练样本数	测试样本数	训练样本数	测试样本数	训练样本数	测试样本数
滚动体故障	0.177 8	1	1 600	400	1 600	400	1 600	400
	0.355 6	2	1 600	400	1 600	400	1 600	400
	0.533 4	3	1 600	400	1 600	400	1 600	400
内圈故障	0.177 8	4	1 600	400	1 600	400	1 600	400
	0.355 6	5	1 600	400	1 600	400	1 600	400
	0.533 4	6	1 600	400	1 600	400	1 600	400
外圈故障	0.177 8	7	1 600	400	1 600	400	1 600	400
	0.355 6	8	1 600	400	1 600	400	1 600	400
	0.533 4	9	1 600	400	1 600	400	1 600	400
正常	0	10	1 600	400	1 600	400	1 600	400

第 2 组采用东南大学齿轮箱中轴承数据集[21]进行实验，选取转速负载为 $D=20$ Hz－0 V 和 $E=30$ Hz－2 V。该数据集中包括 4 种轴承故障状态和 1 种健康状态，故障状态分别为滚动体故障、内圈故障、外圈故障和内圈外圈复合故障。首先，在恒定工况负载 $D=20$ Hz－0 V 下验证所提方法的优越性，然后，利用 2 种不同工况负载下的数据作为训练集和测试集，验证本文方法的泛化性能。振动信号是一维时间序列，每个样本选取连续 1 024 个数据点，并使用 CWT 中的 Morlet 小波将其转换成大小为 224×224 的二维图像。每种状态生成 2 000 张二维图像，共有 5 种状态，因此共生成 10 000 张二维图像。根据 8∶2 的比例划分为训练集和测试集，具体实验数据集见表 2。

表 2　第 2 组实验数据集展示

Tab.2　Dataset display of the second experiment

故障尺类型	故障标签	负载 D		负载 E	
		训练样本数	测试样本数	训练样本数	测试样本数
滚动体故障	1	1 600	400	1 600	400
内圈外圈复合故障	2	1 600	400	1 600	400
正常	3	1 600	400	1 600	400
内圈故障	4	1 600	400	1 600	400
外圈故障	5	1 600	400	1 600	400

3.2　恒定工况下的故障诊断分析

首先，采用凯斯西储大学轴承数据集进行实验，故障诊断模型在恒定工况载荷 $A=0.746$ kW 下通过迭代 50 次得到了在训练集和测试集上的准确率，如图 5 所示。通过曲线图可以看出，训练集一直处于上升阶段，经过 19 次迭代训练后，训练集趋于平稳，在训练集上的准确率达 100%。测试集曲线在前 14 次迭代过程中一直处于震荡状态，在第 6 次迭代时振幅最大，经过 14 次迭代后，测试集曲线的准确率稳定在 99.98%；同时，训练集和测试集的准确率变化基本一致，都保持在较高水平。这表明模型的诊断效果良好，没有出现过拟合等问题。相应的通过迭代 50 次得到训练集和测试集上的损失值，如图 6 所示，训练集的损失值一直处于下降状态，直到迭代 19 次之后，训练集损失值趋于 0。测试集在前 14 次迭代过程中一直处于震荡状态，第 6 次和第 10 次迭代时振幅较为明显，经过 14 次迭代后测试集曲线趋于平稳，损失值为 0.000 2。

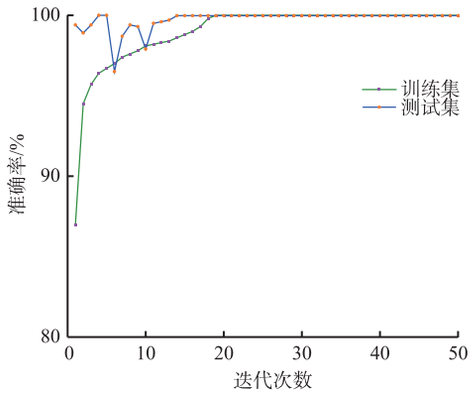


图5 故障诊断模型准确率变化曲线

Fig. 5 Accuracy change curve of fault diagnosis model

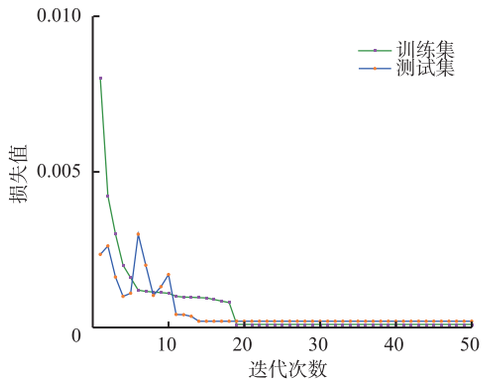


图6 故障诊断模型损失值变化曲线

Fig. 6 Loss rate curve of fault diagnosis model

将测试集训练过程中最佳权重保存下来，使用测试集的最佳权重对已经训练好的模型进行故障分类，结果见图7。从图中可以观察到，除了第3类标签的故障，其余准确率均达到了100%；而对于标签为3的故障，有1个样本被错误地分类为其他类型的故障。总体而言，在4 000个测试样本中，有3 999个样本被正确分类，准确率高达99.98%。

