doi:10.3772/j.issn.1002-0470.2022.10.010

基于深度学习的拉削刀具磨损状态识别模型①

应申舜②* 傅晨泰* 林绿胜** 吕晓敏** 张顺琦*** 易 凯③*

(*浙江工业大学机械工程学院 杭州 310032) (**浙江畅尔智能装备股份有限公司 缙云 321404) (***上海大学机电工程与自动化学院 上海 200444)

摘要 拉削是航空发动机涡轮盘榫槽的关键加工工艺,拉刀在工作过程中的异常状态若不能被及时发现并加以干预,将引起严重的破坏,目前尚无先进的人工智能方法解决这一问题。本文提出了一种拉削刀具磨损状态识别模型(DSBiLSTM),该模型基于具有特征重用优势的稠密连接网络(DenseNet)设计了DenseNet(3-2),实现空间特征提取和降维,并设计了堆叠的双向长短期记忆网络(SBiLSTM),实现时序特征提取,两者优势互补,将多尺度融合提取特征纳入到多个全连接层(FNN)和Softmax 层,实现刀具磨损状态识别。基于创新的识别模型,设计了涡轮盘榫槽拉削加工实验,采用了基于混淆矩阵的性能指标体系,将所提出的模型与单个特征提取模型进行比较。实验结果表明,DSBiLSTM 模型在刀具磨损状态预测中具有突出性能,识别准确率达到98.73%,单个样本的识别速度提高到11 ms。

关键词 深度学习;拉削加工;刀具磨损;状态识别;稠密连接网络(DenseNet);长短期记忆网络(LSTM)

0 引言

涡轮盘是航空发动机机械加工中最具有挑战性 的零件,原因包括:(1)涡轮盘是整机安全的关键, 质量标准较高;(2)部件材料是耐高温超级合金,如 铬镍铁合金718、Rene 104 或者 IN100PM,属于难加 工材料。涡轮盘毛坯经过锻造、车削、磨削和钻孔成 形,最后加工叶片安装槽,这些槽具有复杂的几何结 构和较高的公差要求。制造槽的主要工艺技术就是 拉削。若在加工过程中发生拉刀损伤(磨损、碎屑、 刀齿破裂等),会刮伤工件导致额外的制造成本^[1]。 在加工过程中纳入监测和控制,即可保证加工质量, 又能提高稳定性,因而世界各国都给予了高度重视。 涡轮盘榫槽拉刀最常见的失效方式是磨损失 效。拉削刀具磨损状态识别方法包括直接法和间接 法。直接法由数码显微镜、光学轮廓仪等测量工具 直接测量刀具磨损状态,由于测量工具价格昂贵,并 且需要在机床停机状态下进行,难以在工业环境下 广泛应用^[2]。间接法通过采集切削力^[34]、电流^[5]、 振动^[6-7]和声发射^[8-9]等传感器信号,进一步提取和 优化选择特征,并采用模糊理论^[10-11]、支持向量机 (support vector machine, SVM)^[12-14]和隐马尔科夫 模型^[15-17]等方法建立刀具磨损状态与特征之间的 映射关系模型,利用模型和实时采集的传感器信号, 实现对刀具磨损的在线识别。间接方法具有方法多 样、实时监测等优点,成为当前刀具状态监测的主要 研究方向之一。孙志娟^[15]结合皮尔逊相关分析法 和近似冗余法实现特征筛选与降维,并提出以天牛

① 国家自然科学基金(11972020)和浙江省基础公益研究计划(LGG21E050017)资助项目。

② 男,1978年生,博士,副教授;研究方向:高端装备健康监测与远程运维,机器人应用技术;E-mail:yss9701@163.com。

③ 通讯作者, E-mail:2111902279@ zjut. edu. cn。 (收稿日期:2021-03-17)

粒子群优化算法解决混合隐马尔可夫模型寻优问 题,实现刀具磨损的预测。Kong 等人^[18]提出了基 于高斯混合隐马尔可夫模型的刀具磨损估计模型, 有效地识别铣床刀具的磨损状态。Schwenzer 等 $人^{[19]}$ 研究了支持向量机和随机森林(random forest, RF)两种立铣刀磨损状态分类方法,验证了所提出 的方法在力和电流信号的分类结果对比,得出了 SVM 能够从直切实验中获得更通用的模型。上述 研究在刀具磨损状态识别方面选择机器学习相关模 型,在某些领域取得不错的效果,但是机器学习模型 依赖于前期特征提取、特征降维和特征融合等大量 先验知识与反复测试,建模耗时耗力、效率不高且容 易陷入局部最优。

