

DOI: 10.19650/j.cnki.cjsi.J2310980

多传感器融合下多工况刀具磨损状态 预测的深度森林方法研究*

汪鑫¹, 廖小平², 刘树胜¹, 覃办¹, 鲁娟¹

(1. 北部湾大学广西海洋工程装备与技术重点实验室 钦州 535011; 2. 广西大学广西制造系统与
先进制造技术重点实验室 南宁 530004)

摘要:准确监测加工过程刀具磨损状态有助于避免因刀具失效导致的产品质量问题。建立不同工况的刀具磨损监测模型,往往需要对每组工况调参以保证精度。为减少调参并保证预测精度,结合深度森林的超参数少、参数对模型不敏感和训练过程自适应等优点,利用深度森林建立了多传感器信号及多工况下自主特征选择的刀具磨损状态预测模型。基于3组不同工艺参数下TC18铣削过程的多传感器及磨损数据,以及预测与健康管理(PHM)学会2010年高速数控机床刀具健康预测竞赛的开放数据,深度森林在3组工况的预测精度分别为95.35%、96.63%和97.06%,在PHM数据上为98.95%,验证了深度森林对多工况下刀具磨损预测的高精度和适用性,为在线监测技术提供了有力的指导。

关键词: 深度森林; 刀具磨损状态; 多传感器; 多工况

中图分类号: TH165+.3 TP18 文献标识码: A 国家标准学科分类代码: 510.40

Research on the deep forest method for tool wear state prediction under multiple working conditions with multi-sensor fusion

Wang Xin¹, Liao Xiaoping², Liu Shusheng¹, Qin Ban¹, Lu Juan¹

(1. Guangxi Key Laboratory of Ocean Engineering Equipment and Technology, Beibu Gulf University, Qinzhou 535011, China;
2. Guangxi Key Laboratory of Manufacturing System and Advanced Manufacturing Technology,
Guangxi University, Nanning 530004, China)

Abstract: Accurate monitoring of tool wear during machining helps to avoid product quality problems caused by tool failure. To formulate tool wear monitoring models for different working conditions, it is necessary to adjust the parameters for each group of working conditions to ensure the accuracy. To reduce the number of parameter adjustment and ensure the prediction accuracy, the advantages of deep forest are combined, such as few hyperparameters, parameter insensitivity to the model and adaptive training process. A tool wear state prediction model with multi-sensor signals and autonomous feature selection for multi-conditions is established by using deep forest. The multi-sensor and wear data of TC18 milling process under three sets of different process parameters, and the open data in the predictive and health management (PHM) society 2010 high-speed CNC machine tool health prediction competition are utilized. For the three sets of working conditions, the prediction accuracy values of deep forest are 95.35%, 96.63% and 97.06%, respectively, and 98.95% on PHM data, which evaluate the high accuracy and applicability of deep forest for tool wear prediction under multiple working conditions. It provides a strong guidance for online monitoring technology in intelligent machining technology.

Keywords: deep forest; tool wear state; multi-sensor; multiple working conditions

0 引言

随着传感器、人工智能、大数据和工业物联网等技术的发展,传统制造向智能制造转型,因此,智能加工技术显得尤为重要。在加工过程中,刀具磨损不可避免,以往的研究结果表明,10%~40%的机床总停机时间是由刀具异常状态造成^[1]。而且刀具磨损的不断累积会对加工表面的粗糙度、尺寸精度、平整度等产生不利影响。因此,准确预测刀具磨损状态有利于指导换刀,对提高质量及节约成本具有重要意义。

国内外学者围绕刀具磨损监测做了许多研究,通过采集刀具磨损信息的方式不同,其监测技术可分为直接监测和间接监测^[2]。直接监测方法难以实现刀具磨损的实时在线监测,在实际工程中应用较少^[3]。间接监测是通过传感器采集与刀具磨损相关的物理信号,利用数据驱动模型建立信号的隐含特征与刀具磨损间的映射关系监测刀具磨损状态^[4]。相比直接监测,间接监测具有更好的灵活性和实用性,是目前主要的刀具磨损监测方法。

在间接监测方法中,切削力、振动、声发射、温度、电流和声音等信号都可作为监测信号,从不同层面上反映刀具磨损情况,其中力、振动和声发射信号在刀具磨损监测上较为常用。多传感器信号具有互补性、协同性等特点,且不同传感器可从不同维度对信息进行全面监测。因此,多传感器组合是常采用的数据采集方式。Wu等^[5]从切削力、振动和声发射信号中提取统计特征,利用随机森林(random forest, RF)建立刀具磨损预测模型,取得了较好的预测精度。Wang等^[6]从声音、加速度和切削弯矩信号中提取铣削过程的多传感器特征,提出改进的主成分分析进行特征选择,利用支持向量机(support vector machines, SVM)准确识别了颤振和刀具磨损。He等^[7]提取振动、切削力和声发射信号的时域、频域和时频域特征,通过相关性分析确定最优多传感器特征,利用反向传播神经网络(back propagation neural network, BPNN)建立了准确度高的刀具磨损预测模型。何彦等^[8]将归一化的力、振动、声发射信号输入长短时记忆卷积神经网络,在线监测刀具磨损,并验证了模型的有效性。

分析上述文献可知,利用传统机器学习方法(如SVM、RF和BPNN等)监测刀具磨损时,由于传统机器学习方法数据处理效率和能力有限,不具备或具有较弱的特征选择能力,预测刀具磨损时,需额外的特征选择(如最大信息系数(maximal information coefficient, MIC)^[9]、遗传算法(genetic algorithm, GA)^[10]、基于交叉验证的递归特征消除(recursive feature elimination cross validation, RFECV)^[11]等)步骤来避免“维数灾难”,并获得与刀具模型相关度高的特征,提高预测效率及精度。但一些特征选择方法会依

