

融合 SKNet 与堆叠 LSTM 的 MobileNetV3 齿轮箱故障识别方法^①



杨辰峰¹, 杨喜旺², 黄晋英³, 范振芳³, 刘晶晶¹

¹(中北大学 计算机科学与技术学院, 太原 030051)

²(山西电子科技学院 信息与通信工程学院, 临汾 041000)

³(中北大学 机械工程学院, 太原 030051)

通信作者: 杨喜旺, E-mail: yangxw@nuc.edu.cn

摘要: 当前基于深度学习的故障识别方法普遍面临高数据依赖性、高昂计算成本与时间开销, 以及模型泛化能力受限等挑战. 为此, 本研究提出一种融合 MobileNetV3、选择性核网络 (selective kernel network, SKNet) 及堆叠长短期记忆网络 (stacked long short-term memory network, Stacked LSTM) 的轻量化高精度故障识别模型. 首先进行输入数据预处理, 将处理后的数据转换成适应卷积层的输入格式. 在特征提取阶段, 利用改进的 MobileNetV3 骨干网络进行深度特征挖掘, 其倒置残差模块在保留深度可分离卷积高效性的基础上, 策略性地嵌入 SE (squeeze-and-excitation) 与 SK (selective kernel) 双重注意力机制, 有效兼顾通道信息交互与多尺度特征自适应选择, 显著提升了特征表征能力并降低了计算复杂度. 随后, 堆叠 LSTM 捕获振动信号中的长距离时序依赖关系. 最终通过全连接层实现特征压缩与分类决策, 构建端到端识别系统. 实验结果显示, 本文模型识别准确率达到 99.47%, 与传统的齿轮箱故障识别技术相比, 该方法在识别精准度和模型泛化能力方面均呈现出显著优势.

关键词: 故障识别; 深度学习; MobileNetV3; 选择性核网络; 堆叠长短期记忆网络

引用格式: 杨辰峰, 杨喜旺, 黄晋英, 范振芳, 刘晶晶. 融合 SKNet 与堆叠 LSTM 的 MobileNetV3 齿轮箱故障识别方法. 计算机系统应用, 2026, 35(1): 237-245. <http://www.c-s-a.org.cn/1003-3254/10068.html>

MobileNetV3 Gearbox Fault Identification Method Combining SKNet and Stacked LSTM

YANG Chen-Feng¹, YANG Xi-Wang², HUANG Jin-Ying³, FAN Zhen-Fang³, LIU Jing-Jing¹

¹(School of Computer Science and Technology, North University of China, Taiyuan 030051, China)

²(School of Information and Communication Engineering, Shanxi University of Electronic Science and Technology, Linfen 041000, China)

³(School of Mechanical Engineering, North University of China, Taiyuan 030051, China)

Abstract: Current fault identification methods based on deep learning generally face challenges such as high data dependency, high computational cost and time consumption, and limited model generalization ability. To address these issues, this study proposes a lightweight and high-precision fault identification model that integrates the MobileNetV3, selective kernel network (SKNet), and stacked long short-term memory network (Stacked LSTM). First, the input data is preprocessed, and the processed data is converted into a format suitable for convolutional layer input. In the feature extraction stage, an improved MobileNetV3 backbone network is employed for deep feature mining. On the basis of retaining the efficiency of depthwise separable convolution, an inverted residual module strategically embeds a dual attention mechanism combining squeeze-and-excitation (SE) and selective kernel (SK), effectively enhancing channel information interaction and multi-scale feature adaptive selection, thus significantly improving feature representation and

① 基金项目: 山西省基础研究计划 (202203021211096)

收稿时间: 2025-06-24; 修改时间: 2025-08-01, 2025-08-13; 采用时间: 2025-09-05; csa 在线出版时间: 2025-11-26

CNKI 网络首发时间: 2025-11-27

reducing computational complexity. Subsequently, the stacked LSTM captures long-term temporal dependencies in the vibration signals. Finally, feature compression and classification decision-making are achieved through the fully connected layer, thus constructing an end-to-end recognition system. Experimental results show that the recognition accuracy of the proposed model reaches 99.47%, demonstrating significant advantages in recognition accuracy and model generalization ability compared with traditional gearbox fault identification techniques.