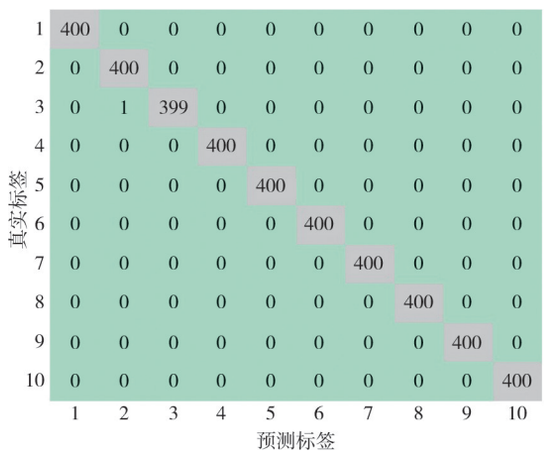


图7 滚动轴承故障分类结果

Fig. 7 Rolling bearing fault classification results

为了验证本文所提故障诊断方法的优越性，采用凯斯西储大学电动机载荷为$A=0.746$ kW下的数据集进行对比实验，使用VGG、ResNet、DenseNet和IDenseNet模型，分别与CWT相结合来进行对比试验。每组模型进行了50次迭代训练，得到4种方法的准确率如图8所示。可以看出4种方法刚开始时都有振幅，CWT-VGG在第4次迭代时振幅最大，本文方法收敛速度最快，故障诊断准确率可达99.98%，高于其他3种方法的故障准确率。4种方法的损失值如图9所示，可以看出本文方法的损失率最低，损失值趋近于0。

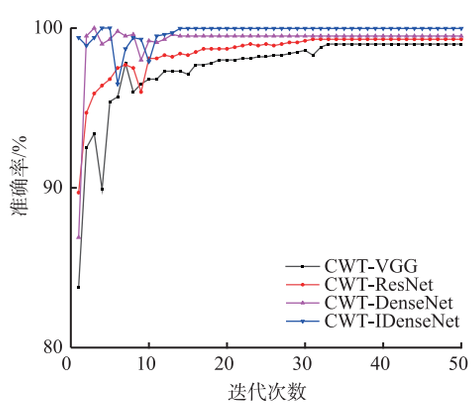


图8 不同方法准确率

Fig. 8 Accuracy of different methods

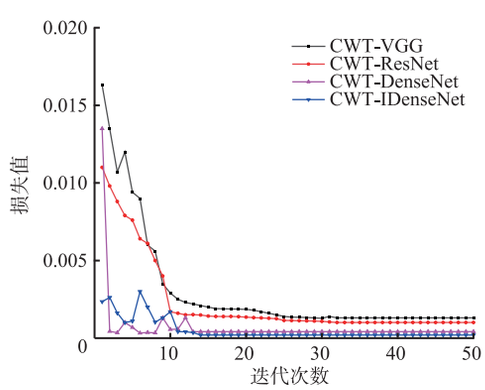


图9 不同方法损失值

Fig. 9 Loss value of different methods

其次，采用东南大学齿轮箱中轴承数据集进行实验，故障诊断模型在恒定工况负载$D=20$ Hz—0 V下，经过50次迭代后，在训练集和测试集上的准确率如图10所示。通过曲线图可以看出，训练集一直处于上升阶段，经过22次迭代训练后，训练集趋于平稳，在训练集上的准确率达100%。测试集曲线在前11次迭代过程中一直处于震荡状态，在第4次迭代时振幅最大，经过23次迭代，测试集曲线的准确率稳定在99.90%；同时，训练集和测试集的准确率变化基本一致，都保持在较高水平。这表明模型的诊断效果良好，没有出现过拟合等问题。相应地，通过迭代50次得到训练集和测试集上的损失值，如图11所示。训练集的损失值一直处于下

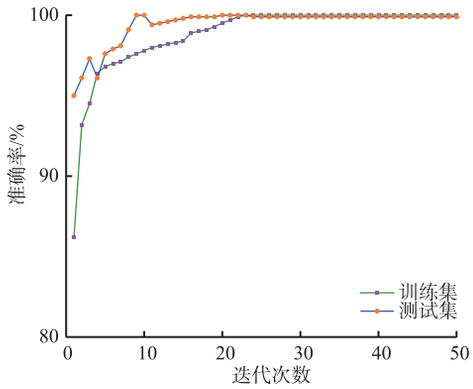


图 10　故障诊断模型准确率变化曲线

Fig. 10　Accuracy change curve of fault diagnosis model

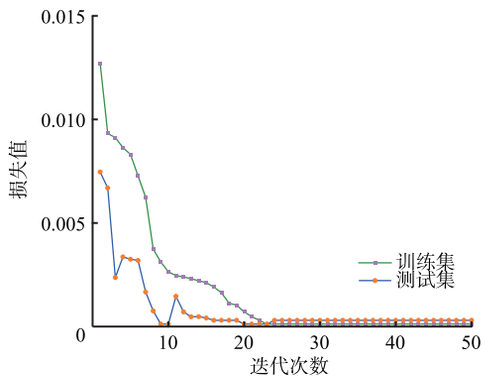


图 11　故障诊断模型损失值变化曲线

Fig. 11　Loss rate curve of fault diagnosis model

降状态，直到迭代 22 次后，训练集损失值趋于 0。测试集曲线在前 11 次迭代过程中一直处于震荡状态，第 3 次和第 11 次迭代时振幅较为明显，经过 23 次迭代后测试集曲线趋于平稳，损失值为 0.000 3。