赖专家知识,存在一定的不确定性。深度神经网络(deep neural network, DNN)具有深度模型的高级特征工程能力、良好的数据处理效率及能力,能从原始信号中自动学习并选择与输出变量相关的特征,可不需额外的特征选择操作,能有效减少对专家知识的依赖。但DNN的学习和预测性能依赖于超参数(学习率、正则化参数和小批量处理等)的调节,且其参数较多,调参过程复杂。

另外,当前利用间接方法进行的刀具磨损监测研究主要集中在单一工况,针对多工况下刀具的磨损情况研究的不多,而不同工况下刀具磨损的情况是不同的。当工况发生变化,选择的特征及预测模型可能都不适合新的工况^[12],需重新调参,耗时耗力。因此,寻求适用多工况、具有自主特征选择且调参简单的学习方法来预测刀具磨损是必要的。

多粒度级联森林(multi-grained cascade forest, gcForest)是2017年提出的一种非神经网络方法的深度森林(deep forest, DF)模型,该模型以RF为基础,采用类似DNN的深层结构建立,与DNN一样能自主进行特征选择^[13]。此外,DF的超参数比DNN少得多,对参数的调节不敏感,级联层数随训练过程自适应调节,并且可以很好地处理小规模训练数据。目前,DF已经成功应用于图像处理^[14]、情绪识别^[15]、生物工程^[16]、癌症疾病诊断^[17]、推荐系统^[18]和故障诊断^[19]等领域,但利用DF监测刀具磨损的研究很少。

综上,为寻求少调参且高精度的适应于多工况的刀具磨损状态预测模型,本文融合提取的切削力、振动和声发射的多传感器信号特征,利用DF实现多传感器融合的多工况刀具磨损状态预测。基于3组不同铣削参数的多传感器及磨损数据,对多工况下深度森林模型的预测能力(训练自适应性、参数敏感性及预测精度、小训练样本预测能力、稳定性及特征选择能力)进行分析和验证。此外,在预测与健康管理(prognostic and health management, PHM)2010年刀具磨损公开数据集上进一步验证了深度森林不同工况下的预测性能。构建了基于深度森林的多传感器融合下多工况刀具磨损状态预测方案,为多工况刀具磨损监测方法选择提供了指导。基于刀具磨损监测,对多工况下深度森林的预测性能进行了系统的分析和验证,确定了深度森林用于多工况下刀具磨损监测的可行性和优异性。

1 基于深度森林的刀具磨损预测框架及方法

1.1 刀具磨损预测框架

本文采用常用的切削力、振动和声发射信号表征刀具磨损,构建基于深度森林的多传感器多工况的刀具磨损状态预测框架(图1),其预测过程如下。

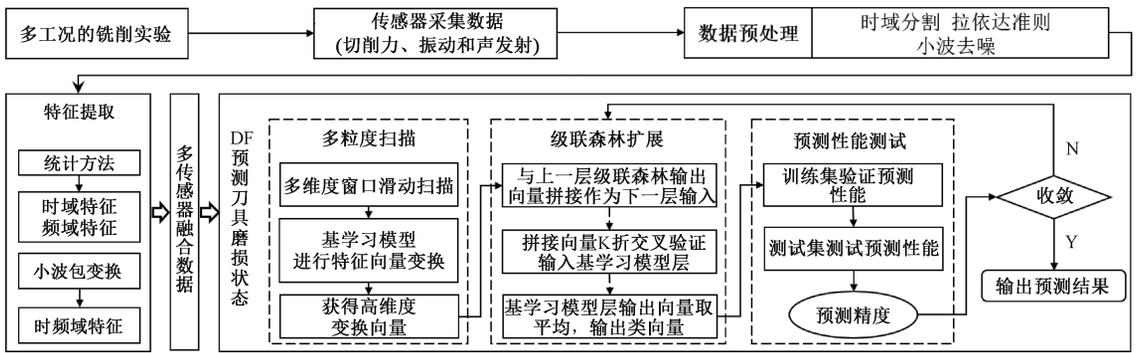


图1 基于深度森林的多工况的刀具磨损状态预测框架图

Fig. 1 Tool wear state prediction frame diagram based on deep forest with multiple working conditions

1) 搭建铣削实验平台采集多工况加工过程的切削力、振动和声发射信号,并对采集的多传感器数据进行预处理:采用时域分割选取每次测量信号的中间稳定铣削阶段作为后期的预测信号,并使用拉依达准则处理异常值和小波去噪去除环境干扰噪声信号。

2) 为提高模型的计算效率,提取每个工况下多传感器的时域和频域统计特征,及基于小波包变换的时频域特征。

3) 融合每个工况下提取的多传感器特征,利用DF构建每个工况的刀具磨损状态预测模型。

1.2 深度森林模型

在gcForest中,多粒度扫描结构的设计灵感来源于卷积神经网络对图像数据和循环神经网络对序列数据的特征处理,其有利于对级联结构的性能进一步增强^[13]。对多粒度扫描结构输入 P 维原始特征向量,可通过 M 维、 Q 维和 W 维等不同的多维度窗口。以通过 M 维窗口进行逐步扫描为例,将得到 N 个 M 维子样本,若扫描步长为 s ,则 $N = \lfloor (P - M) / s + 1 \rfloor$, $\lfloor \cdot \rfloor$ 表示下取整运算^[19]。经过滑动窗口扫描得到的特征分别输入gcForest的基模型(如RF和完全随机森林(completely random forest, CRF))中,每个森林输出 N 个 C 维的特征向量,其中 C 维分别对应数据集 C 类标签的概率,即为 C 类分类问题。因此,对于原始输入特征向量,经过基模型后会分别输出生成两个 $C \times N$ 的特征矩阵。最后,将两个特征矩阵的每一列连接起来,得到一个长度为 $2 \times C \times N$ 的增强特征向量作为多粒度扫描的输出。

在级联森林结构中,多粒度扫描的输出作为级联森林的输入。级联森林是多层的,级联的每一层都是由决策树组成的森林集合,故级联森林是基于决策树的集成。基模型的多样性有利于提高集成学习精度,为了增强多样性,级联森林每一层采用RF和CRF组成。每个级联级别接收上一级处理的特征信息和原始输入特征信息,并将处理结果输出到下一级。当扩展的新层预测性能没