Key words: fault identification; deep learning; MobileNetV3; selective kernel network (SKNet); stacked long short-term memory network (Stacked LSTM)

1 引言

齿轮箱作为工业设备的关键传动部件,其运行健康状况直接关系到整个系统的安全性能。若故障未及时发现,可能会导致设备停机、生产线瘫痪,甚至引发安全事故。随着工业4.0和智能制造的快速发展,设备的可靠性和稳定性愈发关键,面向边缘计算场景的轻量化、高泛化、强特征提取能力的故障识别技术作为保障设备高效运行、减少停机损失的核心手段,已成为学术界与工业界的研究热点^[1,2]。

近年来,深度学习凭借其卓越的特征提取与模式识别能力,在故障识别领域彰显出突破性的应用潜力,诸多基于深度学习的故障识别方法相继被提出。陈超等^[3]构建了以卷积神经网络(CNN)和长短期记忆网络(LSTM)为核心的模型框架,与通道注意力机制(CAM)和自注意力机制(self-attention)相结合进行行星齿轮箱的故障识别。吴启航等^[4]提出了一种面向齿轮箱故障边缘智能诊断的轻量化乘法-卷积网络(MCN)。Shi等^[5]提出了一种基于双向卷积长短期记忆网络(BiConv-LSTM)的新型深度模型,通过同步从振动信号与转速数据中自动提取的时空特征,实现了对行星齿轮箱故障类型、位置及方向的精准辨识。杨青松等^[6]提出了一种基于轻量化神经网络 ShuffleNet V2 的模型,通过引入 GELU 激活函数增强模型的非线性变换能力,并嵌入高效通道注意力(ECA)模块提升网络性能,将其应用于齿轮箱故障识别任务。Zhang等^[7]结合了优化的奇异值分解(SVD)、对称点图案(SDP)和卷积神经网络 SVD-SDP-CNN 的方法提高了故障识别的效率和准确率。虽然以上方法在故障识别领域有一定的研究,但是存在计算资源和时间成本高,泛化能力不足,特征提取能力相对有限等问题。

为了解决上述问题,本文提出了一种故障识别方法——MobileNetV3-SKNet-Stacked LSTM (MSS),以

MobileNetV3 作为轻量化骨干网络,通过深度可分离卷积和动态激活函数的设计,在保证特征提取能力的同时大幅降低模型参数规模与计算复杂度,有效解决现有方法计算资源和时间成本高的问题;融入 SKNet 模块通过自适应选择特征通道,动态聚焦故障识别的关键特征信息,增强模型对复杂工况下微弱故障特征的捕捉能力,弥补传统模型特征提取能力有限的短板;结合 Stacked LSTM 的多层时序建模能力,充分挖掘振动信号中蕴含的时序依赖关系,提升模型对变转速、负载波动等动态工况的适应能力,改善泛化性能。通过各模块协同运作,实现对齿轮箱故障的精准识别。

2 MobileNetV3-SKNet-Stacked LSTM 模型组成模块

2.1 MobileNetV3

MobileNetV3 是 Google 团队开发的基于深度可分离卷积(depthwise separable convolution, DSC)的轻量化网络架构,属于 MobileNet 系列的第3代版本。该版本有 MobileNetV3-small 和 MobileNetV3-large 两个型号,本文选用的是 MobileNetV3-large 模型^[8-11]。表1呈现了该模型的网络结构。Input 代表输入数据;Operator 指该层所采用的操作类型,其中 conv2d 表示标准卷积, bneck 代表 MobileNetV3 的倒置残差块,其后的 $n \times n$ 代表所使用的卷积核尺寸; exp_size 指每个 bneck 块中,第1层 1×1 conv2d 卷积核的数量,即利用 1×1 卷积对输入特征进行升维,升维后的维度就是 exp_size; NBN 表示该卷积层未采用批标准化(batch normalization); out 表示输出特征的通道数; SE 用于标记哪些 bneck 块使用了 SE 模块,表中“√”表示使用 SE 模块的部分; NL 指每层采用的非线性激活函数,其中 HS 是 *h-swish* 函数, RE 为 *ReLU* 激活函数; *s* 表示步长。