随着智能制造的不断发展,深度学习在不使用 先验知识和专家经验的情况下,利用优秀的自适应 特征学习能力,在机器健康监测领域获得广泛关注, 诸如稀疏自编码器、深度置信网络、卷积神经网络 (convolutional neural network, CNN)和循环神经网络 (recurrent neural network, RNN)等深度学习算法在 刀具状态监测应用领域取得了重要进展。Shi等 人^[7]使用无监督学习将多个堆叠的稀疏自动编码 器用于刀具状态监测中,解决了超精密加工离线监 测效率低的问题。Nguyen 等人^[20]基于堆叠式自动 编码器和 Softmax 构造深度学习网络(deep learning network, DLN),实现对刀具磨损状态有效识别。卢 志远等人^[21]提出了在线监测方法,并利用卷积神经 网络建立了刀具磨损状态识别模型,该方法更加适 用于生产加工现场。Cao 等人^[22]提出了一种基于 2-D 卷积神经网络和派生小波框架(derived wavelet frames, DWFs)的高鲁棒铣削刀具磨损监测方法,绕 开了复杂且低便携性的特征工程。Cai 等人^[23]使用 堆叠的长短期记忆网络(long short-term memory, LSTM)提取多传感器时间序列中包含的抽象和深层 特征,将提取的特征与过程信息结合,提高了刀具磨 损状态识别的精度。文献[24-27]集成了深度卷积 神经网络和长短期记忆网络的优点,通过提取时间 序列中的数据依存关系,更加精准地识别刀具磨损 状态。尽管基于深度学习的刀具磨损状态识别技术 得到广泛关注,但是其自适应提取特征的能力还有

待加强,需要进一步研究。此外,上述研究大多针对 铣削加工,基于深度学习的拉削刀具磨损监测在理 论和技术上尚处在验证阶段,高精度的拉削加工刀 具状态识别技术有待深入研究。

本文采用振动信号为刀具磨损状态识别的信号 源,将卷积神经网络和循环神经网络结合,提出了一 种实现精确的拉削刀具状态识别的新方法。振动信 号包含丰富的空间特征和前后时序特征,从特征提 取的角度分析,卷积神经网络更适用于空间特征提 取,而循环神经网络更适用于时间序列特征提取。 本文借助稠密连接网络(densely connected network, DenseNet)网络和 LSTM 网络构建特征提取方法,使 得特征的输出在空间上和前后序列的输入都有联 系,同时考虑到空间和时间问题。首先,研究了基于 深度学习的特征提取理论基础,设计了 DenseNet(3-2) 网络结构和堆叠双向 LSTM 网络。然后,引入密 集层堆叠的全连接神经(fully connected neural network, FNN) 网络, 构建刀具磨损状态识别模型.并开 展模型训练。接着,设计了涡轮盘榫槽拉削加工的 刀具磨损状态识别实验,建立了基于混淆矩阵的性 能指标体系,并开展实验研究。最后,通过与单个特 征提取识别模型进行对比研究,结果表明该模型具 有较好的识别速度和准确性。

深度学习理论基础 1

1.1 基于深度残差网络的空间特征提取

CNN 具有3个关键特性,即权值共享、局部连 接和输入输出数据的结构化。卷积神经网络强调几 个相对的数据在相交和相离之间数据的上下层联 系,数据的处理在时序上相互独立。卷积神经网络 已在各种计算机视觉应用程序中取得了成功,这些 应用程序的输入数据通常为二维数据。由于振动信 号是一维数据,因此本文使用一维 CNN。对于时间 序列数据,卷积内核沿着时间轴移动以执行卷积操 作。激活功能用于通过添加偏差来处理信息,以获 得保留有用信息的特征图。池化层能够减少特征向 量的长度,从而可以进一步减少模型参数的数量,被 广泛应用于缩小特征向量。在处理时序序列时, CNN 中的一维卷积层首先在整个顺序输入多个卷 积核,通过在序列上滑动卷积核来生成特征向量。 然后,池化层从每个特征向量中提取最重要和固定 长度的特征。图1展示了一维 CNN 的局部特征提 取器框架。



另外,卷积层和池化层都可以通过堆叠方式执行。通过堆叠的方式增加网络深度,提升识别准确

率。但是随着网络深度的增加,一方面梯度消失的 问题也愈加明显,梯度更新将会以指数形式衰减,导 致部分参数不再更新。另一方面,模型过拟合,采样 过多的噪声数据,使得准确率下降。深度残差网络 刚好可以解决这些问题。DenseNet 提出了一个更激 进的密集连接机制:即互相连接所有的层,每个层都 会接受其前面所有层作为其额外的输入。通过对特 征的极致利用达到更好的效果和更少的参数。

借鉴 DenseNet 网络,设计 DenseNet(3-2) 网络 结构作为刀具状态的自适应特征提取器。图 2 给出 了 DenseNet(3-2) 网络结构,它共包含 2 个 Dense-Block,各个 DenseBlock 之间通过 Transition 连接在 一起。其中 DenseBlock 是包含多个层的模块,每个 层的特征图大小相同,层与层之间采用密集连接方 式。而 Transition 模块是连接 2 个相邻的 Dense-Block,并且通过 Pooling 使特征图大小降低。