将测试集训练过程中最佳权重保存下来，使用测试集的最佳权重对已经训练好的模型进行故障分类，结果见图 12。从图中可以观察到，除了第 4 类标签的故障，其余准确率均达到了 100%；而对于标签为 4 的故障，有 2 个样本被错误地分类为其他类型的故障。总体而言，在 2 000 个测试样本中，有 1 998 个样本被正确分类，准确率高达 99.90%。

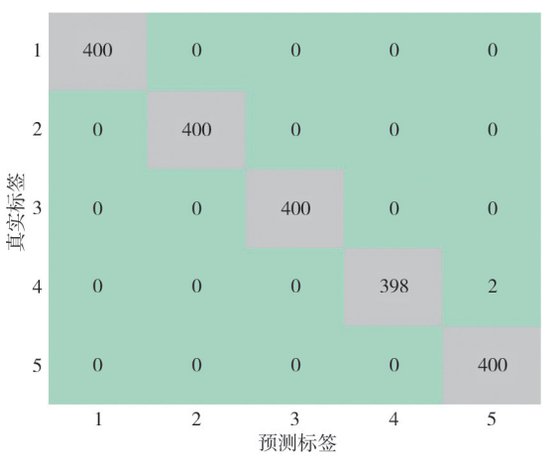


图 12　滚动轴承故障分类结果

Fig. 12　Rolling bearing fault classification results

为了验证本文方法的优越性，采用东南大学转速负载为 $D=20$ Hz－0 V 下的数据集进行对比试验，使用 VGG、ResNet、DenseNet 和 IDenseNet 模型，分别与 CWT 相结合来进行对比实验。每组模型进行了 50 次迭代训练，得到 4 种方法测试集的平均故障识别准确率，如图 13 所示。可以看出 4 个模型测试集的平均故障识别准确率均在 93%以上。其中，ResNet 网络测试集的平均故障识别准确率最低，为 93.56%。本文方法在测试集的平均故障识别准确率最高，为 99.44%。

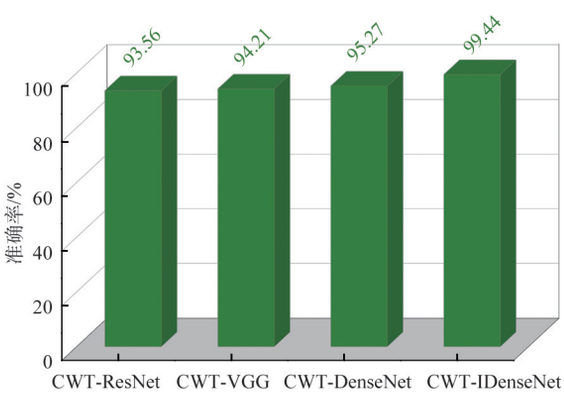


图 13　4 种方法测试集平均故障识别准确率

Fig. 13　Average fault accuracy of the four methods test set

3.3　变工况下的故障诊断分析

为了进一步证明本文方法的优越性，采用 2 种不同的实验数据进行验证。首先，使用凯斯西储大学轴承数据集进行实验验证，并选择采样频率为 12 kHz 的驱动端轴承故障振动信号数据，电动机载荷选取 $A=0.746$ kW、$B=1.491$ kW 和 $C=2.237$ kW 3 种不同工况载荷下的数据作为训练集和测试集。选取 CWT、灰度图、包络谱图和倒频谱图 4 种方法生成二维时频图像，图像大小为 224×224，将 4 种不同图像分别与 VGG、ResNet、DenseNet 和 IDenseNet 4 种网络相结合的方法作为对比实验。其中，灰度图是根据文献[22]的方法生成的，主要是将时域信号转化为二维灰度图像，其中包络谱图通过包络解调方法生成，而倒频谱图则采用倒频谱方法生成。

不同方法的轴承故障诊断结果展示了在恒定工况载荷和变工况载荷下的故障识别准确率，具体结果见表 3，其中 $A\rightarrow B$ 表示以数据集 A（电动机载荷为 0.746 kW）作为训练集，数据集 B（电动机载荷为 1.491 kW）作

表3 不同方法故障诊断结果对比

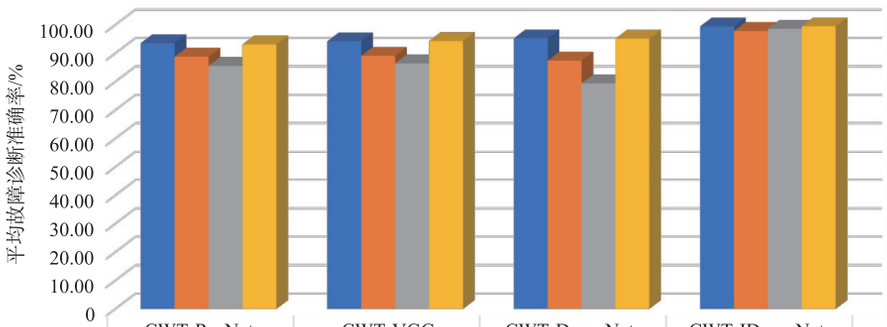
Tab.3 Comparison of fault diagnosis results of different methods

