

基于 SPWVD 与知识蒸馏的行星变速器故障诊断研究

马超^{1,2} 郑鑫辉^{1,2} 王少红^{1,2} 徐小力^{1,2}

(1. 北京信息科技大学现代测控技术教育部重点实验室 北京 100192;

2. 北京信息科技大学机电系统测控北京市重点实验室 北京 100192)

摘要:行星变速器运行工况多是非平稳运行工况,运行过程中齿轮啮合振动信号相互耦合导致测试信号混叠,其隐藏故障诊断难度增大;同时应用复杂的神经网络模型进行故障诊断预测时多数会收到工业现场边缘计算设备硬件的限制。针对相关问题,在保证行星变速器故障诊断准确率的情况下减少网络模型的参数量,提出了一种应用平滑伪维格纳-威利分布(smooth and pseudo Wigner-Vile distribution, SPWVD)与知识蒸馏结合的智能识别模型用于行星变速器故障诊断。首先,利用集合经验模态分解(EEMD)方法将多分量振动信号分解后选取单分量信号进行 SPWVD 计算后线性叠加得到二维时频图作为输入,以 ResNet101 为教师模型指导学生模型 MobileNet 进行训练,复杂教师模型将数据中的知识传授给学生模型,提高了学生模型的精度。将该方法与同类方法进行了对比,结果表明,模型以牺牲 2.43% 准确率为代价,存储成本下降为教师模型的 24.55%,相较于未知识蒸馏的 MobileNet 的准确率提高了 9.61%,实现模型轻量化。本研究方法对提高深度学习模型在工程实际应用,降低边缘计算设备部署成本提供了一种有效且可行的解决方法。

关键词:故障诊断; SPWVD; 知识蒸馏; MobileNet; 行星变速器

中图分类号: TH133.33; TP206.3 **文献标识码:** A **国家标准学科分类代码:** 460.40

Research on fault diagnosis of planetary transmission based on SPWVD and knowledge distillation

Ma Chao^{1,2} Zheng Xinhui^{1,2} Wang Shaohong^{1,2} Xu Xiaoli^{1,2}

(1. Key Laboratory of Modern Measurement and Control Technology Ministry of Education, Beijing Information Science & Technology University, Beijing 100192, China; 2. Beijing Key Laboratory of Measurement and Control of Mechanical and Electrical System, Beijing Information Science & Technology University, Beijing 100192, China)

Abstract: The operating conditions of planetary transmission are mostly non-stationary operating conditions. During the operation, the gear meshing vibration signals are coupled with each other, which leads to the aliasing of test signals, and the difficulty of hidden fault diagnosis increases. At the same time, when applying complex neural network models for fault diagnosis and prediction, most of them will be limited by the hardware of industrial field edge computing equipment. Aiming at the related problems, an intelligent recognition model based on smooth and pseudo Wigner-Vile distribution (SPWVD) and knowledge distillation is proposed to reduce the parameters of the network model while ensuring the accuracy of planetary transmission fault diagnosis. First, the ensemble empirical mode decomposition (EEMD) method is used to decompose the multi-component vibration signal, then the single-component signal is selected for SPWVD calculation and linearly superimposed to obtain a two-dimensional time-frequency diagram as input. The ResNet101 is used as the teacher model to guide the student model MobileNet for training. The complex teacher model imparts the knowledge in the data to the student model, which improves the accuracy of the student model. The method is compared with similar methods. The results show that the storage cost of the model is reduced to 24.55% of the teacher model at the expense of 2.43% accuracy, which is 9.61% higher than that of MobileNet without knowledge distillation. This research method provides an effective and feasible solution to improve the practical application of deep learning model in engineering and reduce the deployment cost of edge computing equipment.

Keywords: Fault diagnosis; SPWVD; knowledge distillation; MobileNet; planetary transmission

0 引言

行星变速器在风力发电机、直升机、舰船等大型机械设备中,因其传动比大、承载效率高等优点具有广泛应用^[1]。在实际运行中,不同于定轴变速器,行星变速器往往处在低速交变重载的复杂恶劣工作环境中,更容易发生故障,因此对行星变速器进行健康监测具有重要意义^[2]。由于传感器安装位置固定,行星齿轮公转导致采集到啮合点的振动响应周期变换,信号呈现非线性^[3]。行星变速器中零部件繁多,振动信号成分复杂,故障响应微弱,故障信息淹没在噪声中不易提取。并且在一个齿轮发生故障后,会加速相啮合的齿轮失效,产生复杂耦合故障,大大增加故障诊断难度^[4]。

行星变速器是一种典型的旋转系统,当行星变速器中的齿轮、轴承等部件发生故障时会产生一系列的冲击信号,这些测信号包含准周期性指数衰减结构的脉冲信号。学者们提出了大量用于提取故障信息的信号处理方法,包括快速傅立叶变换、小波变换、希尔伯特黄变换和维格纳-威利分布(Wigner-Ville distribution, WVD)等,用于分析判断变速器的运行状态、故障类别等^[5]。行星变速器振动信号具有非平稳和分线性特点,采用时频分析方法可以同时反映出信号在时域频域的信息^[6]。WVD对信号的瞬时频率和局域化具有清晰表达^[7],为消除其交叉项的影响产生了伪维格纳-威利分布(PWVD)和SPWVD,文献[8]提出将EEMD方法和PWVD相结合利用卷积神经网络进行分类识别;文献[9]构建EEMD优势分量多重分型与重组二次型SPWVD相结合,更好的揭示爆破振动信号时频特性;文献[10]提出一种基于SPWVD时频纹理图特征的故障诊断方法。