表1 MobileNetV3-large 模型网络结构表

Input	Operator	exp_size	out	SE	NL	s
224 ² ×3	conv2d	—	16	—	HS	2
112 ² ×16	bneck, 3×3	16	16	—	RE	1
112 ² ×16	bneck, 3×3	64	24	—	RE	2
56 ² ×24	bneck, 3×3	72	24	—	RE	1
56 ² ×24	bneck, 5×5	72	40	√	RE	2
28 ² ×40	bneck, 5×5	120	40	√	RE	1
28 ² ×40	bneck, 5×5	120	40	√	RE	1
28 ² ×40	bneck, 3×3	240	80	—	HS	2
14 ² ×80	bneck, 3×3	200	80	—	HS	1
14 ² ×80	bneck, 3×3	184	80	—	HS	1
14 ² ×80	bneck, 3×3	184	80	—	HS	1
14 ² ×80	bneck, 3×3	480	112	√	HS	1
14 ² ×112	bneck, 3×3	672	112	√	HS	1
14 ² ×112	bneck, 5×5	672	160	√	HS	2
7 ² ×160	bneck, 5×5	960	160	√	HS	1
7 ² ×160	bneck, 5×5	960	160	√	HS	1
7 ² ×160	conv2d, 1×1	—	960	—	HS	1
7 ² ×960	pool, 7×7	—	—	—	—	1
1 ² ×960	1×1 conv2d, NBN	—	1280	—	HS	1
1 ² ×1280	1×1 conv2d, NBN	—	k	—	—	1

相较于前两代, MobileNetV3 延续了 MobileNetV1 中深度可分离卷积的设计(将标准卷积拆分为深度卷积和逐点卷积)。同时在 MobileNetV2 基础上对倒置残差模块(inverted residual block)进行优化,新增了 SE 模块,并且放在 depthwise filter 之后,通过两个全连接层(FC)与激活函数实现通道信息的精准交互,强化对关键特征通道的关注,提升特征表征的针对性^[12,13]。并且引入了 *h-swish* 激活函数,其表达式为:

$$h-swish(x) = x \times \frac{ReLU6(x+3)}{6} \quad (1)$$

与 *ReLU6* 等激活函数相比, *h-swish* 激活函数在低精度计算场景中具备多方面优势,不仅能有效保留模型性能,还可通过增强非线性表达能力提升特征提取效果,同时加快模型收敛速度。此外,该函数在改善模型鲁棒性与缓解梯度消失问题上也表现突出^[14]。这些设计使 MobileNetV3 在保持轻量化特性的同时,显著提升了特征提取能力与计算效率,体现了对前两代架构的突破性优化。图1为 MobileNetV3 的倒置残差结构示意图。

2.2 SKNet 注意力机制

SKNet 作为轻量级模块,既能轻松嵌入现有成熟网络,也易与现有架构整合。在 MobileNetV3 中虽然已

经有 SE 模块,但是 SE 模块仅关注通道间信息交互,却无法处理特征尺度差异,难以区分不同大小目标特征的重要性;而 SKNet 强调多分支卷积,利用不同卷积核参数分支(如 3×3 与 5×5 卷积核分支)提取多尺度特征;通过引入残差连接确保梯度有效流经不同卷积分支,避免因梯度消失导致部分分支失效;在通道注意力基础上结合多尺度特征融合,通过全局平均池化、全连接层非线性变换生成通道权重,精准定位不同尺度特征的位置与重要性。这种设计既弥补了 SE 模块在特征尺度处理上的不足,又与 SE 形成“通道交互+多尺度适配”的双重注意力机制,显著提升对复杂工况下微弱故障特征的捕捉能力^[15,16]。

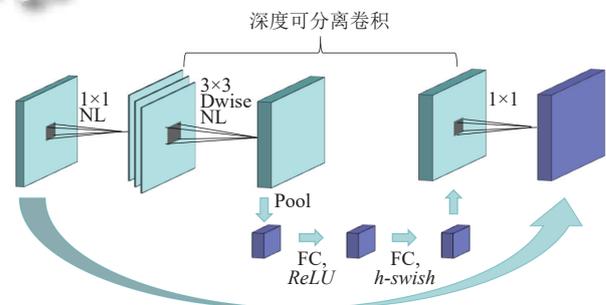


图1 MobileNetV3 的倒置残差结构示意图

SKNet 主要通过多分支卷积、特征融合以及通道注意力机制这 3 个关键步骤来提升模型性能。

首先, SKNet 构建多分支卷积结构,通常设置两个分支,分别采用不同卷积核大小的卷积操作^[17]。假设输入特征图为 X ,第 i 个分支的卷积操作可表示为:

$$Y_i = Conv_i(X) \quad (2)$$

其中, Y_i 是每个分支卷积后得到的特征图, $Conv_i$ 代表第 i 个分支的卷积运算。3×3 卷积核聚焦局部细节,5×5 卷积核捕捉更广泛上下文信息,以此让网络同时获取不同尺度特征。随后,需对多分支输出的特征图进行融合,常见做法是借助元素相加等手段将不同尺度的特征图进行组合。表示为:

$$Y_{merge} = \sum_{i=1}^M Y_i \quad (3)$$

其中, Y_{merge} 是将所有分支的特征图进行融合后的结果, M 为分支数量,综合各分支优势,形成包含更丰富信息的融合特征图。最后,引入通道注意力机制:先对融合特征图 Y_{merge} 执行全局平均池化操作,把各个通道压缩