不同方法	输入图像	准确率/%									
		$A\rightarrow A$	$A\rightarrow B$	$A\rightarrow C$	$B\rightarrow A$	$B\rightarrow B$	$B\rightarrow C$	$C\rightarrow A$	$C\rightarrow B$	$C\rightarrow C$	平均
VGG	灰度图	93.18	71.34	69.57	76.19	93.22	79.44	67.58	79.64	93.96	80.46
	包络谱图	94.32	78.58	76.18	81.64	94.64	84.22	76.18	84.66	94.37	84.98
	倒频谱图	96.15	79.64	81.28	78.64	95.16	85.65	71.33	86.49	96.54	85.65
	CWT图	97.46	89.53	81.67	83.56	97.32	85.79	83.26	81.68	97.51	88.64
ResNet	灰度图	94.46	75.56	73.94	79.57	94.44	85.93	75.19	82.38	93.42	83.88
	包络谱图	94.63	84.23	81.68	88.34	95.74	86.15	72.19	84.76	94.51	86.91
	倒频谱图	97.02	90.04	84.51	82.64	95.23	89.67	81.34	92.96	97.12	90.06
	CWT图	98.42	91.82	86.42	88.57	98.53	91.26	87.64	88.94	98.31	92.21
DenseNet	灰度图	96.31	76.54	74.32	79.63	95.46	79.64	81.67	82.04	95.94	84.62
	包络谱图	96.82	83.36	76.56	86.75	97.79	82.85	74.68	84.36	97.28	86.72
	倒频谱图	98.23	87.16	85.64	88.43	98.12	92.46	84.67	90.81	97.22	91.42
	CWT图	99.21	95.03	86.34	91.14	99.18	94.35	92.12	93.25	99.12	94.42
IDenseNet	灰度图	97.23	78.16	76.82	83.22	96.36	81.12	82.86	82.21	97.01	86.11
	包络谱图	97.74	86.03	79.16	88.94	98.24	86.83	79.45	87.36	97.76	89.06
	倒频谱图	98.94	94.12	88.45	89.86	98.54	97.01	96.36	94.96	97.74	95.11
	CWT图	99.75	99.73	98.85	98.52	99.64	98.82	99.46	98.56	99.70	99.23

为测试集，其余含义以此类推。每组实验迭代50次，计算测试集的平均故障识别准确率。然后每组实验重复5次，最后取5次结果的平均值作为最终的平均故障识别准确率，以确保实验结果的准确性。分析表3中的数据可以发现，所有方法在恒定工况载荷（$A\rightarrow A$、$B\rightarrow B$和$C\rightarrow C$）的情况下都达到了较好的平均故障识别准确率。在工况载荷为$A\rightarrow A$不变时，IDenseNet的平均故障识别准确率达到了99.75%，高于DenseNet的99.21%。这表明，在相同工况下，IDenseNet展现了更优越的诊断能力。当涉及变工况载荷时（如$A\rightarrow C$、$B\rightarrow C$等），一些模型的表现有所下降。特别是DenseNet在工况载荷为$A\rightarrow C$时，其平均故障识别准确率下降至86.34%，表现出过拟合问题，表明其在变工况下的泛化能力不足。然而本文方法在类似的变工况载荷条件下（如$A\rightarrow C$）仍能保持较高的平均故障识别准确率，达到98.85%，显示出优越的泛化性和鲁棒性。在所有实验中，本文方法的平均故障识别准确率在工况载荷为$A\rightarrow A$不变的情况下达到了99.75%，而在变工况载荷为$B\rightarrow A$条件下的平均故障识别准确率达到了98.52%。这不仅解决了DenseNet在变工况下过拟合的问题，还进一步证明了IDenseNet优越的泛化性能。综合分析，本文方法在不同工况下仍具有较高的故障诊断准确率，表明模型的泛化性能较好。

其次，采用东南大学齿轮箱中轴承数据集进行实验，选取转速负载为$D=20$ Hz−0 V和$E=30$ Hz−2 V 2种不同工况负载下的数据作为训练集和测试集。采用CWT中Morlet小波方法生成二维时频图像，图片大小为224×224。将图片分别与VGG、ResNet、DenseNet和IDenseNet 4种网络相结合的方法作为对比实验。每组实验迭代50次，计算测试集的平均故障识别准确率。然后每组实验重复5次，最后取5次结果的平均值作为最终的平均故障识别准确率，以确保实验结果的准确性。不同方法得到的轴承故障诊断结果如图14所示，其中$D\rightarrow E$表示以数据集D（转速负载为$D=20$ Hz−0 V）作为训练集，数据集E（转速负载为$E=30$ Hz−2 V）作为测试集，其余含义以此类推。通过对实验数据的对比可以看出，CWT-DenseNet在$E\rightarrow D$工况负载下的平均故障识别准确率仅为79.42%，表现出明显的过拟合现象。然而，CWT-IDenseNet有效解决了这一问题，其在$E\rightarrow D$工况负载下的平均故障识别准确率达到98.64%，显著优于CWT-DenseNet。此外，CWT-IDenseNet在所有工况负载下均表现出较高的平均故障识别准确率，其中$D\rightarrow D$的平均故障识别准确率达到99.44%，$D\rightarrow E$为97.86%，$E\rightarrow E$为99.52%。这表明本文方法不仅在恒定工况下表现出优异的诊断能力，在变工况下依然保持了稳定的高准确率，较好地解决了DenseNet在变工况下过拟合的问题，展示了较好的泛化性能和适应能力。

为了更直观地看出所改进网络的优越性，展示了IDenseNet、DenseNet、ResNet和VGG 4种方法在变工况载荷$A\rightarrow B$和$D\rightarrow E$时，测试集最优诊断结果的可视化散点分布图和混淆矩阵图，分别如图15和图16所示。