图 2 DenseNet(3-2)网络结构

在 DenseBlock 中,各个层之间的特征图大小一 致,可以在 channel 维度上连接。随着层数增加, DenseBlock 后面层的输入维度会非常大。因此,在 DenseBlock 内部使用1×1卷积层和批量归一化层 (batch normalization,BN),如图3所示。其中1×1 卷积层用于降低通道维数和减少参数量,增加非线 性和跨通道间的信息交互能力,从而提高了网络的 表达能力。BN 用于解决网络更新过程中内部协变 量偏移的问题。



图 3 DenseBlock 密接层

1.2 基于长短期记忆网络的时间特征提取

循环神经网络对于数据处理侧重于序列的顺 序,当前的数据可能对未来的数据发展产生影响,也 可能与之前的数据有关联。常规的循环神经网络的 主要目的是利用历史信息来辅助当前决策,但当循 环体过长时,会出现梯度消失和梯度爆炸等问题。 另一方面,RNN决策主要还是依赖最后输入的一些 信号,更早之前的信号会随着时间的推迟而强度变 得越来越低,辅助作用变得越来越弱,网络无法学习 和利用较久之前的信息。为了解决梯度消失、梯度 爆炸和长期依赖的问题,通常使用长短时记忆网络。 LSTM 不仅能够敏感地应对短期信息,而且能够对 有价值的信息进行长期记忆,以提升网络的学习能 力。

LSTM 具有 RNN 相同的输入和输出,但是 LSTM 引入了门控单元,用于控制信号的储存、利用 和舍弃,如图 4 所示。对于每个时刻 *t*,LSTM 有输 人门 I_i 、遗忘门 F_i 和输出门 O_i 共3个门控单元^[28]。



遗忘门基于当前时间输入、前一时刻的输出和 遗忘门的偏置项共同决定需要遗忘那一部分记忆。 当前时刻第 *i* 个 LSTM 单元的遗忘门的值可以用式 (1)来表示。

 $F_{i}^{(t)} = \sigma(b_{i}^{F} + \sum_{j} U_{i,j}^{F} x_{j}^{(t)} + \sum_{j} W_{i,j}^{F} h_{j}^{(t-1)})$ (1) 其中, $b^{F} \setminus U^{F}$ 和 W^{F} 分别是 LSTM 单元遗忘门的输 入偏差、输入权重和循环权重, $x^{(t)}$ 表示当前时刻的 输入向量, j 是输入向量的数量, $h^{(t-1)}$ 是前一时刻的 输出, σ 是 Sigmoid 函数, 它将遗忘门设置为 0 到 1 之间的值。循环神经网络不仅需要遗忘门忘记部分 之前的记忆, 它还需要补充最新的记忆, 可以通过输 入门来控制。当前时刻第 i 个 LSTM 单元的输入门 的值可以用式(2)来表示。

$$\boldsymbol{I}_{i}^{(t)} = \boldsymbol{\sigma}(\boldsymbol{b}_{i}^{T} + \sum_{j} \boldsymbol{U}_{i,j}^{T} \boldsymbol{x}_{j}^{(t)} + \sum_{j} \boldsymbol{W}_{i,j}^{T} \boldsymbol{h}_{j}^{(t-1)})$$

$$(2)$$

其中,**b**[']、**U**[']和 **W**[']分别是 LSTM 单元遗忘门的输入 偏差、输入权重和循环权重。

LSTM 单元的内部状态将会以式(3)更新。

$$S_{i}^{(t)} = F_{i}^{(t)}S_{i}^{(t-1)} + I_{i}^{(t)} \tanh(b_{i}^{t} + \sum_{j} U_{i,j}x_{j}^{(t)} + \sum_{j} W_{i,j}h_{j}^{(t-1)})$$
(3)

其中,**b**、**U**和**W**分别是 LSTM 单元中的输入单元的 输入偏差、输入权重和循环权重。当前时刻第*i*个 LSTM 单元的输出门的值可以用式(4)来表示。

 $O_{i}^{(\iota)} = \sigma(b_{i}^{0} + \sum_{j} U_{i,j}^{0} \mathbf{x}_{j}^{(\iota)} + \sum_{j} W_{i,j}^{0} h_{j}^{(\iota-1)})$ (4) 其中, b^{0} 、 U^{0} 和 W^{0} 分别是 LSTM 单元输出门的输入 - 1092 — 偏差、输入权重和循环权重。LSTM 在得到当前时刻 状态值 $S_{(i)i}$ 后会进一步产生当前时刻的输出 $h_i^{(i)}$ 。

$$\boldsymbol{h}_{i}^{(t)} = \tanh(\boldsymbol{S}_{i}^{(t)}) \cdot \boldsymbol{O}_{i}^{(t)}$$
(5)