有显著提升时,级联的训练过程停止,因此,级联的层数是模型根据输入特征的训练效果自适应确定的。基于级联森林的预测模型的实质为级联的级联,且每个级联由多个级别组成,每个级别对应一个滑动窗口的多粒度扫描^[13]。最后一层得到12维类向量(3分类问题),对12维类向量对应类的概率值求平均得到一个三维类向量,即3个类的概率值,然后将最大值对应的类作为最终的预测结果。

2 实验设计与数据处理

2.1 实验设计

刀具切削过程是个复杂的过程,影响刀具磨损的因素有很多,主要包括工件材料、刀具材料及结构、切削参数、切削方式及加工环境等。在这些影响因素中,切削参数是最容易被选择和调控来保证加工要求的加工质量。因此,本实验选择的刀具磨损影响因素为切削速度、每齿进给量、切削宽度和切削深度,研究的多工况也主要针对不同的加工参数组合工况。TC18钛合金是一种广泛应用于航空航天、造船和武器工业的难加工材料^[20]。相比普通材料,准确预测难加工材料加工过程的刀具磨损,科学指导换刀和参数调控,意义更重大也更迫切。因此,实验工件材料选择为TC18。实验机床为大连机床厂生产的立式加工中心(VDL-600A),铣刀直径为16mm,刀片为型号APMT1135的硬质合金刀片。工件加工面的尺寸为60mm×80mm,其中60mm为进给方向的长度。

铣削过程中采用切削力、振动和声发射传感器进行数据采集,实验现场如图2所示。三向铣削力采用压电式三向测力系统YDCB-III05B测量,力传感器安装在装夹工件的虎钳下面,采样频率为20000Hz; x 和 y 方向的振动信号由DH5901压电式加速度测量系统采集,加速度传感器安装在夹具的正面与侧面,采样频率为20000Hz;声发射信号采用PCI-8声发射检测系统进行采集,将声发

射探头贴合在工件正面。采用 AXIO SCOPE A1 蔡司金相显微镜测量刀具后刀面的磨损值。依据刀具制造商推荐的参数范围,本文选择的3个工况的铣削参数如表1所示。



图2 实验现场环境

Fig. 2 Environmental of the experimental site

表1 刀具磨损实验的铣削参数

Table 1 Milling parameters of tool wear test

工况序号	切削速度/ ($m \cdot min^{-1}$)	每齿进给量/ ($mm \cdot r^{-1}$)	切削宽度/ mm	切削深度/ mm
C1	100.48	0.06	3	0.1
C2	75.36	0.05	5	0.1
C3	75.36	0.04	3	0.4

刀具磨损阶段一般划分为初期磨损、正常磨损和剧烈磨损。在实际加工中,不同工况下刀具的磨损速率可能不一致,故其划分的阶段数值可能也不一样。文献[21-22]表明,磨损阶段的划分主要通过刀具磨损曲线的趋势和磨损变化率。因此,本文结合刀具磨损曲线的变化趋势和前后磨损值的变化率对磨损阶段进行划分。具体为:首先通过磨损曲线确定大致的分段区域,然后选择前后磨损值变化率最大的位置作为最终划分阶段位置。对磨损阶段进行划分后,得到每个工况3个阶段对应的样本数如表2所示。由表2可知,3个工况因切削参数的综合影响,其3个磨损阶段的样本数是不同的。

2.2 数据预处理及特征提取

为保证提取数据的稳定性,基于1.1节的数据预处理过程,选取每次采集的中间平稳阶段的信号作为后期特征提取信号。拉依达准则异常值处理过程为将信号中大于整体值3倍标准差的数据视为异常值进行剔除,用前一个非异常值和后一个非异常值的平均值填充剔除的值。由于实验使用PCI-8声发射检测系统对声发射的特征参数进行实时提取,并利用AEwin软件进行声发射采集参数的设置,可减少噪声干扰,故未对采集到的声发射特征参数进行数据预处理。

表2 C1、C2和C3的刀具磨损阶段划分

Table 2 Tool wear stage division of C1, C2 and C3

工况序号	测量次数	磨损状态	磨损值范围/ μm	样本数
C1	1~16	初期磨损	0~49	16
	17~254	正常磨损	49~88	238
	255~285	剧烈磨损	88~ $+\infty$	31
C2	1~13	初期磨损	0~71.2	13
	14~271	正常磨损	71.2~112.5	258
C3	272~296	剧烈磨损	112.5~ $+\infty$	25
	1~35	初期磨损	0~78.2	35
	36~177	正常磨损	78.2~145	142
	178~224	剧烈磨损	145~ $+\infty$	47

为从与刀具状态密切相关的信号中获取特征参数,给监测模型提供可靠的输入,从时域、频域和时频域上,全面提取了3组工况刀具的切削力和振动信号的刀具磨损特征参数。声发射信号特征则直接选用AEwin数据采集系统提取的5个特征(幅值、能量、计数、持续时间和上升时间)。

为表征切削力与振动信号的时域特性,提取均值、方差、标准差、均方根、偏度、峭度、波形因子、脉冲因子和峰值因子9个时域统计量作为力和振动信号的特征。三向切削力和两向振动信号提取时域特征共45个。在功率谱上对信号提取重心频率、均方频率、均方根频率和频率方差4个频域特征,共获得20个切削力和振动信号频域特征。利用小波包分解对切削力和振动信号进行时频域特征提取。采用“db3”小波基函数进行3层小波包分解,共提取到切削力和振动信号时频域特征40个。

每个工况下,切削力和振动信号的时域、频域和时频域共105个,再加上声发射5个特征。因此,从3种传感器信号中总计提取110个特征。

3 DF 刀具磨损预测效果评价及分析

3.1 DF 多工况模型的参数敏感性和预测精度分析

以特征提取的110个特征为输入,对应的刀具磨损为输出,利用DF分别构建C1、C2和C3的刀具磨损状态预测模型。每个工况样本数量的70%和30%分别作为训练集和测试集,预测精度评价指标为准确率。