	CWT-ResNet	CWT-VGG	CWT-DenseNet	CWT-IDenseNet
$D\rightarrow D$	93.56%	94.21%	95.27%	99.44%
$D\rightarrow E$	88.83%	89.12%	87.43%	97.86%
$E\rightarrow D$	85.56%	86.44%	79.42%	98.64%
$E\rightarrow E$	93.12%	94.34%	95.14%	99.52%

图 14　不同方法故障诊断对比图

Fig. 14　Comparison diagram of fault diagnosis of different methods

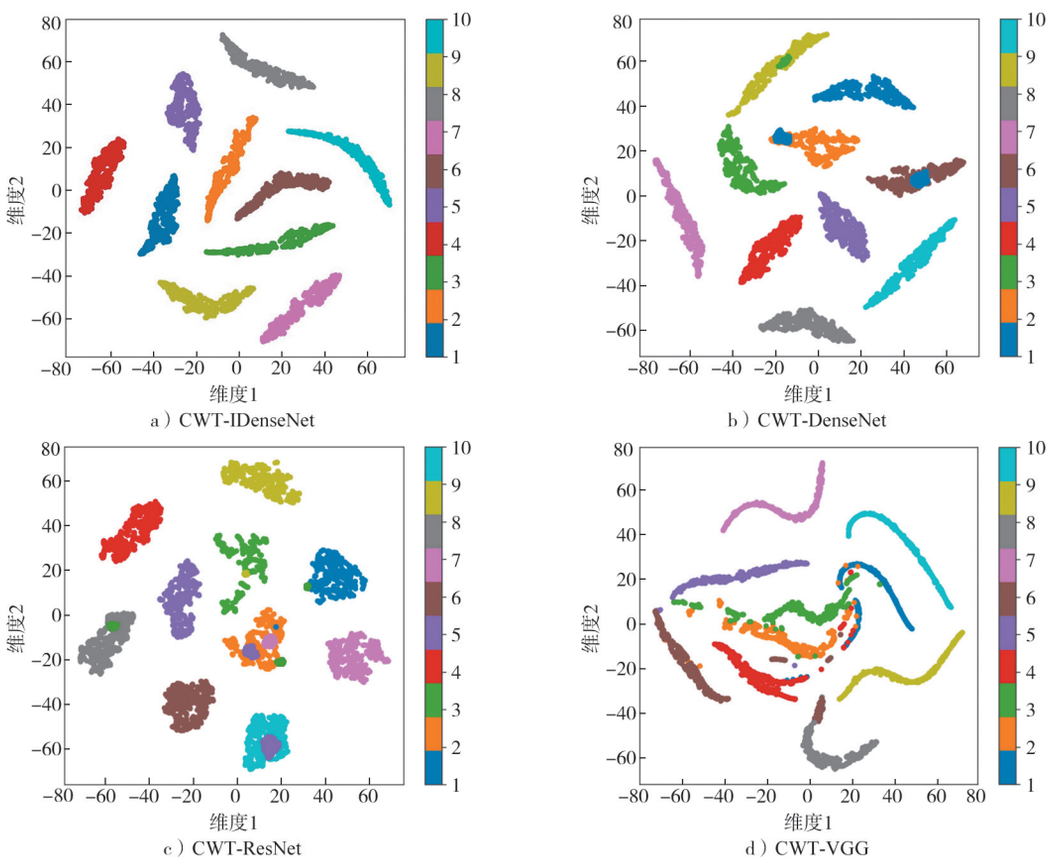


图 15　变工况 $A\rightarrow B$ 条件下 4 种方法测试集数据分布可视化

Fig. 15　Visualization of test set data distribution of four methods under variable condition $A\rightarrow B$

图 15 为凯斯西储大学数据集下工况载荷为 $A\rightarrow B$ 的可视化散点分布图，其中每一堆代表一类故障特征点，CWT-IDenseNet 模型表现出良好的分类效果，特征点之间有明显边界且同类之间联系紧密。相比之下，其他 3 种方法的表现出现了混乱和错分的情况，整体效果不及 CWT-IDenseNet 网络。这些结果清晰地表明了在变工况下，CWT-IDenseNet 网络对特征点的分类效果较好。

图 16 为东南大学数据集下工况为 $D\rightarrow E$ 的混淆矩阵图，可以看出本文方法分类效果最好，没有出现错分的情况，而 ResNet、VGG 和 DenseNet 都出现了错分的情况，其中 ResNet 分类效果最差，有 184 个测试样本没有被正确识别。这些结果清晰地表明了在变工况下，CWT-IDenseNet 网络对特征点的分类效果较好。