从工程应用角度出发,虽然上述方法取得了良好的诊断效果,但是由于所应用的深度神经网络深度较高、模型参数量较多,难以直接部署在嵌入式边缘计算设备。因此,考虑分析模型诊断效果的同时还需要考虑到工程应用可行性等因素^[11]。Hinton等^[12]于2015首次提出了知识蒸馏(knowledge distillation, KD),通常是利用在相同带标签数据训练好的复杂深层网络教师模型,通过蒸馏训练向简单浅层学生模型传递知识,用以提高学生模型性能接近教师网络,从而实现模型压缩^[13-15]。通过知识蒸馏技术,可以实现半监督迁移决策、类别增量中的灾难性遗忘缓解、模型轻量化等问题^[16-19]。

针对上述有关问题,本文提出一种将EEMD与SPWVD相结合通过知识蒸馏来实现模型的轻量化,解决了深度模型的轻量化部署问题。首先利用EEMD将多分量振动信号分解后选取单分量信号进行SPWVD计算以避免SPWVD种交叉项干扰影响,将一维振动信号转为

具有清晰时频特征的二维时频分布图像作为模型输入,选取庞大且高识别精度的教师模型ResNet,选取小巧且高计算速度的轻量化学生模型MobileNet,然后利用训练好的教师模型指导学生模型训练,利用知识蒸馏方法得到一个识别精度较高的低复杂度学生模型。本文方法为提升简单模型的诊断精度,在边缘计算设备部署提供了可行性方案。

1 理论方法

1.1 EEMD-SPWVD 时频分析

行星变速器中零件繁多,齿轮振动响应复杂,测取信号受噪声干扰,需进行预处理突出特征信息。EEMD是一种具有自适应性的辅助信号分解方法,不必选择基函数和阈值,可以根据信号自身特性,将耦合的非平稳随机信号分解为若干频率的窄带信号,将调制信号剥离出来,一定程度上抑制了EMD中模态混叠问题,保留故障信息的同时增强了消噪效果,对非平稳信号具有很好的识别能力。其原理时通过在原始信号中添加白噪声并对其进行经验模态分解(empirical mode decomposition, EMD),用添加噪声来除去干扰信号,最后将多次分解后的结果进行集总平均计算。EEMD分解后的IMF定义为:

$$c_j(t) = \frac{1}{M} \sum_{i=1}^M c_{i,j}(t) \quad (1)$$

式中: $c_j(t)$ 为第 j 个IMF, j 为分解层。

信号EEMD分解后可以得到一系列IMF和一个残余分量 $r(t)$,原始振动信号 $x(t)$ 可表示为:

$$x(t) = \sum_j c_j(t) + r(t) \quad (2)$$

WVD是一种基于时频能量分布的非平稳信号的非线性时频分析工具,具有很好的时频聚集性、平移不变性等优点。由于振动信号具有非平稳性,时频分析具有很好的应用前景。SPWVD方法是为保证时频分辨率及能量正值性的同时消除WVD交叉干扰项影响,在三维准确率上有很大提升,可以更加清晰的表现时频聚集性,具有更好的时频特征分析能力。其原理是通过在时域、频域同时加入窗函数,通过在WVD与平滑窗函数进行卷积,可以起到消除交叉干扰的作用。信号 $x(t)$ 的SPWVD的表达式为:

$$\begin{aligned} \text{SPWVD}_z(t, f) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(u) h(\tau) x(t - u + \frac{\tau}{2}) \\ & x^*(t - u - \frac{\tau}{2}) e^{-j2\pi f \tau} \text{d}u \text{d}\tau \end{aligned} \quad (3)$$

式中: $z(t)$ 是分析信号 $x(t)$ 的解析信号, τ 为积分变量, $h(t)$ 为时域平滑窗函数, $g(t)$ 频域平滑窗函数。

由于SPWVD是一种二次型时频分布,服从二次叠

加原理,即满足一个复杂信号的每一个分量都有一个对应的信号项以及各分量彼此间的交叉项,信号越复杂,尤其在信号特征频率相近时交叉项干扰越严重。通过 EEMD 可以将信号分解成多个 IMF,即多个单分量信号。将每个 IMF 进行 SPWVD 计算,并将各分量信号的 SPWVD 计算结果进行线性叠加,可以抑制干扰项并最终得到一个更加清晰的信号时频分布图。

1.2 知识蒸馏

知识蒸馏通常是将一个模型结构复杂、特征提取能力强的教师模型进行训练,将学习得到的“暗知识”(软标签)传授给结构简单、运行速率快的学生模型,作为训练时的标签信息指导其进行训练,通过最优化蒸馏损失函数优化学生模型的参数,得到一个接近教师模型准确率的模型,知识蒸馏原理如图 1 所示。