基于参考文献[23]中DF的参数设置以及本文针对的刀具磨损状态预测的3分类问题,对DF模型中的多粒度扫描和级联森林的结构进行如下设置:多粒度扫描设置3个窗口,其大小分别为 1×2 维, 1×4 维, 1×6 维;多粒度扫描的基学习器(分类器)采用一个CRF和一个RF,

最大深度设置为 10,并采用 3 折交叉验证获得森林输出的类向量;级联森林结构每层分类器为两个 CRF 和两个 RF,森林的最大深度为每棵树生长为纯叶节点后的深度,采用 5 折交叉验证进行训练。

由于集成学习的多样性对整个模型的构建有着重要意义,每一种森林数目不一致会导致整个模型预测效果的不平衡性。因此,为了更好的体现 DF 模型性能,并简化参数设置过程,本文将两种基学习模型(CRF 和 RF)森林的数量和森林中树木的数量设置为相同。确定多粒度扫描和级联森林的结构后,对每个森林中的决策树树木棵数的设置进行测试,以 50 棵为基础,并以 50 棵为步长累加,进行 10 次预测,结果如图 3 中。从图 3 可知,在 10 组预测中,树木棵数对 C3 的预测精度无影响,C2 的预测精度只在 100 棵树木时发生变化,C1 的预测精度随着树木棵数的增加在两个精度切换,这体现了 DF 对内部参数不敏感的特点。此外,随着决策树棵数的增加,整个模型的内存消耗和训练时间会增加。因此,权衡预测精度与时间,本文将每个森林的树木数量设置为 200。

确定 DF 模型结构后,预测 3 个工况下的刀具磨损状态,获得的分类精度分别为 95.35%、96.63% 和 97.06%。表 3 为 DF 在 3 种工况下获得的自适应层数,预测的 3 组工况每个磨损状态精度的混淆矩阵如图 4 所示。图 4 中

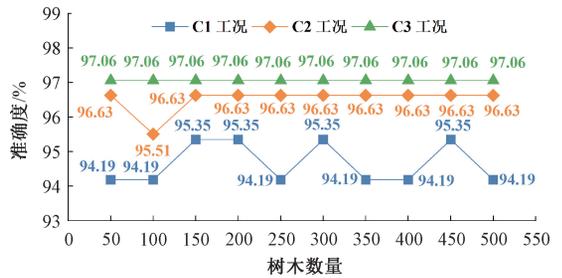


图 3 C1、C2 和 C3 工况的 DF 在不同树木棵数的准确度
Fig. 3 Prediction accuracy of DF in C1, C2 and C3 working conditions under different tree numbers

每个小方块的横坐标对应刀具的真实磨损状态,纵坐标对应 DF 模型预测的刀具磨损状态,方块中的数字表示模型对每个状态的预测精度,括号中的数字表示每个状态的预测样本数量。

表 3 C1、C2 和 C3 下 DF 的自适应层数

Table 3 Adaptive layers of DF under C1, C2 and C3

工况序号	C1	C2	C3
层数	1	2	2

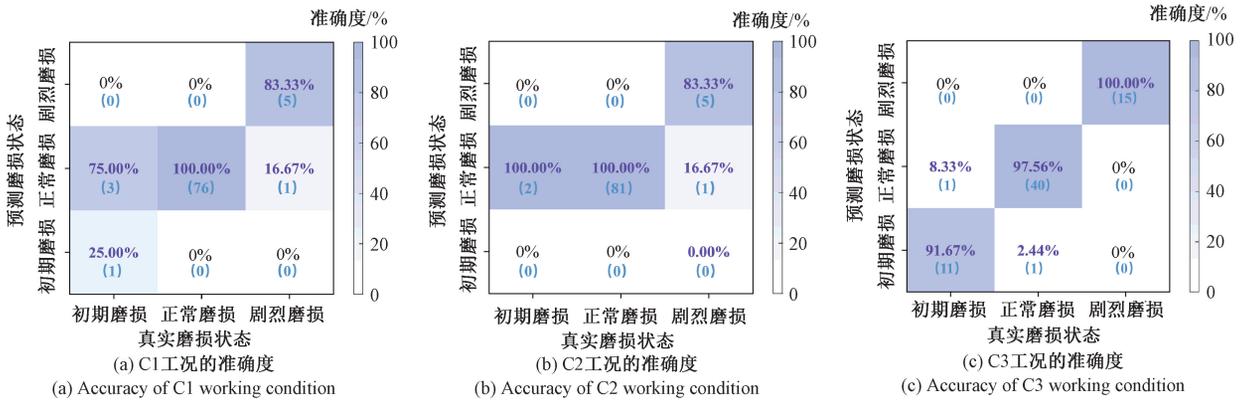


图 4 DF 在 C1、C2 和 C3 工况的预测结果混淆矩阵

Fig. 4 DF prediction results confusion matrix at C1, C2 and C3 working conditions

图 4 中,C1、C2 和 C3 的正常磨损和剧烈磨损的分类精度都较高,其中 C1 和 C2 的正常磨损及 C3 的剧烈磨损分类精度都高达 100%;初期磨损的分类精度除 C3 为 91.67% 外(预测总样本为 12 个,预测正确的样本为 11 个),C1 和 C2 都较低。这主要是因为初期磨损阶段的训练样本过少,导致模型对初期磨损知识没有得到充分的学习和训练。实际生产过程中,刀具初期磨损阶段较短,且加工前会选择合适的切削参数加工,此阶段的加工质量和性能相对稳定。另外刀具的换刀发生在剧烈磨损阶段,准确预测刀具的剧烈磨损阶段,可科学指导换刀

时间。因此,相对识别初期磨损阶段,准确识别正常磨损和剧烈磨损阶段对刀具磨损监测实际应用的意义更显著。基于以上分析,可认为构建的刀具磨损状态预测框架及采用的方法是有效的。