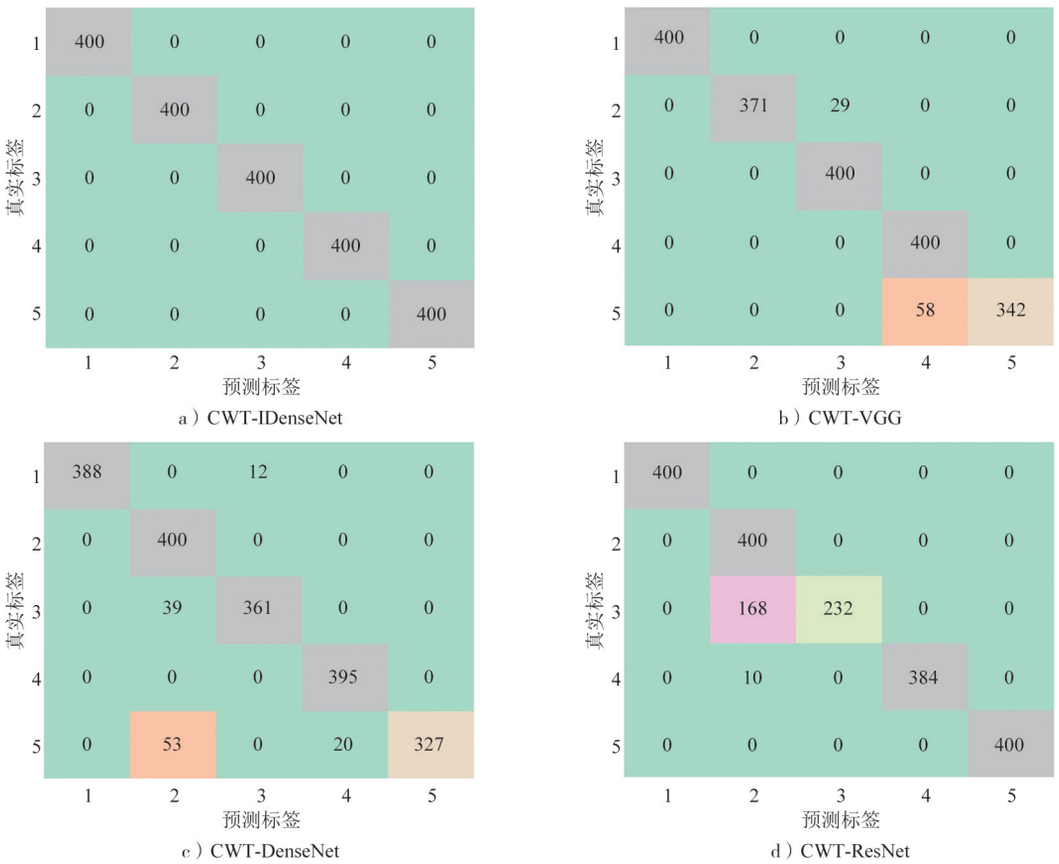


图 16 变工况 $D\rightarrow E$ 条件下 4 种方法测试集混淆矩阵

Fig. 16 Confusion matrix of test set of four methods under variable working condition $D\rightarrow E$

通过上述分析可知，本文方法在变工况下的故障诊断具有良好的泛化性。

3.4 消融实验

采用东南大学数据集在变工况负载为 $E\rightarrow D$ 时进行实验，通过消融实验分别探讨各改进项对模型分类准确率的影响。迭代 50 次，取测试集最高故障诊断率，实验结果如表 4 所示。

表 4 消融实验结果

Tab. 4 Results of ablation experiment

%

改进点	ResNet	VGG	DenseNet
原始	87.26	88.54	81.12
Swish	89.64	90.67	84.04
Swish+SRM	92.26	93.31	91.86
Swish+SRM+CBAM	95.98	96.06	99.92

由表 4 可知，将第 1 个卷积块中的 ReLU 激活函数替换为 Swish 激活函数后，神经网络的准确率得到了提升。Swish 激活函数更具平滑性，可以使梯度传播更加稳定，缓解梯度消失问题，非单调性的特性可以保持更多的梯度信息，减少神经元死亡现象。但 DenseNet 变工况下过拟合仍然存在。

加入 SRM 注意力机制后，DenseNet 准确率得到了较大提升。SRM 可以帮助模型更好地关注特征图中的重要区域特征和上下文信息，关注特征通道权重，提高特征表示能力和泛化能力，从而在一定程度上缓解过拟合现象。SRM 机制的主要作用包括提高特征表达能力、减少参数冗余以及提高模型的泛化能力。但 DenseNet 的故障识别准确率不是很高，需进一步改进。

加入 CBAM 注意力机制后，DenseNet 准确率得到了大幅提升，CBAM 可以同时从通道和空间 2 个维度增强特征表示能力，使模型能够更加准确地捕捉图像中的重要信息，提高特征表达能力和泛化能力，使 DenseNet 在变工况下具有较高分类准确率，没有出现过拟合现象。

消融实验结果表明，本文方法有效地解决了 DenseNet 变工况下过拟合以及故障诊断效果不佳的问题。

3.5 基于不同迁移学习模型的方法对比

采用东南大学齿轮箱中轴承数据集进行实验，选择恒定工况负载为 $D=20$ Hz−0 V 下的数据，以评估不同故障诊断模型的性能。为了验证迁移学习方法在故障诊断中的有效性，采用官网上已经训练好的模型权重进行迁移学习，并对比几种不同的迁移学习方法。使用 ResNet50、VGG16 和 DenseNet121 迁移学习模

型，并在相同的数据集和实验设置下对比其性能，每个模型均迭代训练 50 次，计算其在测试集上的平均故障识别准确率。不同迁移学习方法测试集故障诊断准确率曲线如图 17 所示，不同迁移学习方法的平均故障识别准确率如表 5 所示。

表 5　不同迁移学习模型方法对比

Tab. 5　Comparison of different transfer learning model methods

参数	TL-ResNet50	TL-VGG16	TL-DenseNet121	IDenseNet
平均故障识别准确率/%	97.25	98.76	98.05	99.44