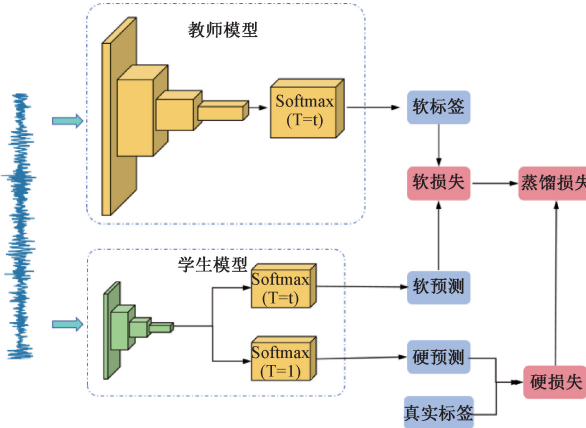


图 1 知识蒸馏原理

Fig. 1 Knowledge distillation principle

软标签是教师模型给学生模型训练时的“知识”,是在位于教师模型的 softmax 输出层引入一个温度系数 T ,“软化”给出的预测概率向量,可以反映出类别之间的相似度。因此可以使学生模型保持较少参数接近学习教师模型性能的目的。KD-Softmax 表达式为:

$$q_i = \frac{\exp(z_i/T)}{\sum_j \exp(z_j/T)} \quad (4)$$

式中: q_i 为软化后的概率向量; z_i 为输出层结果; 温度系数 T 增大会导致输出层的结果越平滑。

软损失是在温度系数 $T=t$ 下,由两个交叉熵损失函数(教师模型软标签损失和学生模型软预测损失)加权求和得到,其函数表达式为:

$$L_{soft} = - \sum_j p_j^t \log(q_j^t) \quad (5)$$

式中: q_j^t 表示温度系数为 t 的条件下学生模型分布概率。

硬标签是在温度系数 $T=1$ 下,由学生模型的预测结果与数据真实标签计算出的交叉熵损失函数,其函数表

达式为:

$$L_{hard} = - \sum_j c_j \log(q_j^1) \quad (6)$$

式中: c_j 为真实标签。

两个模型利用同一样本训练,其具有相似性,教师模型中的标签特征信息迁移至学生模型,同时保留各自任务的输出层。将两个目标函数教师模型 L_{soft} 和学生模型 L_{hard} 进行权重平均可以的到蒸馏损失函数,其函数表达式为:

$$L_{KD} = \alpha L_{soft} + (1 - \alpha) L_{hard} \quad (7)$$

式中: α 表示蒸馏强度。

将蒸馏损失函数反馈至学生模型,优化学生模型的参数,使得学生模型学习得到教师模型的“暗知识”进行正则化训练,最终实现学生模型的效果提升。

1.3 MobileNet

MobileNet^[20] 是一种轻量级卷积网络架构。他将传统网络中的标准卷积替代为深度可分离卷积深度可分离卷积取代了传统网络中的普通卷积从而降低参数量,减轻了计算和存储成本。

MobileNet 的核心为 Bneck 模块,其结构如图 2 所示,主要融合了逆残差连接、深度可分离卷积和注意力机制。

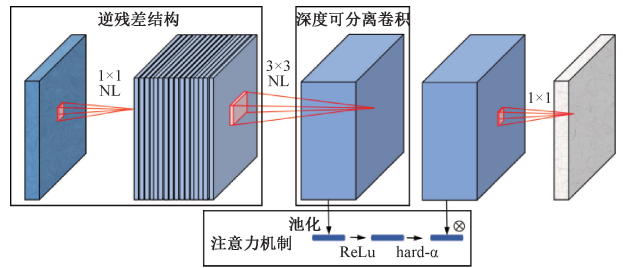


图 2 Bneck 模块结构

Fig. 2 Bneck module structure

深度可分离卷积由深度卷积 (depthwise convolution, DW) 和逐点卷积 (pointwise convolution, PW) 组成。DW 与普通卷积的区别在于 DW 一个通道有且仅由一个一维卷积核进行卷积, DW 的表达式为:

$$G_{i,j,m} = \sum_{w,h} K_{w,h,m} \cdot X_{i+w,j+h,m} \quad (8)$$

式中: X 为输入, G 为输出, 卷积核尺寸为 $W \times H$ 。

由于 DW 完成后,无法对通道进行扩展,并且由于每个通道之间的每次卷积操作都是独立进行的,因此无法有效利用同一空间位置上不同通道的特征信息。因此,需要进行 PW 进行转换通道, PW 的卷积核边长为 1, PW 作用在 DW 输出特征映射上。

将边长 D_s , 通道数 M 的特征图作为输入, 卷积核边长为 D_k , 将边长 D_g , 通道数 N 的特征图作为输出, 对比深度可分离卷积与标准卷积的计算量比值如下:

$$\frac{p_1}{p_2} = \frac{D_f^2 D_k^2 M + D_f^2 MN}{D_f^2 D_k^2 MN} = \frac{1}{N} + \frac{1}{D_k^2} \quad (9)$$