成单个数值,以此来获取全局范围的统计信息;接着借助全连接层搭配 *ReLU* 激活函数来实现非线性变换;之后通过另一个全连接层生成通道注意力权重 *S*;最后将此权重与原始的融合特征图相乘,进而完成通道的加权处理,即:

$$Y_{final} = Y_{merge} \times S \quad (4)$$

其中, Y_{final} 是通道加权后生成最终特征图, 强化关键特征并抑制无效信息, 提升模型对复杂特征的提取能力.

2.3 堆叠 LSTM

堆叠 LSTM (Stacked LSTM) 作为深度学习领域中一类应用广泛的循环神经网络 (RNN) 架构, 在序列数据处理方面优势突出, 特别适用于时间序列分析、自然语言处理等场景. 相较于单层 LSTM, 其通过有序堆叠多个 LSTM 单元, 构建“层级信息传递”机制——前一层 LSTM 的输出作为下一层的输入, 使网络能逐层学习更高级、更抽象的序列特征. 同时, 其还对堆叠层数进行了优化: 实验验证 2-4 层为理想范围 (本文采用 2 层结构), 既能通过深层结构挖掘隐藏信息、提升训练效率与识别精度, 又可避免层数过多导致的参数爆炸、算力消耗激增及训练困难问题^[18-20]. 这种“多层次时序建模+合理层数控制”的设计, 非常适配振动信号等序列数据的长距离时序依赖关系挖掘. 图 2 为 Stacked LSTM 的基本结构图.

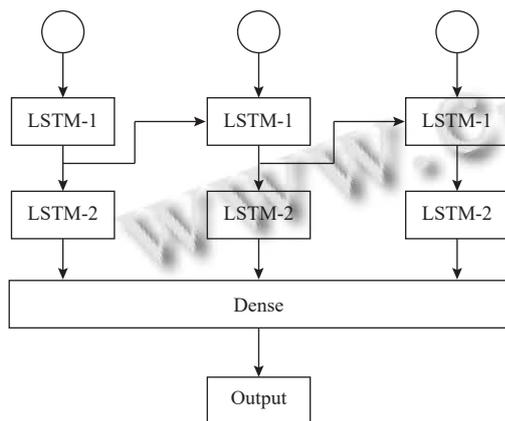


图 2 Stacked LSTM 的基本结构图

输入层接收原始的序列数据. 第 1 层 LSTM 对输入序列进行初步处理, 学习到序列的一些基础特征. 第 2 层 LSTM 以前一层输出为输入, 学习更复杂特征; 全连接层接收第 2 层 LSTM 的输出, 将时序特征转化为

分类 (如故障类型识别)、回归等任务的可解释结果. 最终输出层基于全连接层处理后的特征, 完成具体任务输出.

3 基于 MobileNetV3-SKNet-Stacked LSTM 模型的故障识别

本文通过模型融合, 构建了 MobileNetV3-SKNet-Stacked LSTM (MSS) 的故障识别模型. 该模型深度融合各模块特性优势, 实现对齿轮箱故障的高精度识别. 模型结构如图 3 所示. 该模型具有 3 大核心特点: 一是轻量化设计与效率优化并存, 依托 MobileNetV3 的深度可分离卷积降低基础计算量, 同时选择性在深层网络引入 SKNet 模块, 既避免冗余计算又增强关键特征捕捉; 二是动态特征学习能力突出, 通过 SKNet 模块的多尺度特征选择与通道注意力机制, 结合 *ReLU* 与 *h-swish* 混合激活函数的适配性设计, 提升复杂工况下特征表达能力; 三是时序建模精准, 借助 Stacked LSTM 的多层时序依赖挖掘能力, 契合振动信号的动态特性.