基于上述门控单元,LSTM 能够学习到长期依赖 性。利用 LSTM 监测刀具状态,现有研究只考虑了正 向依赖,而对倒序数据的反向依赖关注较少。对于预 测问题中的依赖关系,序列数据进入 LSTM 网络是按 时间顺序排列的,结果 LSTM 网络中的信息沿着链结 构从时间步长 *t*-1 传递到时间步长 *t*,但有价值的信息 可能被过滤掉。因此,获得反向依赖关系很有必要。 双向 LSTM 网络,既可以捕获正向依赖关系,也可以 捕获反向依赖关系。双向 LSTM 能够使用 2 个单独 的隐藏层在 2 个方向处理序列数据,然后反馈到同一 输出层。以下定义了相应的隐藏层函数,其中→和← 分别表示前进和后退的过程,如式(6)所示。

$$\boldsymbol{h}_{j}^{(t)} = \vec{\boldsymbol{h}}_{j}^{(t)} \oplus \vec{\boldsymbol{h}}_{j}^{(t)}$$
(6)

堆叠多个 LSTM 层形成一个深度 LSTM 神经网 络——堆叠的双向 LSTM 网络(stacked Bi-directional long short term memory, SBiLSTM),可以让模型从 输入数据中获取更深层次的特征。当涉及到深度 LSTM 时,模型的输入可以通过多个 LSTM 层传递, 图 5 所示为 SBiLSTM 网络结构。一个 LSTM 层传递, 图 5 所示为 SBiLSTM 网络结构。一个 LSTM 层的隐 藏输出不仅随时间传播,而且还用于下一 LSTM 层 的输入数据。在双向 LSTM 框架中,每个隐藏层都 接收一个输入序列,该输入序列由下面一级的前向 和后向层的输出序列组成。在终端时间步长的最后 LSTM 层的输出被用作 SBiLSTM 网络的输出。 LSTM 层堆叠的优点是双重的,一个是堆叠层使模 型能够学习不同时间尺度上原始信号的特性。另一



个是参数可以分布在空间层上,而不是增加存储器 大小,有助于输入原始信号实现更有效的非线性操 作。

1.3 分类机制

卷积神经网络和循环神经网络可以视作特征提 取器,在对刀具磨损状态识别时需要加入分类结构 的全连接神经网络。对于多分类问题,输入、输出满 足一个映射关系,信息从输入 *x* 的经过定义的映射 关系,最终达到输出 *y*。映射关系中间计算过程通 过完全连接的密集层堆叠在一起,其中一层的输出 用作下一层的输入。每层的计算公式为

 $y_i = f(w_i^T x_i + b_i)$ (7) 其中, $y_i 和 x_i$ 分别表示第 i 个全连接层的输出和输 入。 $w_i^T 和 b_i$ 分别表示第 i 个完全连接层中的权重矩 阵和偏置项。函数 f()设置为 ReLu 函数。最终,全 连接神经网络最后一层将输出分类状态。

2 拉削刀具磨损状态识别模型

2.1 模型构建

DSBiLSTM 模型分为特征提取网络和分类网络。从特征提取的角度分析,一般认为 DenseNet(3-2)网络适用于不同方向的振动信号空间特征提取, 而 SBiLSTM 网络更适用于同方向振动信号时间序

列特征提取。DenseNet(3-2)网络强调几组独立数 据在组别之间的上下层联系,数据处理在序列上相 互独立。SBiLSTM 网络对于数据处理则侧重于序列 顺序关系上,考虑当前数据对未来数据发展产生的影 响和对前期数据的关联关系。DenseNet(3-2)网络和 SBiLSTM 网络可以视作特征提取器,在对拉刀磨损 状态识别时需要加入分类结构的全连接神经网络。

本文基于 DenseNet(3-2)、SBiLSTM 和 FNN 网 络建立了端到端的刀具磨损状态识别模型,如图 6 所示,使得一个特征的输出在空间上和前后序列上 都有联系,同时考虑到数据空间和时间关联关系,实 现自适应特征提取,并对拉刀磨损状态精准识别。

表1给出了所建模型的每层详细信息。首先对 输入序列数据进行预处理归一化操作,接着通过 DenseNet(3-2)网络提取空间特征,空间特征提取之 后保持较小的大小输出。然后,SBiLSTM 网络提取 时间序列信息,以弥补 DenseNet(3-2)网络在时序上 特征提取的不足。此外,每个 BILSTM 网络的前向 层和后向层共享相同数量的隐藏单元。在 SBiLSTM 输出之后,将最后一个 SBiLSTM 隐藏单元使用 FNN 网络添加非线性输出。最后,将学习到的高级表示 形式接入线性分类层(softmax)以估计刀具磨损状 态。此外,模型中使用 Dropout 层以避免过度拟合 和引入 ReLU 函数标准化输出。