通常卷积核选用边长为 3。因此计算量比值约为标准卷积的 1/9。

MobileNet 在残差连接的基础上引申出反向残差结构,先升维处理可以接受小维度输入特征图,后进行降维处理减小输出维度减小参数量从而减小计算量。同时残差连接可以提升梯度传播效率,使得网络层更深。通过给每个特征图分配不同的权重,使得模型关注到更有用的特征。通过设置超参数宽度因子进一步实现了模型计算量和参数量的降低和分类识别准确率的提高,极大的提高了识别效率。

1.4 基于 SPWVD 和 KD 的故障诊断方法

在以上研究成果的基础上,开展基于知识蒸馏和 SPWVD 时频图的行星变速器故障诊断方法研究,本文方法的流程如图 3 所示,主要步骤如下:

步骤 1 信号采集

利用加速度传感器在行星传动试验台上采集变速器各种故障状态振动信号。

步骤 2 时频图获取

将采集到的行星变速器故障振动多分量信号先进行 EEMD 分解并筛选出蕴含振动特征的 IMF,将选择好的 IMF 进行 SPWVD 处理后线性叠加得到二维时频特征分布图,并划分数据集。

步骤 3 模型搭建与训练

将带有标签的二维时频特征分布图作为输入,设定温度系数 $T=t$ 分别训练教师模型和学生模型以得到软标签和软预测,计算得到软损失;设定温度系数 $T=1$ 训练学生模型得到硬预测,与真实标签计算得到硬损失,最终计算得到蒸馏损失函数以反馈并优化学生模型。

步骤 4 故障分类识别

测试集时频分布图输入到蒸馏后的学生模型中测试并评价模型的诊断性能。

2 实验验证

2.1 试验说明

为验证方法的可行性,搭建行星传动试验台如图 4 所示,主要包括:伺服电机、扭矩传感器、行星变速器、制动控制器、磁粉制动器、信号采集仪和上位机采集系统。

为探究模型在相同工况下的不同故障类型故障诊断性能,以行星变速器为实验对象进行了研究。其中,减速箱的型号为:PLS060-4,传动比为 1:4,振动信号采集仪型号为:INV3062T2,上位机采集系统为东方所的 DASP V10。模拟了正常及七种常见的变速器故障共 8 种

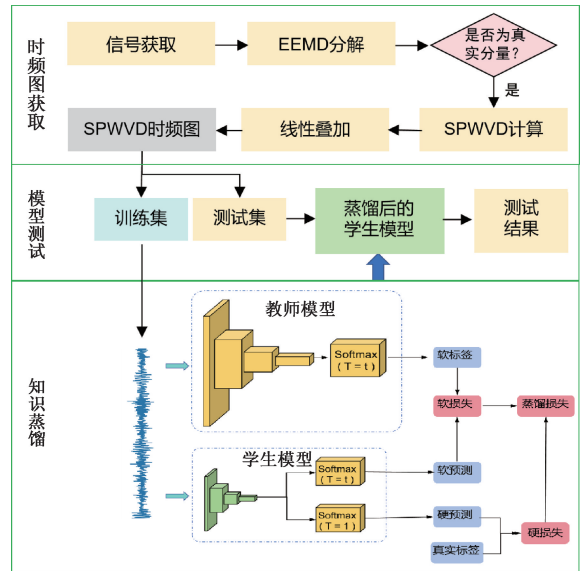


图 3 基于 SPWVD 和 KD 的故障诊断流程

Fig. 3 Fault diagnosis process based on SPWVD and KD



图 4 行星传动试验台

Fig. 4 Planetary transmission test bench

运行状态,具体标签与对应的故障状态如表 1 所示。

表 1 数据标签及对应故障类型

Table 1 Data labels and corresponding fault types

标签	对应故障类型
0	正常
1	太阳轮齿根裂纹
2	内齿圈齿根裂纹
3	内齿圈齿根裂纹 + 太阳轮齿根裂纹
4	行星轮齿根裂纹
5	行星轮齿面磨损
6	行星轮断齿
7	行星轮齿面磨损 + 断齿

选择太阳轮转速为 1 980 rpm 作为输入轴,采用振动加速度传感器型号为: INV9832A, 采样频率为 20.48 kHz,采样时间为 10 s。各故障类型在测点的时域波形如图 5 所示。

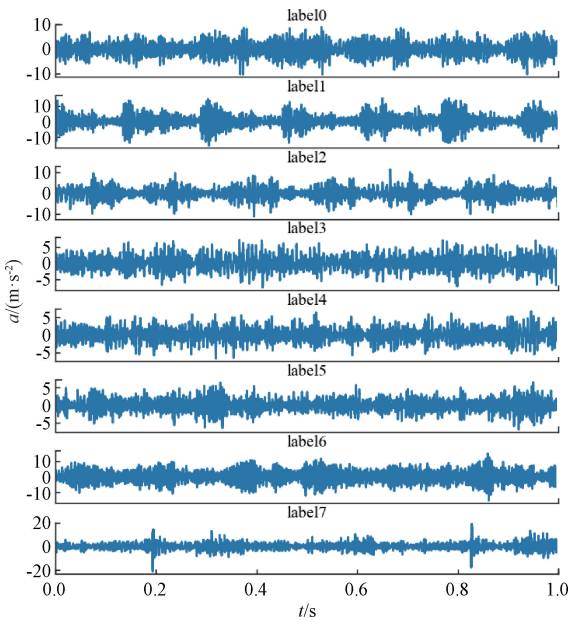


图5 各故障类型时域波形

Fig. 5 Time domain waveform of each fault type

各故障类型在测点的频谱图如图6所示,由于行星变速器存在加工安装误差、故障信号有多重途径传递到传感器等因素影响,会对故障信号的频率及其谐波产生幅值调制和频率调制,信号的幅值产生周期性变化,信号的啮合频率附近出现间隔为转频的调制边频带,具有时变特性,传统的频谱分析法会出现频率模糊,需要选择具有良好分辨、抗噪的时频分析方法。