3.2 DF 在小训练样本的预测能力及 DF 的稳定性分析

新刀具从开始加工到剧烈磨损阶段会耗费较长的加工时间,频繁采集传感器数据和测量磨损会增加人力和时间成本,数据处理成本也会增加。寻求小训练样本下高精度的预测模型,可有效减少数据采集和处理成本。因此,本文对 DF 在小训练样本上的预测性能进行测试

和分析。对3组工况磨损数据下训练样本占比为20%、30%、40%、50%、60%和70%的6种情况分别进行性能测

试,并对比DF、RF、SVM和BPNN的分类精度,对应的结果如图5所示。

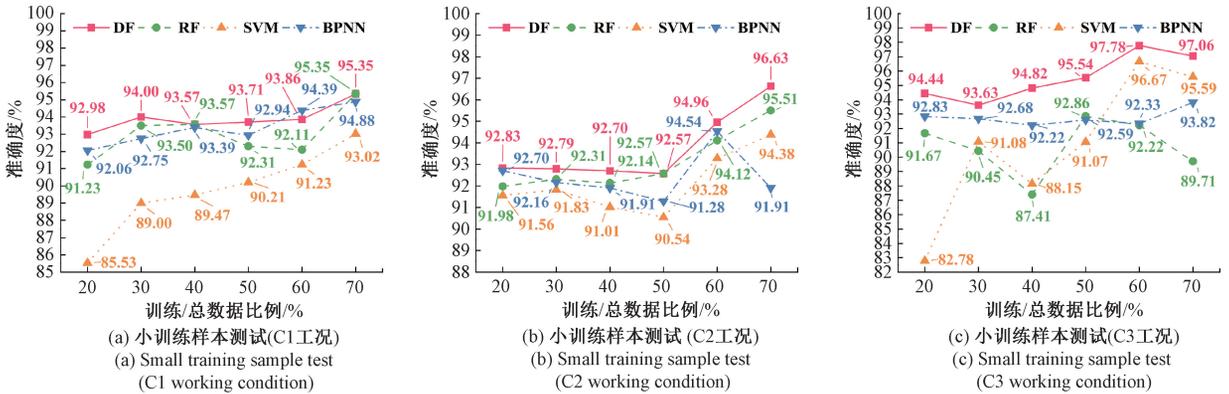


图5 C1、C2和C3工况下训练集占比为20%~70%时的预测精度

Fig. 5 Prediction accuracy of C1, C2 and C3 working conditions with training set accounting for 20%~70%

预测模型在Python平台中构建,RF、SVM和BPNN调用Python平台中相应的函数实现。SVM采用“RBF”核函数,惩罚系数设置为1,其余参数均采用默认参数。RF的树木数量为20,最大深度为3,其余参数为默认参数。BPNN设置两个隐藏层,第1层的神经元数目为20,第2层神经元数目为10,学习率为0.01,网络迭代训练100次。DF的结构及参数设置与3.1节相同。从图5可知,在3组工况下的训练样本占比为20%~70%的6种情况中,DF获得的3个工况的刀具磨损分类精度都在92%以上,且训练样本占比为20%、30%、40%、50%、70%的情况预测精度优于RF、SVM和BPNN。因此,DF获得的刀具磨损预测精度和稳定性都优于RF、SVM和BPNN。

为进一步测试基于DF的刀具磨损状态预测模型的稳定性,从3种工况的样本中随机选取30%作为测试集,运行10次,利用DF进行预测并记录预测精度(图7)。图8所示为3种工况下DF随机运行10次的层数。图9所示为图7中预测精度的最大值、最小值、平均值和标准差。

进一步分析图5可知,对3组工况下不同训练样本占比的预测模型,DF、RF、SVM和BPNN都没有重新调参,RF、SVM和RF并不能在每次预测中都获得优秀的预测精度,而DF由于在每次训练的中都能自适应的确定合适层数(图6),故可保持良好的预测精度。综上可知,DF可在小训练样本条件下得到很好的训练,表现出良好的预测精度和稳定性。

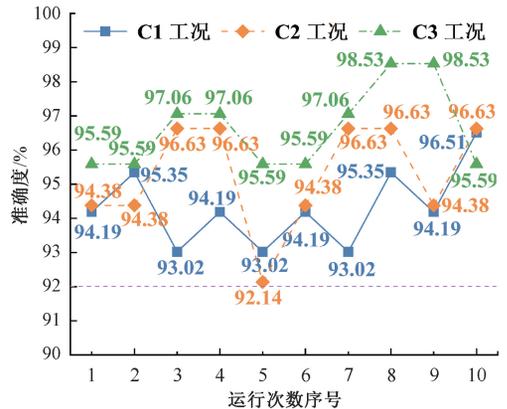


图7 DF在C1、C2和C3上随机运行10次的准确度

Fig. 7 Accuracy of DF running 10 times randomly on C1, C2 and C3

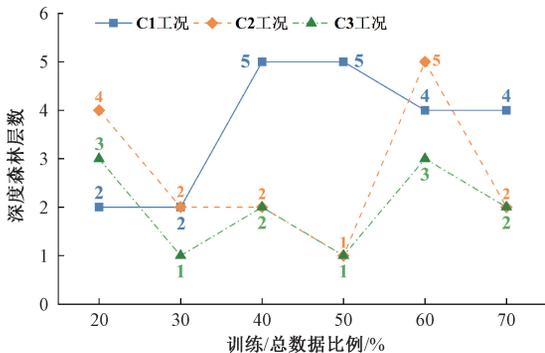


图6 C1、C2和C3工况不同训练集占比的DF层数

Fig. 6 DF layers with different training set accounting under C1, C2, and C3 working conditions