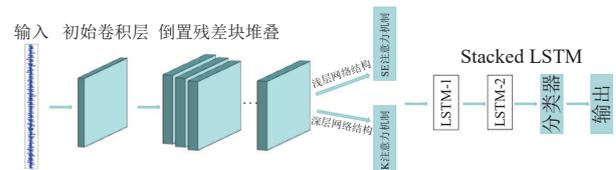


图 3 MSS 模型结构图

在神经网络中, 各层提取的特征抽象层次存在差异: 浅层网络主要捕捉边缘、纹理等低级特征, 而深层网络则聚焦于更高级的抽象特征. SKNet 模块的多尺度特征提取与通道注意力机制, 更适配于处理包含复杂语义信息的高级特征, 因此在深层网络添加该模块, 有助于更高效地捕捉和利用这些特征, 进而提升模型性能. 因此选择性地在部分层引入 SKNet 模块, 既能避免冗余计算, 又能提高模型的针对性与运行效率. 表 2 为 MobileNetV3 中倒置残差模块配置参数表.

在倒置残差模块配置表里部分模块选用 *ReLU* (RE) 激活函数, 部分采用 *h-swish* (HS) 激活函数, 这主要是基于二者特性、不同模块需求以及模型整体性能优化的考量. 在模型早期阶段, 特征图维度高、数据分布复杂, 使用 *ReLU* 可利用其计算简单和缓解梯度消失的特性快速收敛, 同时减少计算量. 到了模型后期, 特征图维度降低, 需学习更复杂抽象的特征, *h-swish* 能凭借

其非线性特性增强模型表达能力, 还可优化梯度传播, 使训练更稳定.

上述设计有效破解了传统深度模型高昂成本和时空开销的痛点, MobileNetV3 的轻量化架构缩减参数规模与计算复杂度, 选择性 SKNet 模块部署减少无效计算, 混合激活函数的针对性应用平衡学习效率与表达能力, 显著降低模型训练与部署的资源消耗.

表2 MobileNetV3 倒置残差模块配置参数表

输入通道数	卷积核大小	扩展通道数	输出通道数	是否使用 SKNet	激活函数类型	步长
16	3	16	0	—	RE	1
16	3	64	0	—	RE	2
24	3	72	0	—	RE	1
24	5	96	1	√	RE	2
40	5	120	1	√	RE	1
40	5	120	1	√	RE	1
40	3	240	1	√	HS	2
80	3	200	1	√	HS	1
80	3	184	1	√	HS	1
80	3	184	1	√	HS	1
80	3	480	1	√	HS	1
112	3	672	1	√	HS	1
112	5	672	1	√	HS	2
160	5	960	1	√	HS	1
160	5	960	1	√	HS	1

表3 为 MSS 模型结构的关键参数, 系统梳理了 MSS 整体模型的关键参数, 涵盖网络结构、激活函数等维度. 在网络结构方面, MobileNetV3 通过 cfg 列表定义倒置残差块的详细配置信息, 决定网络对图像特征提取的方式与深度.

表3 MSS 模型结构的关键参数表

模型名称	参数名称	参数作用
MobileNetV3	input_channels	输入数据通道数
	cfg	配置倒置残差块的具体参数
SELayer	channel	输入特征图的通道数
	reduction	通道压缩比, 用于调控全连接层的维度
	features	输入特征图的通道数
SKLayer	M	分支数量
	G	分组卷积的组数
	r	压缩比率
Stacked LSTM	input_size	输入特征的维度
	hidden_size	隐藏层的维度
	num_layers	LSTM 层的堆叠数量

图4 为完整的齿轮箱识别流程图. 首先对原始振动数据进行系统性预处理. 通过数据清洗剔除异常值、重复数据等干扰信息, 同时采用标准化技术对数据归一化, 消除特征尺度差异, 提升数据质量与可用性.

通过系统性预处理优化数据分布、增强数据有效性, 结合模型对关键特征的高效捕捉能力, 有效缓解了传统深度学习模型的高数据依赖性问题, 在有限数据场景下仍能保持稳定性能. 完成预处理后, 将其划分为训练集和测试集. 其中, 训练集用于对 MSS 模型进行参数优化与模型训练, 使模型充分学习故障数据特征; 测试集则用于对训练后的模型进行严格验证与全面评估, 客观衡量模型的准确性、泛化能力等性能指标, 确保模型在实际应用中的可靠性和有效性.