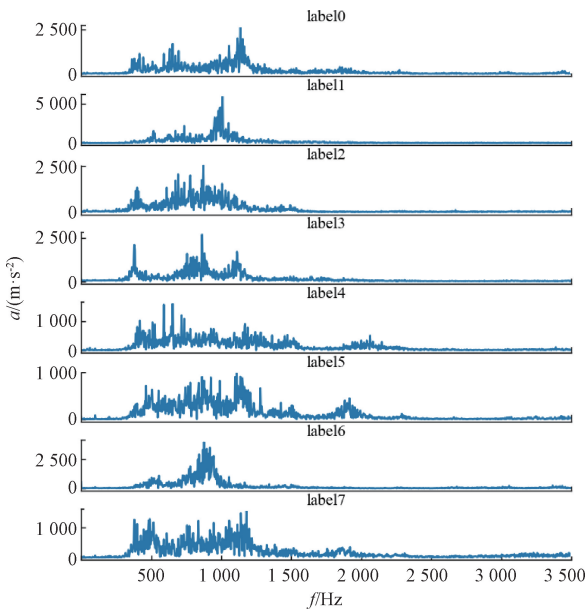


图6 各故障类型频谱图

Fig. 6 Frequency domain waveform of each fault type

2.2 EEMD-SPWVD 时频分布图数据集生成

为了更好的提取故障特征,首先进行数据预处理,选择每个样本数据的采样点个数为2048。由图5、6可知行星变速器的振动信号复杂,为避免直接进行SPWVD分析产生严重的干扰项影响,首先对振动信号进行EEMD分解,以行星轮断齿即标签为label6的振动信号为例,其EEMD分解得到的各阶分量如图7所示。

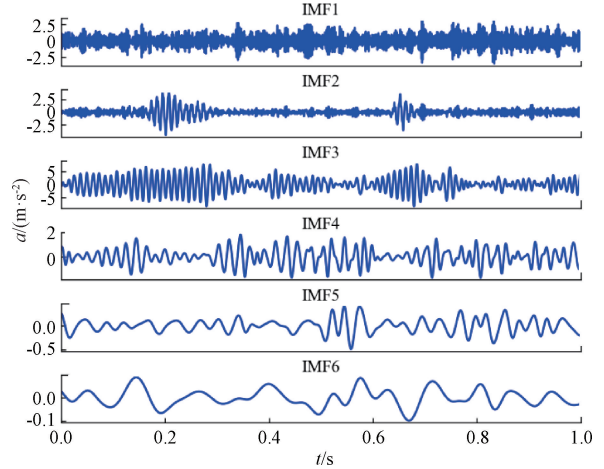


图7 EEMD分解结果

Fig. 7 EEMD decomposition results

根据图7可以很好的观察到信号的周期特性,其中IMF1包含了信号的大部分能量和信息,IMF2为信号的高频分量;IMF3分量振幅较大,但频率较低;IMF4~IMF6为频率更小的微弱分量。因此选择IMF1对其进行SPWVD计算,得到时频分布图,以此类推得到行星变速器8种运行状态的时频图如图8所示。

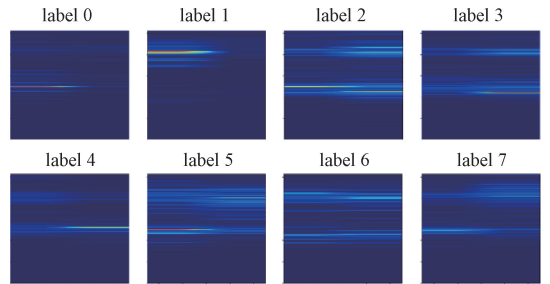


图8 8种运行状态的SPWVD时频图

Fig. 8 SPWVD time-frequency diagrams of eight health states

从8图中可以看出:8种行星变速器工作状态的SPWVD时频分布云图中,其信号能量在频域上具有明显的1或2个集中区域,同时其分布由于运行状态不同而有明显区别,呈现出不同的亮条纹。因此,EEMD和SPWVD相结合可以很好的提取非平稳信号的时变特征。

为更好的利用神经网络进行分类识别更快收敛,进行 SPWVD 处理生成时频分布图时进行标准化处理,即不显示坐标系、图像空白部分的尺寸为 128×128 的图片。每一类故障状态生成 500 张时频分布图,并将数据集均匀打乱后按照 7 : 3 划分训练集和测试集。