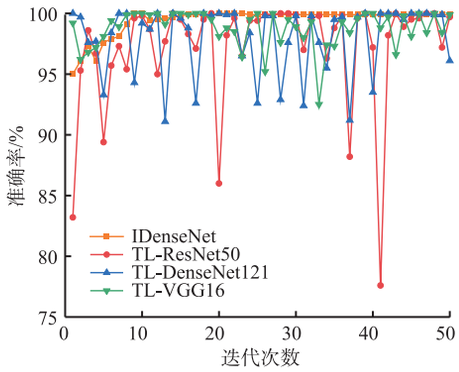


图 17　不同方法测试集的准确率

Fig. 17　Accuracy of test sets of different methods

由图 17 可知，随着迭代次数的增加，各模型的测试集故障诊断准确率变化各不相同。TL-ResNet50 的准确率波动较大，尤其在第 41 次迭代时，出现了显著的下降，这表明该模型在稳定性方面存在一定的问题。TL-VGG16 和 TL-DenseNet121 在迭代过程中也出现了相对明显的波动，其准确率低于 IDenseNet。IDenseNet 在每个迭代阶段都表现出较高的准确率，波动幅度也较小，显示出更好的鲁棒性和稳定性。

由表 5 可见，所提出的 IDenseNet 在故障诊断任务中表现出最优的性能，其测试集平均故障识别准确率达到了 99.44%，显著高于其他迁移学习模型。这表明，本文方法在处理轴承故障诊断任务时具有更高的鲁棒性和准确性。

4　结　语

本文将 CWT 与 IDenseNet 网络相结合，特别是引入 Swish 激活函数和双注意力机制(SRM 和 CBAM)对 DenseNet 网络进行改进，提出了 CWT-IDenseNet 的滚动轴承故障诊断方法，并将该方法在恒定工况以及变工况下进行了实验分析和验证，得到以下结论。

1)将 Swish 激活函数、SRM 和 CBAM 注意力机制相结合，并应用于变工况下的轴承故障诊断，解决了 DenseNet 变工况下过拟合以及诊断效果不佳的问题，显著提升了模型的泛化能力和诊断准确率。该模型在凯斯西储大学轴承数据集和东南大学齿轮箱中轴承数据集下进行了实验验证。实验结果表明，所提方法在恒定工况及变工况下均表现出较高的平均故障识别准确率，达到了 97.80%以上。

2)在凯斯西储大学的数据集下进行实验，通过将 CWT、灰度图、包络谱图和倒频谱图分别与 VGG、ResNet、DenseNet 和 IDenseNet 相结合，并在变工况下进行对比实验，验证了所提 CWT-IDenseNet 方法故障诊断准确率高于其他方法，在不同工况下仍具有较高的故障诊断准确率，没有出现过拟合现象，平均故障识别准确率达到 99.23%，验证了本模型具有较好的泛化性。

3)在东南大学齿轮箱中轴承的数据集下进行实验，通过将 CWT 与 VGG、ResNet、DenseNet 和 IDenseNet 相结合，并在变工况下进行对比实验，验证了所提 CWT-IDenseNet 方法故障诊断准确率高于其他方法，在不同工况下仍具有较高的故障诊断率，没有出现过拟合现象，平均故障识别准确率都达到了 97.80%以上，验证了本文方法具有较好的泛化性。通过系统的消融实验和迁移学习对比实验，验证了本文方法的有效性和优越性。

本文使用的 CWT 时频图像输入 IDenseNet 网络的池化层后，有时会出现部分特征丢失的问题，这不可避免地影响了模型对细节特征的识别。为了弥补这一不足，未来拟探讨使用未池化卷积层或自适应池化方法，通过自适应调整池化区域的大小，保留更多的细节特征，从而进一步提升模型的性能。

参考文献/References：

[1] LIU Yunpeng，JIANG Hongkai，LIU Chaoqiang，et al. Data-augmented wavelet capsule generative adversarial network for rolling bearing fault diagnosis[J]. Knowledge-Based Systems，2022. DOI：10.1016/j.knosys.2022.109439.

[2] 杨婧婷. 机械加工过程中轴承故障诊断方法研究[J]. 机械设计与制造，2021(2)：112-116.
YANG Jingting. Bearing fault diagnosis method in the machining process[J]. Machinery Design & Manufacture，2021(2)：112-116.

[3] 刘东波. 基于振动信号的冶金机械中轴承故障特征提取及诊断分析[J]. 机械管理开发，2024，39(1)：41-42.

LIU Dongbo. Feature extraction and diagnosis analysis of bearing failure in metallurgical machinery based on vibration signal[J]. Mechanical Management and Development,2024,39(1):41-42.

[4] 郭艳平,龙涛元. 振动信号模型和散度在诊断滚动轴承故障中的应用[J]. 机械设计与制造,2024,403(9):311-315.
GUO Yanping,LONG Taoyuan. The application of vibration signal model and divergence in fault diagnosis of rolling bearing[J]. Machinery Design & Manufacture,2024,403(9):311-315.

[5] 周建民,王云庆,李家辉. 结合振动信号图像化和 RepVGG 的滚动轴承故障诊断方法[J/OL]. 机械科学与技术(2023-03-15)[2024-04-15]. https://link.cnki.net/doi/10.13433/j.cnki.1003-8728.20230154.
ZHOU Jianmin,WANG Yunqing,LI Jiahui. Rolling bearing fault diagnosis method combined with vibration signal visualization and RepVGG[J/OL]. Mechanical Science and Technology(2023-03-15)[2024-04-15]. https://link.cnki.net/doi/10.13433/j.cnki.1003-8728.20230154.