DSBiLSTM 模型	网络层	输出层	细节
Input		$1 \times 14 \times 30\ 000$	
	Conv1	$1\times32\times15\;000$	Conv1D, 32, kernelsize = 3, stride = 2
	Pool1	$1 \times 32 \times 7500$	MaxPooling1D, stride = 2
	Conv2	$1 \times 64 \times 3750$	Conv1D, 64, kernelsize = 3, stride = 2
	Pool2	$1 \times 64 \times 1873$	MaxPooling1D, stride = 2
DenseNet (3-2)	DenseBlock1	1 × 192 × 1873	$ \begin{pmatrix} 1 \times 1 \text{ Conv1D} \\ 1 \times 3 \text{ Conv1D} \end{pmatrix} \times 4 $
	Conv3	$1 \times 96 \times 1873$	Conv1D,96, kernelsize = 3; stride = 1
	Pool3	$1 \times 96 \times 935$	MaxPooling1D, stride = 2
	DenseBlock2	1 × 224 × 935	$\binom{1 \times 1 \text{ Conv1D}}{1 \times 3 \text{ Conv1D}} \times 4$
	ReLU	$1 \times 224 \times 935$	ReLU, inplace = true
	$\operatorname{Avg}_\operatorname{pool}$	1 × 224 × 935	Avg_pool1d , kernelsize = 3, stride = 1

表1 刀具状态识别模型详细信息

SBiLSTM	BiLSTM	$1 \times 935 \times 64$	LSTM, hidden _ size = 32, num _ layers = 2, bidirectional = True, dropout = 0.5
	BiLSTM	$1 \times 935 \times 128$	LSTM, hidden $_$ size = 64, num $_$ layers = 2, bidirectional = True, dropout = 0.5
	F 1	129	Fully connect, bias = True, ReLU, inplace = true, BatchNorm, Momentum = 0.1 ,
	rci	128	eps = 1e-05, Dropout, P = 0.5, inplace = False
	E O	64	Fully connect, bias = True, ReLU, inplace = true, BatchNorm, Momentum = 0.1 ,
	Fc2	64	eps = 1 e-05, Dropout, P = 0.5, inplace = False
	Classification	3	Softmax, dim = 1

(续表1)

注: Conv 表示 BatchNorm-ReLU-Conv, Fc 表示 Fc-BatchNorm-ReLU-Dropout



图 6 刀具状态识别模型

2.2 模型训练

振动信号从加工现场收集,考虑到不同的加速 度传感器数值上的差异,对数据进行归一化处理,以 降低计算复杂度。原始数据根据式(8)进行归一 化。



其中,x_{max}表示信号中的最大值。

对于做统一变换的监测信号,不会影响信号对 刀具磨损状态的表征能力,归一化前后的信号如 图7所示。



图 7 数据预处理

在模型训练过程中,交叉熵损失误差被用作模型的损失函数,定义为

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij} \log(p_{ij})$$
(9)

其中,N表示样本的数量,M表示类别的数量,y_{ij}指 示变量(如果该类别和样本 *i* 的类别相同就是1,否 则就是0),*p_{ii}为观测样本 i* 于类别*j*的预测概率。

将数据经归一化后,随机选取带磨损状态标注 的振动信号样本的 80% 为训练集训练模型,剩余 20%用作测试集评估训练优化后的模型,模型的训 练和测试过程如图 8 所示。利用训练集对刀具磨损 状态模型进行训练,将训练集样本预处理和归一化 后送入 DSBiLSTM 用于刀具磨损状态分类,计算实 际类别与预测类的交叉熵误差。模型采用 Adam (adaptive moment estimation)优化算法更新网络的 超参数,使得交叉熵损失函数的值不断下降,让模型 预测值更趋于真实值。若损失函数未呈现逐渐减小 趋势,则模型存在过拟合,调整模型结构进行训练。 反之,模型收敛,调整模型参数直到具有较高的精 度,保存模型结构和参数用于实际加工过程。将测 试集作为训练有素的 DSBiLSTM 模型的输入,通过 正向传播得到刀具磨损状态,评估 DSBiLSTM 模型 的训练效果。



3 实验验证及结果分析

3.1 实验装置和流程

实验研究在高速卧式榫槽拉床(型号: LG6516zx-2800)上进行,拉削刀具采用涂层硬质合 金,拉削工件为某型号飞机涡轮盘,材料为高温镍基 粉末合金FH97。拉削过程中采用14个加速度传感 器(型号:PCB333B30,量程:±50g,灵敏度:((±10%) 100 mV/g),按照 x, y, z 3 个方向粘附在拉床刀盒上 (见图 9),用于实时采集拉刀在加工过程中的振动 信号。采用 SCADAS III-305 采集卡处理实时信号 并传送到 Dell/M90 计算机上,信号的采样频率为 2.56 kHz。实验训练深度学习的硬件平台采用高性 能服务器,InterCore i7-9700 CPU,主频 3.00 GHz, 16 GB内存,GPU 选择 NVIDIAGeForce RTX2070 SU-PER 图像处理器。软件平台使用 Ubuntu 16.04 操 作系统,深度学习框架选用 pytorch。