2.3 使用 KD 进行模型轻量化训练

ResNet 核心思想是引入残差学习模块,残差学习可以避免复杂深度卷积神经网络深度增加性能反而下降的问题^[21],因此将其作为教师模型。为验证教师模型深度对学生模型性能的影响,分别将并将 ResNet34、ResNet50、ResNet101 作为教师模型,以 MobileNet 作为学生模型数据集输入进行训练,设置迭代数为 200,学习率为 0.001,知识蒸馏前后结果对比如表 2 所示,其中深度层数为教师模型网络层数的深度,输入尺寸为所保存的 SPWVD 时频二维图的图片尺寸参数,准确率为模型测试准确率,模型大小为训练好的模型所需存储空间。

表 2 教师模型参数

Table 2 Teacher model parameters

模型	准确率/%	模型大小/MB
MobileNet	86.21	13.6
ResNet34	97.62	28.9
ResNet50	98.25	30.3
ResNet101	98.91	55.4
ResNet34-MobileNet	94.78	13.6
ResNet50-MobileNet	95.22	13.6
ResNet101-MobileNet	95.82	13.6

通过实验结果可以看出,学生模型在使用知识蒸馏后准确率都有显著提高,说明原学生模型通过知识蒸馏技术在不同教师模型种都获取了知识,在保证模型大小的前提下,准确率与教师模型相接近。

通过实验结果发现,在模型大小方面,虽然 ResNet101 高于 ResNet34 和 ResNet50,但蒸馏后的学生模型大小差距不大,说明教师模型的复杂程度对蒸馏后学生模型大小影响不大,但当教师模型选用 ResNet101 时,识别准确率最高,准确率为 95.82%,这说明教师模型复杂,传授给学生模型的知识具有更好的指导作用。ResNet101 作为教师模型蒸馏后的学生模型仅为原教师模型大小的 25%,可以看出,知识蒸馏处理能以少量准确率为代价,大大减少模型计算成本和存储成本,有利于模型部署,ResNet101-MobileNet 模型训练结果的准确率和损失值如图 9 所示。

为了进一步分析模型在测试集上的性能,使用混淆矩阵对结果进行分析,T-SNE 算法用于显示最终全局学习到的特征,结果如图 10 和 11 所示。如图所示,行星变速器准确率为 100%,太阳轮齿根裂纹、内齿圈齿根裂纹

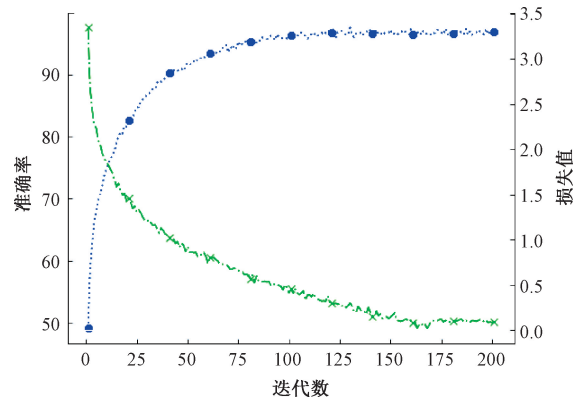


图 9 ResNet101-MobileNet 训练结果

Fig. 9 ResNet101-MobileNet training results

及他们的耦合故障相互混淆,行星轮齿面磨损、行星轮断齿及他们的耦合故障相互混淆,两种耦合故障的准确率分别为 88.78% 和 90.60%,8 种行星变速器平均识别率 95.82% 表明模型未知数据集上具有较好的泛化性能。从图可以看出模型可以准确地将行星变速器的每种故障特征完全剥离,证明模型具有很好的特征提取能力。

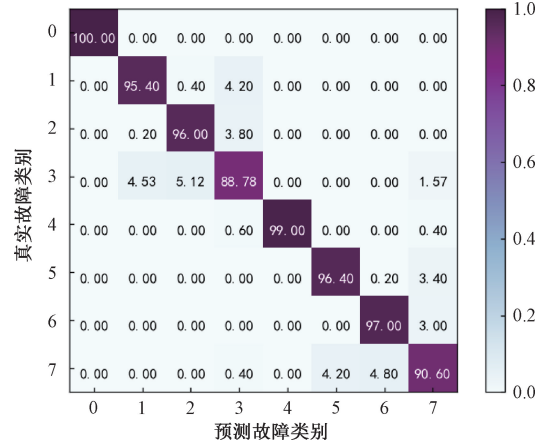


图 10 混淆矩阵

Fig. 10 Confusion matrix

2.4 不同工况对比

为了进一步验证本文所提方法信号混叠条件下的有效性,在不改变采样频率的条件下,改变太阳轮转速对比实验结果如表 3 所示。

表 3 不同工况下模型性能

Table 3 Model performance at different speeds

转速/rpm	准确率/%
1 140	95.87
1 380	95.86
1 740	95.79
1 980	95.82
2 220	95.66

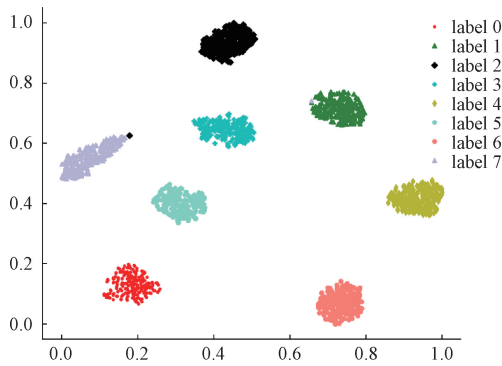


图 11 T-SNE 可视化结果

Fig. 11 T-SNE visualization results

通过表 4 可以看出,模型在不同转速工况下均获得了较高准确率且基本相等,说明具有较好的有效性。

2.5 不同算法对比

为了进一步验证本文所提方法的有效性和优越性,选取了经典的卷积神经网络 CNN 以及选择 ShuffleNet 作为学生模型的故障诊断模型作为对比,实验结果如表 4 所示。