3.3 DF的特征选择能力分析

DF具有自动选择特征的能力,在进行刀具磨损状态预测时,可不进行额外的特征选择操作。为分析DF的特征选择能力,本文选择常用的特征选择方法(RFECV、MIC和GA)和传统机器学习方法(RF、SVM和BPNN)构建刀具磨损状态的预测模型,与DF的预测模型进行对比。RF、SVM和BPNN的关键参数设置与3.2节相同。在RFECV中,决策树作为基学习器,并利用3折交叉验证计算平均准确度;设置MIC阈值为0.7,即选取与刀具

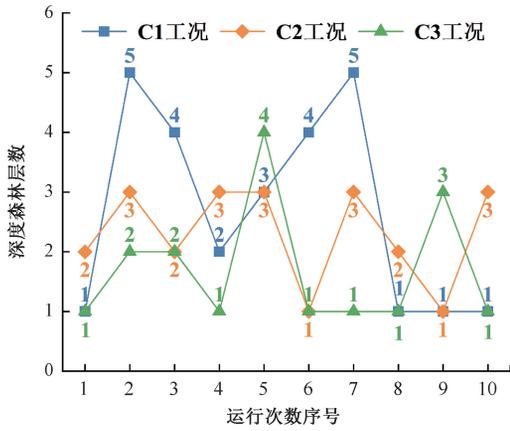


图8 DF在C1、C2和C3上随机运行10次的层数
Fig. 8 Layers of DF running 10 times randomly on C1, C2 and C3

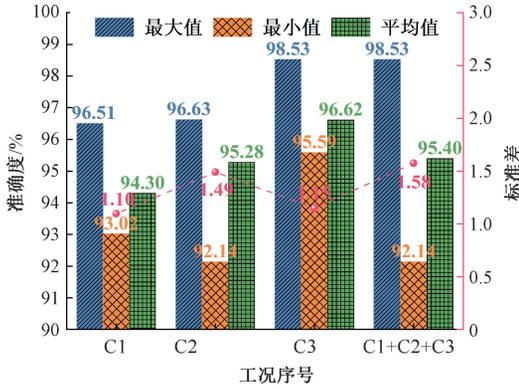


图9 DF在C1、C2、C3和整体的准确度的评价指标
Fig. 9 Evaluation metrics of DF accuracy in C1, C2, C3 and the whole

磨损相关信息系数大于0.7的特征;GA的种群大小、交叉率、变异率和迭代次数分别为设置为28、0.8、0.001和40,皮尔逊系数作为适应度函数。利用RFECV、MIC和GA分别对3组工况提取的110维特征进行特征选择,获得与刀具磨损相关度高的最优特征子集。

对于C1、C2和C3工况的数据,DF、RF、SVM和BPNN不采取额外特征选择操作的预测精度,和RF、SVM和BPNN分别基于3种特征选择方法后的预测精度结果如表4所示,模型的训练集都为70%。

从表4可以看出,在C1、C2和C3工况数据上,相比RF、SVM和BPNN,DF都获得了最高预测准确度,分别为95.35%、96.63%和97.06%。

针对C1、C2和C3工况,进行额外的特征提取操作的模型共27个,其中对比3个传统的机器学习模型(RF、SVM和BPNN),预测精度提升的模型有14个,

表4 在C1、C2、C3工况测试集上的预测精度对比
Table 4 Comparison of prediction accuracy on test set of C1, C2 and C3 working conditions %

模型	C1	C2	C3
RF	95.35	95.51	89.71
RFECV-RF	93.02	95.51	94.12
MIC-RF	93.02	96.63	91.18
GA-RF	91.86	95.51	94.12
SVM	93.02	94.38	95.59
RFECV-SVM	94.19	94.38	95.59
MIC-SVM	94.19	96.63	95.59
GA-SVM	94.19	95.51	94.12
BPNN	94.88	91.91	93.82
RFECV-BPNN	93.14	92.36	91.91
MIC-BPNN	93.37	95.96	95.44
GA-BPNN	95.23	92.36	93.09
DF	95.35	96.63	97.06

精度保持不变的有5个。因此,整体而言,进行额外的特征提取操作可提高传统机器学习方法的预测精度及效率,但同时也会使预测精度存在一定程度的不确定性。DF模型可自主进行特征选择,获得高的预测精度,避免了额外的特征选择操作带来的不确定性问题,同时展示了良好的特征选择能力。

3.4 DF模型的预测性能在PHM数据上的验证

为了进一步验证DF刀具磨损状态识别模型在多工况下的预测性能,本文选取PHM 2010年高速数控机床刀具健康预测竞赛的第1组开放数据^[24]测试。

基于3.1节确定的深度森林内部参数,对PHM数据进一步测试DF在不同刀具结构和材料、工件材料及加工参数组合下的小训练样本预测能力、稳定性及特征选择能力。使用2.1节中的磨损阶段划分方法来划分PHM数据的磨损阶段,初始磨损阶段为第1~29个样本(0~75.94 μm),正常磨损阶段为第30~225个样本(75.94~125.1 μm),剧烈磨损阶段为第226~315个样本(125.1~+∞ μm)。采用与2.2节相同的数据预处理和特征提取方法,共获得127维特征,包括63维切削力特征、63维振动特征和1维声发射特征。

图10所示为DF、RF、SVM和BPNN在PHM数据下训练样本占比为20%~70%的6种情况的分类精度。

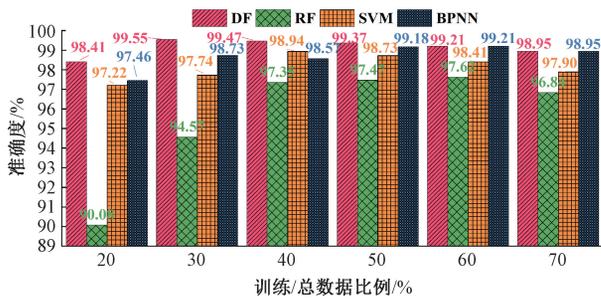


图10 PHM数据下训练集占比为20%~70%时的预测精度

Fig. 10 Prediction accuracy of PHM data with training set accounting for 20% ~ 70%