[6] 刘湘楠,赵学智,何宽芳. 圆柱滚子轴承振动信号时频特征提取及状态识别[J]. 振动工程学报,2022,35(4):932-941.
LIU Xiangnan,ZHAO Xuezhi,HE Kuanfang. Time-frequency feature extraction and state recognition of vibration signal of cylindrical roller bearing[J]. Journal of Vibration Engineering,2022,35(4):932-941.

[7] TAHMASBI D,SHIRALI H,SAJAD MOUSAVI NEJAD SOUQ S,et al. Diagnosis and root cause analysis of bearing failure using vibration analysis techniques[J]. Engineering Failure Analysis,2024. DOI: 10.1016/j.engfailanal.2023.107954.

[8] 赖奎,戴雄杰,潘松波,等. 基于 STFT 和 CNN-Attention 的配电终端采集模块故障诊断研究[J]. 自动化仪表,2023,44(9):37-41.
LAI Kui,DAI Xiongjie,PAN Songbo,et al. Fault diagnosis study of power distribution terminal acquisition module based on STFT and CNN-attention[J]. Process Automation Instrumentation,2023,44(9):37-41.

[9] 刘鑫,贾云献,苏小波,等. 基于灰度图像纹理分析的柴油机失火故障特征提取[J]. 振动与冲击,2019,38(2):140-145.
LIU Xin,JIA Yunxian,SU Xiaobo,et al. Fault feature extraction for diesel engine misfires based on the gray image texture analysis[J]. Journal of Vibration and Shock,2019,38(2):140-145.

[10] 李振宇,邓向阳,张立民,等. 基于 Swish 激活函数的双通道 CNN 结构[J]. 计算机与数字工程,2020,48(6):1413-1416.
LI Zhenyu,DENG Xiangyang,ZHANG Limin,et al. Dual-channel CNN structure based on swish activation function[J]. Computer and Digital Engineering,2020,48(6):1413-1416.

[11] ZHANG Bangcheng,SUN Shiqi,YIN Xiaojing,et al. Bearing fault diagnosis based on wavelet transform and convolution neural network[C]//2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS). Yibin:IEEE,2023:1-5.

[12] 王焜,刘鑫,杨嘉其,等. 基于改进 DenseNet 模型的滚动轴承变工况故障诊断[J]. 组合机床与自动化加工技术,2022(6):78-81.
WANG Kun,LIU Xin,YANG Jiaqi,et al. Fault diagnosis of rolling bearing under variable working condition based on improved DenseNet model[J]. Modular Machine Tool & Automatic Manufacturing Technique,2022(6):78-81.

[13] 程建刚,毕凤荣,张立鹏,等. 基于 MACNN 的柴油机故障诊断方法研究[J]. 振动与冲击,2022,41(10):8-15.
CHENG Jiangang,BI Fengrong,ZHANG Lipeng,et al. Fault diagnosis method for diesel engines based on MACNN[J]. Journal of Vibration and Shock,2022,41(10):8-15.

[14] MA Xinna,ZHENG Xuepeng,LIU Qinging,et al. Research on bearing fault diagnosis based on improved dense connection network[C]//2023 7th International Symposium on Computer Science and Intelligent Control (ISCSIC). Nanjing:IEEE,2023:216-220.

[15] 熊鹏,汤宝平,邓蕾,等. 基于动态加权密集连接卷积网络的变转速行星齿轮箱故障诊断[J]. 机械工程学报,2019,55(7):52-57.
XIONG Peng,TANG Baoping,DENG Lei,et al. Fault diagnosis for planetary gearbox by dynamically weighted densely connected convolutional networks[J]. Journal of Mechanical Engineering,2019,55(7):52-57.

[16] 王太勇,宫立明,王鹏,等. 基于 KD-DenseNet 的旋转机械故障诊断模型[J]. 振动与冲击,2020,39(16):39-45.
WANG Taiyong,GONG Liming,WANG Peng,et al. A fault diagnosis model of rotating machinery based on KD-DenseNet[J]. Journal of Vibration and Shock,2020,39(16):39-45.

[17] YANG Li,HU Tianhan,ZHANG Xueyuan,et al. Enhanced classification of COVID-19 CT images using CDenseNet with CBAM attention and swish activation[C]//2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA). Dalian:IEEE,2023:160-165.

[18] LEE H,KIM H E,NAM H. SRM:A style-based recalibration module for convolutional neural networks[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul:IEEE,2019:1854-1862.

[19] 韩争杰,牛荣军,马子魁,等. 基于注意力机制改进残差神经网络的轴承故障诊断方法[J]. 振动与冲击,2023,42(16):82-91.
HAN Zhengjie,NIU Rongjun,MA Zikui,et al. Bearing fault diagnosis methods based on an attentional-mechanism-improved residual neural network[J]. Journal of Vibration and Shock,2023,42(16):82-91.

[20] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the case western reserve university data: A benchmark study [J]. Mechanical Systems and Signal Processing, 2015, 64: 100-131.

[21] SHAO S Y, MCALEER S, YAN R Q, et al. Highly accurate machine fault diagnosis using deep transfer learning[J]. IEEE Transactions on Industrial Informatics, 2018, 15(4): 2446-2455.

[22] 曹思灿. 基于生成对抗网络的轴承故障诊断方法研究[D]. 武汉:华中科技大学,2019.
CAO Sican. Generative Adversarial Network Based Methods for Rolling Bearing Fault Diagnosis[D]. Wuhan:Huazhong University of Science and Technology,2019.