图 9 刀具磨损实验装置图

实验采用8把拉削刀具完成拉削操作,总共得 到790个原始样本信号。拉削刀具加工工件一次记 为一个拉削行程,每个拉削行程结束后,将拉削刀具 取下放到显微镜下,测量并记录最易磨损的刀齿后 刀面的1/2处得刀具平均磨损量。

根据刀具磨损过程并结合实验情况,在后刀面磨 损量为0~0.05 mm 时,拉削刀具在较短的时间内迅 速磨损,将该阶段划分为初期磨损阶段;在后刀面磨 损量为0.05~0.2 mm 时,拉削刀具切削过程平稳,工 件表面质量较好,将该阶段划分为中期磨损阶段;在 后刀面磨损量大于0.2 mm 时,刀具磨损量迅速增加, 工件加工表面粗糙,刀具在较短的时间内迅速达到失 效状态,需要及时换刀或刃磨,将该阶段划分为磨钝 阶段。将刀具磨损状态划分为3种数据标签,并采用 one-hot 编码形式对刀具磨损状态进行编码。

3.2 实验结果与分析

实验将训练集样本送入 DenseNet(3-2) 网络和 - 1096 --

SBiLSTM 网络自适应提取空间和时间上的特征,特征提取后使用 FNN 网络用于刀具磨损状态分类,计 算实际类别与预测类的交叉熵误差,采用 Adam 优 化算法更新网络的权重参数,使得交叉熵损失函数 的值不断下降,让模型预测值更趋于真实值,并保存 训练结果。同时,为了能更好地阐述量化特征提取 方法在刀具状态识别中的优越性,将 DenseNet(3-2) 分类模型和 SBiLSTM 分类模型与 DSBiLSTM 模型进 行对比,训练过程中3个模型设置相同的参数,模型 的具体参数表如表2 所示。

表 2 模型具体参数表

参数	DSBiLSTM	DenseNet(3-2)	SBiLSM
学习率	0.001	0.001	0.001
权重衰减	0.01	0.01	0.01
迭代次数	200	200	200
批处理数	16	16	16
随机失活	0.5	0.5	0.5
优化算法	Adam	Adam	Adam

经过训练后,得到不同的准确率和损失函数值, 3个模型的准确率和损失函数变化曲线如图 10 所 示。由图 10 可知,随着迭代次数(epoch)的增加,各 个模型的交叉熵损失函数总体呈现下降趋势,准确 率逐渐升高,没有出现梯度爆炸或梯度弥散现象,最 终模型都以较快速度收敛。



图 10 各模型准确率和损失函数变化曲线

测试样本总量为158个,其中初期磨损样本为 39个、中期磨损样本为100个、磨钝样本为19个, 将模型训练收敛后保存的DenseNet(3-2)分类模型、 SBiLSTM 分类模型和 DSBiLSM 模型的结构和权重 参数,放入测试集中进行测试。通过混淆矩阵展示 测试结果如图 11 所示。



由图 11(a)和图 11(b)识别结果可以得出,只 关注空间特征信息的 DenseNet(3-2)模型准确率为 95.56%;只关注时序特征信息的 SBiLSTM 模型准 确率为 84.81%,虽然 DenseNet(3-2)模型准确率比 较高,但在对刀具初期磨损状态识别中精确率和召 回率分别为 90.24%、94.87%,仍不能满足精度要 求。由于初始振动信号没有经过降维处理,直接将 标准化后的振动信号作为 SBiLSTM 模型的输入,在

完整序列中提取时序特征信息运算成本巨大,且 SBiLSTM 模型未能学习到刀具磨损状态的空间特征 映射,导致 SBiLSTM 模型识别效果较差。以上 DenseNet(3-2)模型和 SBiLSTM 模型对拉削刀具磨 损状态识别,其结果都不理想。

从图 11(c)分析可以得出,DSBiLSTM 模型的准确率为 98.73%,比 DenseNet(3-2)模型准确率提升了 2.17%;DSBiLSTM 模型中的 SBiLSTM 网络的输

入是经过 DenseNet(3-2)网络特征提取降维后的特征向量(*d* = 224),所以 DSBiLSTM 模型相对 SBiL-STM 模型提取时序特征的情况,参数计算量小,运算速度快;DSBiLSTM 模型在对磨钝状态识别中,精确率和召回率为100%、95%。

表 3 测试集的单次测试时间和准确率

模型	单次测试时间/ms	准确率/%
DSBiLSTM 模型	11	98.73
DenseNet(3-2)模型	13	95.56
SBiLSTM 模型	89	84.81