DF的结构及参数设置与3.1节相同,RF、SVM和BPNN的参数设置与3.2节相同。从图10可以看到,DF在训练集占比为20%~70%的6种情况下的预测精度都达到98%以上,且大部分情况其预测精度都高于RF、SVM和BPNN。由此,表明DF模型在小训练样本下可得到有效训练,具有良好的预测精度和稳定性。

从315组样本中随机选取30%作为测试集,运行10次进行稳定性测试,得到DF每次的预测精度如图11所示。图11中10次的预测精度都在96.5%以上,表明DF在PHM数据工况下也具有较高稳定性。

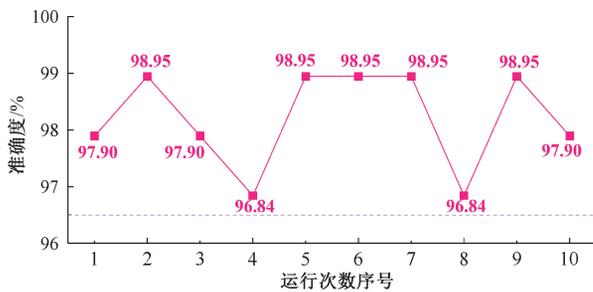


图11 DF在PHM数据上运行10次的准确度

Fig. 11 Accuracy of DF running 10 times on PHM data

使用与3.3节相同的特征选择方法(RFECV、MIC和GA)对PHM数据的特征进行选择,将得到的最优特征子集分别输入RF、SVM和BPNN中得到对比结果如表5所示。

由表5可以看出,在PHM数据上,DF除了与GA-SVM和BPNN达到了相同的预测精度(98.95%)外,比其他模型上的预测精度都高,BPNN也是没有经过特征选择直接进行预测的,但在3.3节的实验数据中,其获得的预测精度没有DF高。由此,进一步验证了DF良好的特征选择能力,没有经过特征选择步骤也能获得高的预测精度。

表5 在PHM数据测试集上的预测精度

Table 5 Prediction accuracy on test set of PHM data

模型	准确度/%
RF	96.84
RFECV-RF	97.90
MIC-RF	95.79
GA-RF	95.79
SVM	97.90
RFECV-SVM	97.90
MIC-SVM	97.90
GA-SVM	98.95
BPNN	98.95
RFECV-BPNN	97.90
MIC-BPNN	98.21
GA-BPNN	98.11
DF	98.95

4 结论

为寻求适应于多工况的少调参、高精度的刀具磨损状态预测模型,本文针对多工况下多传感器和磨损数据,提取的切削力、振动和声发射的时域、频域和时频域特征,利用深度森林分别构建适应多工况的刀具磨损状态预测模型,并分析及验证了DF模型的训练自适应性、参数敏感性、预测精度、小训练样本预测能力、稳定性及特征选择能力。

本文对3组不同加工参数组合,利用DF不同内部参数(树木棵树)下的模型预测3组工况的磨损状态,验证了DF模型对内部参数的不敏感性;在选定的树木棵树及自适应确定的层数下,3组工况的预测精度分别为95.35%、96.63%和97.06%,表明DF模型无需重新调参,便可高精度的预测不同工况下的刀具磨损。利用DF分别预测训练样本占比为20%~70%的6种情况下的磨损状态,3组工况磨损状态分类精度都在92%以上,且结果比RF、SVM和BPNN稳定,表明了DF在小训练样本下具有良好的预测精度;对3组工况的测试集随机划分10次,并运行DF模型,准确度均在92%以上,验证了DF模型的稳定性。对比特征选择方法(RFECV、MIC和GA)和机器学习方法(RF、SVM和BPNN)构建的磨损状态识别模型,DF模型均优于其他模型。表明了DF模型良好的特征选择能力。使用PHM公开数据集测试DF模型在变工况(不同工件材料、刀具材料和结构)的预测性能,进一步验证了DF在不同工况下优秀的预测性能。

本文测试了不同加工参数组合等多工况下 DF 预测刀具磨损的效果,但实际工程中的多工况是一个极为复杂的加工环境,常见多工况还包括不同铣削方式和不同工件形状等。因此,后期将研究不同切削方式及工件形状等多工况,系统性探索复杂工业环境下 DF 模型的磨损状态识别能力。