表 4 模型性能对比

Table 4 Model performance comparison

模型	大小/MB	训练时间/s	准确率/%
CNN	31.6	2.6	89.68
MoblieNet	13.6	1.9	86.21
ResNet101	55.4	9.4	98.25
ResNet101+MoblieNet	13.6	2.5	95.82
ResNet101+ShuffleNet	13.6	2.6	95.42

通过表 3 可以看出,经过知识蒸馏后的模型在保证大小不变的情况下均获得了较高的准确率,教师模型同样是 ResNet101 对 ShuffleNet 的蒸馏指导效果略差于 MoblieNet,原因是 ResNet101 和 MoblieNet 都是基于残差思想,网络结构相近,取得了较好的蒸馏效果。

3 结论

行星变速器故障的诊断和预警,是保障具有传动系统的大型设备正常运行的关键技术之一。由于具有行星变速器的传动系统多数是在非平稳运行工况中工作运行,传感器采集到的故障信号受到时变噪声干扰;同时由于齿轮啮合等因素导致其故障波形叠加现象多。针对相关问题,为提高行星变速器运行状态的精准监测,提出了一种应用平滑伪维格纳-威利分布与知识蒸馏结合的智能识别模型用于行星变速器故障诊断预测。主要结论

如下:

1) 利用 SPWVD 方法对微弱的故障信号十分敏感的特点,充分挖掘信号时频信息,通过 EEMD 将多分量信号分解后进行 SPWVD 处理后累加以避免交叉项影响,有效揭示了信号规律,提高识别精度,对于滚动轴承早期故障检测十分有利。

2) 当教师模型为 Resnet101 时,学生模型 MobileNet 的试验准确率为 95.82%。学生模型蒸馏后准确率提高 9.61%,提高了小模型的识别精度,说明知识蒸馏的可行性。

3) 经知识蒸馏后的 MobileNet 学生模型准确率与未蒸馏的教师模型 Resnet101 相差 2.43%,而模型大小仅为教师模型 Resnet101 的 24.55%,很好的实现了模型轻量化,在边缘计算设备部署模型提供了可行性方案。

本文的主要目的是解决深度学习模型在边缘计算设备的轻量化部署的模型压缩与加速问题,未来研究中会在参数剪枝、参数量化、紧凑网络几个方向进行进一步优化。知识蒸馏主要是以教师模型的 softmax 层输出作为学习信息指导学生模型进行训练,未来结合旋转信号特点利用白噪声与测试信号叠加构建中间特征层后将与 softmax 层结合,从而改进模型结构以及提升模型的可靠性。

参考文献

[1] 雷亚国,何正嘉,林京,等.行星变速器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 progress of fault diagnosis technology for planetary transmission[J]. Journal of Mechanical Engineering, 2011, 47 (19) : 59-67.

[2] 胡芑庆,陈微鹏,程哲,等.基于经验模态分解和深度卷积神经网络的行星变速器故障诊断方法[J].机械工程学报,2019,55(7):9-18.
HU Q Q, CHEN H P, CHENG ZH, et al. A planetary gearbox fault diagnosis method based on empirical mode decomposition and deep convolutional neural network[J]. Journal of Mechanical Engineering, 2019,55(7) : 9-18.

[3] 李肖,冯志鹏.行星变速器齿轮组合故障振动频谱特征[J].振动与冲击,2020,39(1):15-23.
LI X, FENG ZH P. Vibration spectrum characteristics of planetary transmission gear combination fault[J]. Journal of Vibration and Shock, 2020,39 (1) : 15-23.

[4] 魏秀业,程海吉,贺妍,等.基于特征融合与 ResNet 的行星齿轮箱故障诊断[J].电子测量与仪器学报,2022,36(5):213-222.
WEI X Y, CHENG H J, HE Y, et al. Fault diagnosis of planetary gearboxes based on feature fusion and ResNet [J].