由表3可知, DenseNet(3-2)和 SBiLSTM 模型测 试集的准确率分别为95.56%和84.81%,预测精度 较低,表明单独的空间特征提取和时间特征提取可 以对刀具磨损状态进行预测,但由于受模型自适应 特征提取能力的限制,无法捕获刀具振动信号中更 深层次隐藏的特征。DSBiLSTM 模型相较于 DenseNet(3-2)和SBiLSTM 模型结构较深的情况下, 单次识别时间分别快 2 ms 和 78 ms。DSBiLSTM 模 型利用 DenseNet(3-2)模型特征重用优势有效地提 取振动信号中的空间特征信息,并对振动信号进行 降维,便于后续时序特征提取到振动信号时序特征 的依赖关系,同时利用 SBiLSTM 模型的优势,从正 向和反向捕捉振动信号过去和未来的特征信息,挖 掘出振动信号丰富的时间特征信息。最后,测试集 的准确率达到 98.73%。在性能方面, DSBiLSTM 模 型的单次测试时间为11 ms,满足工业生产在线拉削 刀具状态识别应用。因此,实验结果表明 DSBiL-STM 模型具有较高的识别精度,更适合在线拉削刀 具状态识别。

4 结论

本文提出了一种用于刀具磨损状态识别的改进 的深度学习模型,该模型整合了 DenseNet 网络和 LSTM 网络的优势,由 DenseNet(3-2)网络结构挖掘 出振动信号与刀具磨损状态间的空间相关性特征, 而 SBiLSTM 网络能够补充与刀具磨损状态间的时 序相关性特征。此外,所设计实验装置可以有效地 - 1098 - 采集拉削刀具在操作过程中输出的振动信号和测量 拉削刀具的磨损量,具有简单实用的特点。对比实 验研究发现,DSBiLSTM 模型兼顾 DenseNet(3-2)模 型和 SBiLSTM 模型的特性,具有较深的模型结构, 且比 DenseNet(3-2)模型和 SBiLSTM 模型有较高的 准确率和较快的运算速度,适用于涡轮盘榫槽拉削 刀具在线磨损状态识别。

深度学习是数据驱动的方法,需要大量数据进 行模型训练,且刀具状态识别模型随着拉削系统参 数的变化也需要进行动态调整。后续研究将进一步 增加传感型号的类型和数量,以获取更多的运行数 据,实现对刀具磨损状态提取更加全面的状态特征, 并动态更新模型参数,从而达到更准确的识别效果。

参考文献

- [1] SEIMANN M, PENG B X, FISCHERSWORRING-BUNK A, et al. Model-based analysis in finish broaching of inconel 718[J]. International Journal of Advanced Manufacturing Technology, 2018, 97(9-12): 3751-3760
- [2] YANG Y F, HAO B J, HAO X Q, et al. A novel tool (single-flute) condition monitoring method for end milling process based on intelligent processing of milling force data by machine learning algorithms[J]. International Journal of Precision Engineering and Manufacturing, 2020, 21(11): 2159-2171
- [3] 廖小平,黎宇嘉,陈超逸,等.基于核主成分和灰狼优化算法的刀具磨损状态识别[J].计算机集成制造系统,2020,26(11):139-147
- [4] KONG D D, CHEN Y J, LI N, et al. Tool wear monitoring based on kernel principal component analysis and v-support vector regression[J]. The International Journal of Advanced Manufacturing Technology, 2017, 89(1-4): 175-190
- [5] 王民,刘利明,宋铠钰,等. 基于主轴驱动电流杂波的 立铣刀复杂工况下磨损状态辨识[J]. 计算机集成制 造系统,2021,27(12):3429-3438
- [6] RMILI W, OUAHABI A, SERRA R, et al. An automatic system based on vibratory analysis for cutting tool wear monitoring[J]. *Measurement*, 2016, 77:117-123
- [7] SHI C M, PANOUTSOS G, LUO B, et al. Using multiple-feature-spaces-based deep learning for tool condition monitoring in ultraprec's on manufacturing [J]. IEEE