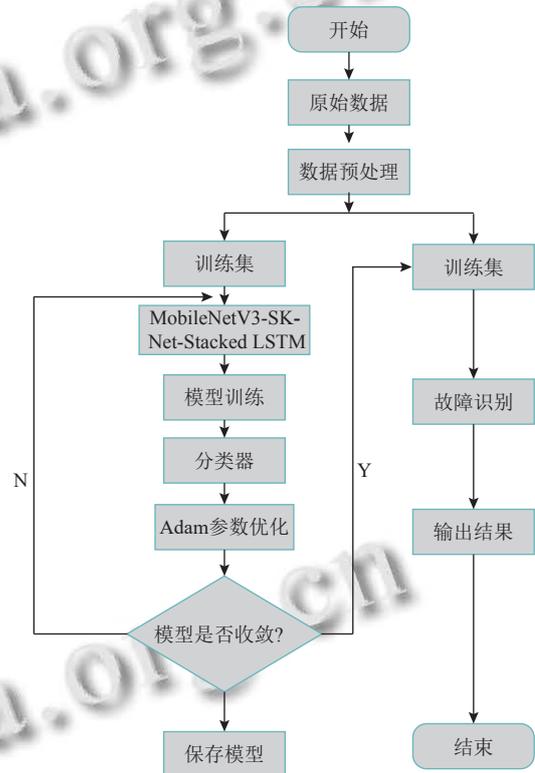


图4 故障识别流程图

4 实验分析

4.1 实验验证

本文采用实验室齿轮箱采集的数据集进行实验, 采样频率为 10.24 kHz, 将采集到的连续振动数据切成长度均为 2 048 个采样点的独立样本. 如表4 所示, 共有 6 种工况类型, 其中包含正常工况, 用 0-5 分别对 6 类工况进行标记, 每种工况选取 500 个样本, 总共收集 3 000 个样本. 将采样数据按照 7:3 分为训练集和测试集.

表5 为模型训练时设置的一系列关键参数, 学习率 0.000 1, 采用 Adam 优化器和 Cross entropy loss 损

失函数, 训练轮次为 30, Dropout 值为 0.2 以防过拟合, Early stopping 值为 10 (连续 10 轮测试集性能无提升则终止), 训练/测试集数据加载器批大小为 32. 这些参数共同作用, 使模型在训练过程中平衡训练速度、优化性能并避免过拟合, 以实现较好的故障识别效果.

表 4 数据集故障类型与划分

故障类型	训练样本长度	测试样本长度	标签
正常	350	150	0
行星轮点蚀	350	150	1
行星轮单齿齿根裂纹	350	150	2
太阳轮外圈裂纹	350	150	3
太阳轮全齿断裂	350	150	4
太阳轮内圈裂纹	350	150	5

表 5 模型关键参数表

参数	参数值
损失函数	Cross entropy loss
优化器	Adam
学习率	0.0001
训练轮次	30
Dropout	0.2
Early stopping	10
batch_size	32

图 5 展示了准确率随训练轮次的变化曲线. 从图中可见, 随着训练轮次增加, 准确率呈上升趋势, 前期上升明显, 24 轮次后趋于稳定. 这表明模型随着训练的持续, 逐步趋近于收敛状态, 充分说明模型具备良好的学习和拟合能力.

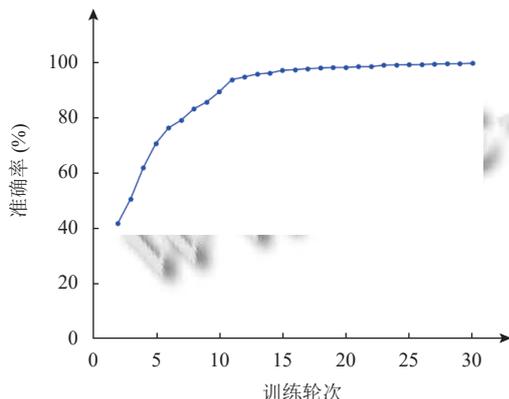


图 5 准确率变化曲线

图 6 呈现的是训练损失值随训练轮次的变化曲线. 在训练的起始阶段, 损失值处于高位. 随着训练轮次的增加, 损失值迅速下降, 这意味着模型在自适应调整参数, 缩小识别值与真实值之间的差距. 在第 24 轮次后最终稳定在一个较低的值附近, 从而达到了一个较为

稳定的拟合状态, 清晰地反映出模型训练逐步收敛的过程.