- Journal of Electronic Measurement and Instrumentation, 2022, 36 (5): 213-222.
- [5] MAQSOOD A, OSLEBO D, CORZINE K, et al. STFT cluster analysis for DC pulsed load monitoring and fault detection on naval shipboard power systems [J]. IEEE Transactions on Transportation Electrification, 2020, 6(2):821-831.
- [6] HU Q, QIN A S, ZHANG Q H, et al. Fault diagnosis based on weighted extreme learning machine with wavelet packet decomposition and KPCA [J]. IEEE Sensors Journal, 2018, 18(20):8472-8483.
- [7] YANG B Y, LIU R N, CHEN X F. Sparse time-frequency representation for incipient fault diagnosis of wind turbine drive train [J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67 (11): 2616-2627.
- [8] 张安安,黄晋英,冀树伟,等. 基于卷积神经网络图像分类的轴承故障模式识别 [J]. 振动与冲击, 2020, 39(4):165-171.
ZHANG AN AN, HUANG J Y, JI SH W, et al. Bearing fault pattern recognition based on convolutional neural network image classification [J]. Vibration and Shock, 2020, 39 (4): 165-171.
- [9] 杨仁树,付晓强,张世平,等. 基于 EEMD 分形与二次型 SPWV 分布的爆破振动信号分析 [J]. 振动与冲击, 2016, 35(22):41-47.
YANG R SH, FU X Q, ZHANG SH P, et al. Blasting vibration signal analysis based on EEMD fractal and quadratic SPWV distribution [J]. Vibration and Shock, 2016, 35 (22): 41-47.
- [10] 王亚萍,许迪,葛江华,等. 基于 SPWVD 时频图纹理特征的滚动轴承故障诊断 [J]. 振动. 测试与诊断, 2017, 37(1):115-119, 203.
WANG Y P, XU D, GE J H, et al. Rolling bearing fault diagnosis based on texture features of SPWVD time-frequency map [J]. Journal of Vibration, Test and Diagnosis, 2017, 37 (1): 115-119, 203.
- [11] 姚博文,彭喜元,于希明,等. 深度神经网络的自适应联合压缩方法 [J]. 仪器仪表学报, 2023, 44 (5): 21-32.
YAO B W, PENG X Y, YU M X, et al. Adaptive joint compression method for deep neural networks [J]. Chinese Journal of Scientific Instrument, 2023, 44 (5): 21-32.
- [12] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network [J]. arXiv preprint arXiv: 1503. 02531, 2015.
- [13] 黄震华,杨顺志,林威,等. 知识蒸馏研究综述 [J]. 计算机学报, 2022, 45(3):624-653.
HUANG ZH H, YANG SH ZH, LIN W, et al. Review of knowledge distillation [J]. Journal of Computer Science, 2022, 45 (3): 624-653.
- [14] 邵仁荣,刘宇昂,张伟,等. 深度学习中知识蒸馏研究综述 [J]. 计算机学报, 2022, 45(8):1638-1673.
SHAO R R, LIU Y, ZHANG W, et al. A review of knowledge distillation in deep learning [J]. Journal of Computer Science, 2022, 45 (8): 1638-1673.
- [15] ELSKEN T, METZEN J H, HUTTER F. Neural architecture search: A survey [J]. Journal of Machine Learning Research, 2019, 20(55):1-21.
- [16] TZELEPI M, PASSALIS N, TEFAS A. Online subclass knowledge distillation [J]. Expert Systems with Applications, 2021, 181:115132.
- [17] 王廷轩,刘韬,王振亚,等. 知识蒸馏与迁移学习的轴承故障诊断应用研究 [J]. 计算机工程与应用, 2023, 59(13):289-297.
WANG T X, LIU T, WANG ZH Y, et al. Application of knowledge distillation and transfer learning in bearing fault diagnosis [J]. Computer Engineering and Application, 2023, 59 (13): 289-297.
- [18] 曲志昱,李根,邓志安. 基于知识蒸馏与注意力图的雷达信号识别方法 [J]. 电子与信息学报, 2022, 44(9): 3170-3177.
QU ZH Y, LI G, DENG ZH AN. Radar signal recognition method based on knowledge distillation and attention map [J]. Electronics and Informatics, 2022, 44 (9): 3170-3177.
- [19] 王成林. 基于知识蒸馏和深度可分离卷积的轴承故障轻量化诊断 [J]. 噪声与振动控制, 2023, 43 (3): 139-144.
WANG CH L. Lightweight bearing fault diagnosis based on knowledge distillation and deep separable convolution [J]. Noise and Vibration Control, 2023, 43 (3): 139-144.
- [20] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South).
- [21] 周兴康,余建波. 基于深度一维残差卷积自编码网络的变速器故障诊断 [J]. 机械工程学报, 2020, 56(7): 96-108.
ZHOU X K, YU J B. Transmission fault diagnosis based on deep one-dimensional residual convolutional

autoencoder network [J]. Journal of Mechanical Engineering, 2020, 56 (7): 96-108.

作者简介



马超, 2003 年于沈阳航空工业学院获得学士学位, 2009 年于北京理工大学获得博士学位, 现为北京信息科技大学副研究员/副教授, 主要研究方向为机械系统动力学与检测诊断、智能康复机械设计。

E-mail: mach2006@126.com

Ma Chao received his B. Sc. degree from Shenyang Institute of in 2003 and Ph. D. degree from Beijing Institute of Technology in 2009, respectively. Now he is an Associate Professor in Beijing Information Science and Technology

University. His main research interests include mechanical system dynamics and detection diagnosis, intelligent rehabilitation machinery design.



郑鑫辉, 2018 年于北京信息科技大学获得学士学位, 现为北京信息科技大学硕士研究生, 主要研究方向为机械系统故障诊断。

E-mail: 15001053343@163.com

Zheng Xinhui received his B. Sc. degree from Beijing Information Science & Technology University in 2018. Now he is a M. Sc. candidate at Beijing Information Science & Technology University. His main research interest includes fault diagnosis of the mechanical system.