Transactions on Industrial Electronics, 2019, 66 (5): 3794-3803

- [8] 王丽华,杨家巍,张永宏,等. 基于堆叠降噪自编码的 刀具磨损状态识别[J]. 中国机械工程,2018,29(17): 2038-2045
- [9] 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-Transactions of the ASME, 2017, 139(7): 1-9
- [10] AZMI A I. Monitoring of tool wear using measured machining forces and neuro-fuzzy modelling approaches during machining of GFRP composites [J]. Advances in Engineering Software, 2015, 82:53-64
- [11] GAJATE A, HABER R, DEL TORO R, et al. Tool wear monitoring using neuro-fuzzy techniques: a comparative study in a turning process[J]. Journal of Intelligent Manufacturing, 2012, 23(3): 869-882
- [12] 谢楠,马飞,段明雷,等. 基于主成分分析与 C-支持向 量机的刀具磨损状态监测[J].同济大学学报(自然科 学版),2016,44(3):434-439
- [13] GAO C, XUE W, REN Y, et al. Numerical control machine tool fault diagnosis using hybrid stationary subspace analysis and least squares support vector machine with a single sensor [J]. Applied Sciences-Basel, 2017, 7(4): 346-357
- [14] GANGADHAR N, KUMAR H, NARENDRANATH S, et al. Condition monitoring of single point cutting tools based on machine learning approach [J]. International Journal of Acoustics and Vibration, 2018, 23(2): 131-137
- [15] 孙志娟. 基于扩展隐马尔可夫模型的刀具磨损识别与 寿命预测研究[D]. 武汉:华中科技大学机械科学与工 程学院,2019:25-50
- [16] YU J S, LIANG S, TANG D Y, et al. A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction [J]. International Journal of Advanced Manufacturing Technology, 2017, 91(1-4): 201-211
- [17] KONG D D, CHEN Y J, LI N. Hidden semi-Markov model-based method for tool wear estimation in milling process[J]. International Journal of Advanced Manufacturing Technology, 2017, 92(9-12): 3647-3657
- [18] KONG D D, CHEN Y J, LI N. Force-based tool wear es-

timation for milling process using Gaussian mixture hidden Markov models [J]. International Journal of Advanced Manufacturing Technology, 2017, 92(5-8): 2853-2865

- [19] SCHWENZER M, MIURA K, BERGS T. Machine learning for tool wear classification in milling based on force and current sensors [C] // The 4th International Conference on Mechanical, Manufacturing, Modeling and Mechatronics (IC4M), Nice, France, 2019: 1-10
- [20] NGUYEN V, NGUYEN V, PHAM V. Deep stacked auto-encoder network based tool wear monitoring in the face milling process[J]. Strojniski Vestnik-Journal of Mechanical Engineering, 2020, 66(4): 227-234
- [21] 卢志远,马鹏飞,肖江林,等. 基于机床信息的加工过 程刀具磨损状态在线监测[J]. 中国机械工程,2019, 30(2):220-225
- [22] CAO X C, CHEN B Q, YAO B, et al. An intelligent milling tool wear monitoring methodology based on convolutional neural network with derived wavelet frames coefficient[J]. Applied Sciences-Basel, 2019, 9(18): 1-26
- [23] CAI W L, ZHANG W J, HU X F, et al. A hybrid information model based on long short-term memory network for tool condition monitoring [J]. Journal of Intelligent Manufacturing, 2020, 31(6): 1497-1510
- [24] ZHAO R, YAN R Q, WANG J J, et al. Learning to monitor machine health with convolutional bi-directional LSTM networks[J]. Sensors, 2017, 17(2): 273-290
- [25] SUN H B, ZHANG J D, MO R, et al. In-process tool condition forecasting based on a deep learning method [J]. Robot and Computer-Integrated Manufacturing, 2020, 64: 1-8
- [26] DAI L L, LIU T Y, LIU Z Y, et al. An improved deep learning model for online tool condition monitoring using output power signals[J]. Shock and Vibration, 2020, 8: 211490-211505
- [27] 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
- [28] TAO Z, AN Q, LIU G, et al. A novel method for tool condition monitoring based on long short-term memory and hidden Markov model hybrid framework in high-speed milling Ti-6Al-4V [J]. The International Journal of Advanced Manufacturing Technology, 2019, 105 (7-8): 3165-3182

Recognition model of broaching tool wear state based on deep learning

YING Shenshun*, FU Chentai*, LIN Lvsheng**, LV Xiaomin**, ZHANG Shunqi***, YI Kai*

(*School of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310032)

(** Zhejiang CHR Intelligent Equipment Co., Ltd, Jinyun 321404)

(*** School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444)

Abstract

Broaching is a key processing technology for the tongue and groove of aero-engine turbine disks. If the abnormal state of the broach during the working process cannot be detected and intervened in time, it will cause serious damage. This paper proposes a model (DSBiLSTM) for identifying the wear status of broaching tools. In the proposed model, densely connected network (DenseNet)(3-2) is designed based on the DenseNet network to achieve spatial feature extraction and dimensionality reduction with the advantage of feature reuse, and a stacked Bi-directional long short-term memory (LSTM) network (SBiLSTM) is realized for time feature extraction. Features are extracted by multi-scale fusion and incorporated into multiple fully connected layers (FNN) and Softmax layers to realize tool wear status recognition. Based on the innovative recognition model, a broaching experiment of the tongue and groove of the turbine disk is designed, and the performance index system based on the confusion matrix is used to compare the proposed model with the single feature extraction model. The experimental results show that DSBiL-STM has outstanding performance in the prediction of tool wear status, and the recognition accuracy rate reaches 98.73%, and the recognition speed of a single sample is increased to 11 ms.

Key words: deep learning, broaching process, tool wear, state recognition, densely connected network (DenseNet), long short-term memory (LSTM)