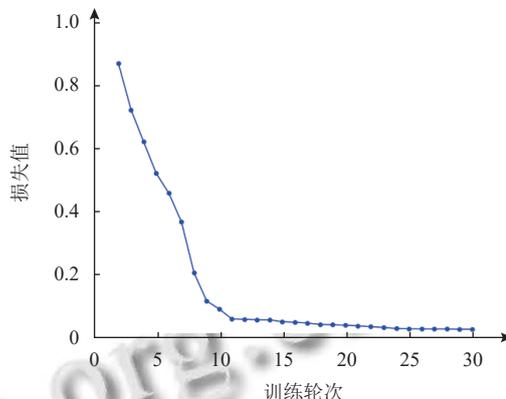


图 6 损失值变化曲线

将在训练集中训练好的 MSS 模型应用到测试集中进行分类, 将分类结果可视化得到的混淆矩阵如图 7 所示, 结果显示, 6 种工况基本全部正确分类, 仅存在个别误分类情况, 分类准确率达 99.47%.

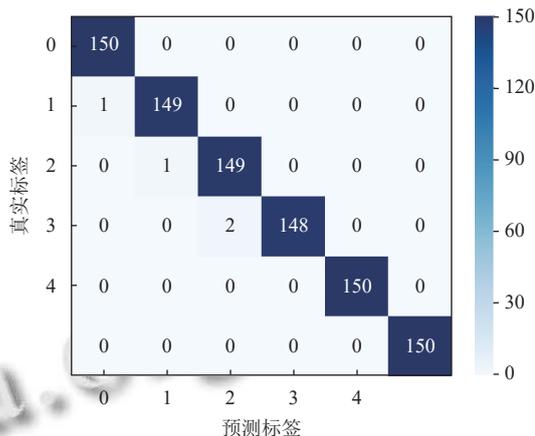


图 7 分类结果混淆矩阵

为直观展现出 MSS 模型的性能, 采用 t-分布随机邻域嵌入 (t-SNE) 技术对模型训练前后的故障特征进行可视化分析. 如图 8 所示, 在训练前的 t-SNE 散点图中, 不同故障类型的原始数据点高度聚合、界限模糊. 而训练后的 t-SNE 结果呈现显著差异, 同一故障类型的样本点形成清晰的聚类中心, 这一变化直观验证了 MSS 模型能够有效挖掘故障特征的本质差异, 通过对数据的深度处理, 实现齿轮箱故障的精准特征提取与高效分类.

4.2 对比实验

在故障识别研究中, CBAM-CNN^[21]是结合卷积神经网络 (CNN) 与通道注意力机制 (CBAM) 的经典模

型,其核心优势在于通过注意力机制增强关键特征的捕捉能力,将其作为对比模型,可验证 SKNet 在多尺度特征自适应选择上相较于 CBAM 的优势,同时凸显 Stacked LSTM 引入的时序依赖挖掘能力对模型性能的增益;深度可分离卷积 (DSC)^[22]是 MobileNet 系列的核心基础组件,其通过将标准卷积拆分为深度卷积与逐点卷积实现轻量化计算,选取 DSC 作为对比模型可直观体现 MSS 模型在基础轻量化特征提取架构上,融合 SKNet 与 Stacked LSTM 后的性能提升。

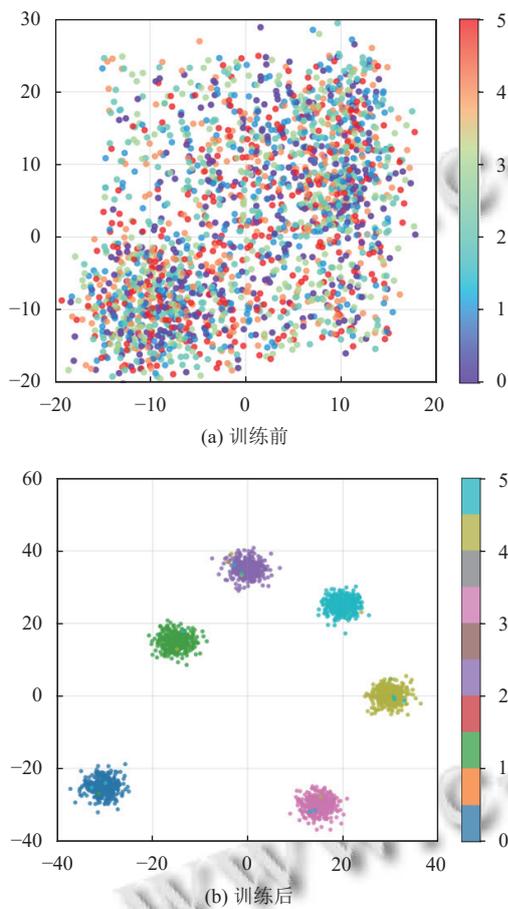


图8 t-SNE 可视化分析

为科学评估本文所提模型的性能,本文用精确率 (Precision)、召回率 (Recall)、F1 分数 (F1-Score) 和准确率 (Accuracy) 作为综合评判的标准,开展系统性对比实验.不同模型的对比实验结果如表 6 所示。