参考文献

- [1] LIU C, LI Y, HUA J, et al. Real-time cutting tool state recognition approach based on machining features in NC machining process of complex structural parts [J]. The International Journal of Advanced Manufacturing Technology, 2018, 97(1): 229-241.
- [2] 刘献礼,李雪冰,丁明娜,等. 面向智能制造的刀具全生命周期智能管控技术 [J]. 机械工程学报, 2021, 57(10): 196-219.
- LIU X L, LI X B, DING M N, et al. Intelligent management and control technology for the whole life cycle of cutting tools for intelligent manufacturing [J]. Journal of Mechanical Engineering, 2021, 57(10): 196-219.
- [3] 陈启鹏,谢庆生,袁庆霓,等. 基于深度门控循环单元神经网络的刀具磨损状态实时监测方法 [J]. 计算机集成制造系统, 2020, 26(7): 1782-1793.
- CHEN Q P, XIE Q SH, YUAN Q N, et al. A real time monitoring method for tool wear based on depth gated cyclic unit neural network [J]. Computer Integrated Manufacturing System, 2020, 26(7): 1782-1793.
- [4] 刘会永,张松,李剑峰,等. 采用改进 CNN-BiLSTM 模型的刀具磨损状态监测 [J]. 中国机械工程, 2022, 33(16): 1940-1947.
- LIU H Y, ZHANG S, LI J F, et al. Tool wear condition monitoring using improved CNN-BiLSTM model [J]. China Mechanical Engineering, 2022, 33(16): 1940-1947.
- [5] WU D, JENNINGS C, TERPENNY J, et al. A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests [J]. Journal of Manufacturing Science and Engineering, 2017, 139(7): 071018.
- [6] WANG R, SONG Q, LIU Z, et al. Multi-condition identification in milling Ti-6Al-4V thin-walled parts based on sensor fusion [J]. Mechanical Systems and Signal Processing, 2022, 164: 108264.
- [7] HE Z, SHI T, XUAN J. Milling tool wear prediction using multi-sensor feature fusion based on stacked sparse autoencoders [J]. Measurement, 2022, 190: 110719.
- [8] 何彦,凌俊杰,王禹林,等. 基于长短时记忆卷积神经网络的刀具磨损在线监测模型 [J]. 中国机械工程, 2020, 31(16): 1959-1967.
- HE Y, LING J J, WANG Y L, et al. Tool wear online monitoring model based on long short memory convolution neural network [J]. China Mechanical Engineering, 2020, 31(16): 1959-1967.
- [9] RESHEF D N, RESHEF Y A, FINUCANE H K, et al. Detecting novel associations in large data sets [J]. Science, 2011, 334: 1518-1524.
- [10] 董慧芬,郑坤,杨占刚. 基于 GA-BRBPNN 的航空自耦变压整流器故障诊断方法 [J]. 电子测量与仪器学报, 2022, 36(9): 217-225.
- DONG H F, ZHENG K, YANG ZH G. Fault diagnosis method of aviation autocoupled variable voltage rectifier based on GA-BRBPNN [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(9): 217-225.
- [11] 刘献礼,秦怡源,岳彩旭,等. 递归特征消除与极端随机树在铣刀磨损监测中的研究 [J]. 机械科学与技术, 2023, 42(6): 821-828.
- LIU X L, QIN Y Y, YUE C X, et al. Research on recursive feature elimination and extreme random tree in milling cutter wear monitoring [J]. Mechanical Science and Technology, 2023, 42(6): 821-828.
- [12] 黄星华,吴天舒,杨龙玉,等. 一种面向旋转机械的基于 Transformer 特征提取的域自适应故障诊断 [J]. 仪器仪表学报, 2022, 43(11): 210-218.
- HUANG X H, WU T SH, YANG L Y, et al. A domain adaptive fault diagnosis for rotating machinery based on Transformer feature extraction [J]. Chinese Journal of Scientific Instrument, 2022, 43(11): 210-218.
- [13] ZHOU Z H, FENG J. Deep forest [J]. National Science Review, 2019, 6(1): 74-86.
- [14] XIA J, MING Z, IWASAKI A. Multiple sources data fusion via deep forest [C]. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 2018: 1722-1725.
- [15] CHENG J, CHEN M, LI C, et al. Emotion recognition from multi-channel EEG via deep forest [J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(2): 453-464.
- [16] YU B, CHEN C, WANG X, et al. Prediction of protein-protein interactions based on elastic net and deep

- forest[J]. *Expert Systems with Applications*, 2021, 176: 114876.
- [17] GUO Y, LIU S, LI Z, et al. Towards the classification of cancer subtypes by using cascade deep forest model in gene expression data[C]. *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine*, 2017: 1664-1669.
- [18] WEN H, ZHANG J, LIN Q, et al. Multi-level deep cascade trees for conversion rate prediction in recommendation system[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019: 338-345.
- [19] 姜万录,李满,张培尧,等. 基于全矢增强深度森林的旋转设备智能故障诊断方法[J]. *中国机械工程*, 2022,33(11):1324-1335.
- JIANG W L, LI M, ZHANG P Y, et al. Intelligent fault diagnosis method for rotating equipment based on full vector enhanced depth forest [J]. *China Mechanical Engineering*, 2022, 33 (11): 1324-1335.
- [20] MA J Y, LUO D CH, LIAO X P, et al. Tool wear mechanism and prediction in milling TC18 titanium alloy using deep learning[J]. *Measurement*, 2021, 173(1): 108554.
- [21] PAN T, ZHANG J, ZHANG X, et al. Milling force coefficients-based tool wear monitoring for variable parameter milling [J]. *The International Journal of Advanced Manufacturing Technology*, 2022, 120(7-8): 4565-4580.
- [22] WEI X, LIU X, YUE C, et al. Tool wear state recognition based on feature selection method with whitening variational mode decomposition[J]. *Robotics and Computer-Integrated Manufacturing*, 2022, 77: 102344.
- [23] 丁家满,吴晔辉,罗青波,等. 基于深度森林的轴承故障诊断方法[J]. *振动与冲击*,2021,40(12):107-113.
- DING J M, WU Y H, LUO Q B, et al. Bearing fault diagnosis method based on deep forest[J]. *Vibration and Shock*, 2021, 40(12): 107-113.

- [24] JIA X, HUANG B, FENG J, et al. A review of PHM data competitions from 2008 to 2017: Methodologies and analytics[C]. *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, 2018: 1-10.

作者简介



汪鑫,2021年于湖南城市学院获学士学位,现为北部湾大学硕士研究生,主要研究方向为机器学习的应用及过程智能监测与调控。

E-mail:2102011124@stu.bbgu.edu.cn

Wang Xin received his B. Sc. degree from Hunan City University in 2021. He is currently a M. Sc. candidate at Beibu Gulf University. His main research interests include the application of machine learning and intelligent monitoring and regulation of processes.



廖小平,2005年于华中科技大学获博士学位,现为广西大学教授,博士生导师,主要研究方向为数字化设计与智能系统开发。

E-mail:xpfeng@gxu.edu.cn

Liao Xiaoping received his Ph. D. degree from Huazhong University of Science and Technology in 2005. He is currently a professor and a Ph. D. advisor at Guangxi University. His main research interests include digital design and intelligent system development.



鲁娟(通信作者),2020年于广西大学获博士学位,现为北部湾大学副教授,硕士生导师,主要研究方向为机器学习的应用与智能系统开发。

E-mail:lujuan3623366@163.com

Lu Juan (Corresponding author) received her Ph. D. degree from Guangxi University in 2020. She is currently an associate professor and a M. Sc. advisor at Beibu Gulf University. Her main research interests include the application of machine learning and the development of intelligent systems.