表6 不同模型对比实验结果

模型	Precision	Recall	F1-Score	Accuracy
DSC	0.9523	0.9534	0.9551	0.9545
CBAM-CNN	0.9745	0.9755	0.9756	0.9776
MSS	0.9964	0.9975	0.9940	0.9947

实验结果显示, MSS 模型的精确率、召回率、F1 分数及准确率均显著优于 CBAM-CNN 和 DSC. 这一数据表明,本文提出的模型在故障特征提取能力上更具优势,且在综合性能表现上更为优异,能够更精准地实现故障检测与分类任务. 具体而言,相较于 DSC, MSS 性能的提升说明融合 SKNet 与 Stacked LSTM 可有效弥补单纯深度可分离卷积在特征表征与动态时序建模上的不足;而相较于 CBAM-CNN, MSS 的优势则体现了 SKNet 在注意力机制上的优越性,以及 Stacked LSTM 对振动信号时序特性的适配性,充分证明了多模块融合的协同效应。

图 9 更直观地展示各模型准确率和损失值的变化曲线,本文提出的 MSS 模型在训练过程中的准确率与损失值优化效率明显优于 CBAM-CNN 和 DSC。

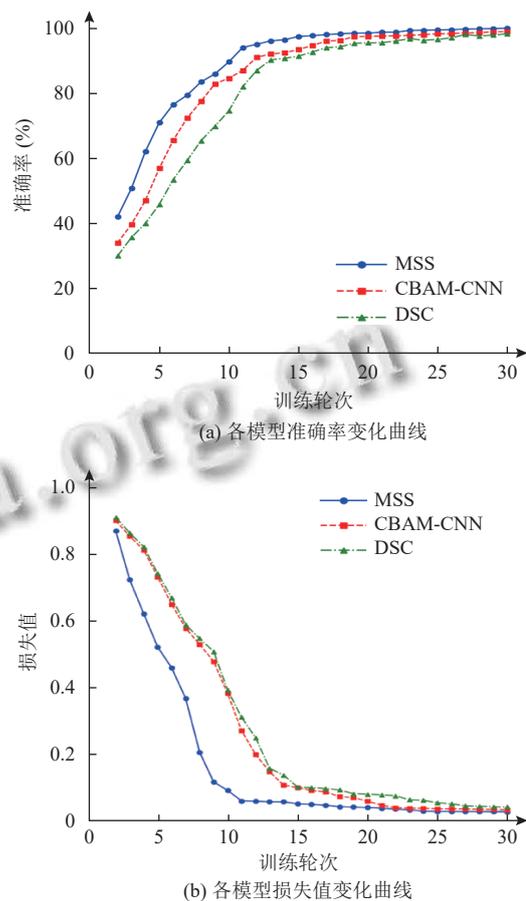


图9 各模型准确率和损失值的变化曲线

4.3 消融实验

为探究模型各组成部分对整体性能的贡献,设计消融实验验证其有效性。

- (1) 模型 1: 仅使用 MobileNetV3, 侧重轻量级特征提取.
- (2) 模型 2: 在 MobileNetV3 模型的基础上嵌入 SK 注意力机制, 提升多尺度特征融合能力.
- (3) 模型 3: 在 MobileNetV3 模型的基础上接入 Stacked LSTM, 强化序列特征建模能力.
- (4) 模型 4: 在 MobileNetV3 的基础上结合 SKNet 注意力机制和 Stacked LSTM 的完整模型.

由图 10 可以清晰地看出, MSS 的分类效果最好, 通过 SKNet 增强空间特征的多尺度表达, 有效解决单一卷积核对复杂物体特征提取的局限性; 借助 Stacked LSTM 引入时序建模能力, 解决了处理序列数据时难以捕捉长距离依赖的问题, 大幅提升模型在复杂场景下的泛化能力. 消融实验证明 3 种的结合并非简单的叠加, 而是通过深度的技术融合与优势互补, 展示了显著的协同效应.

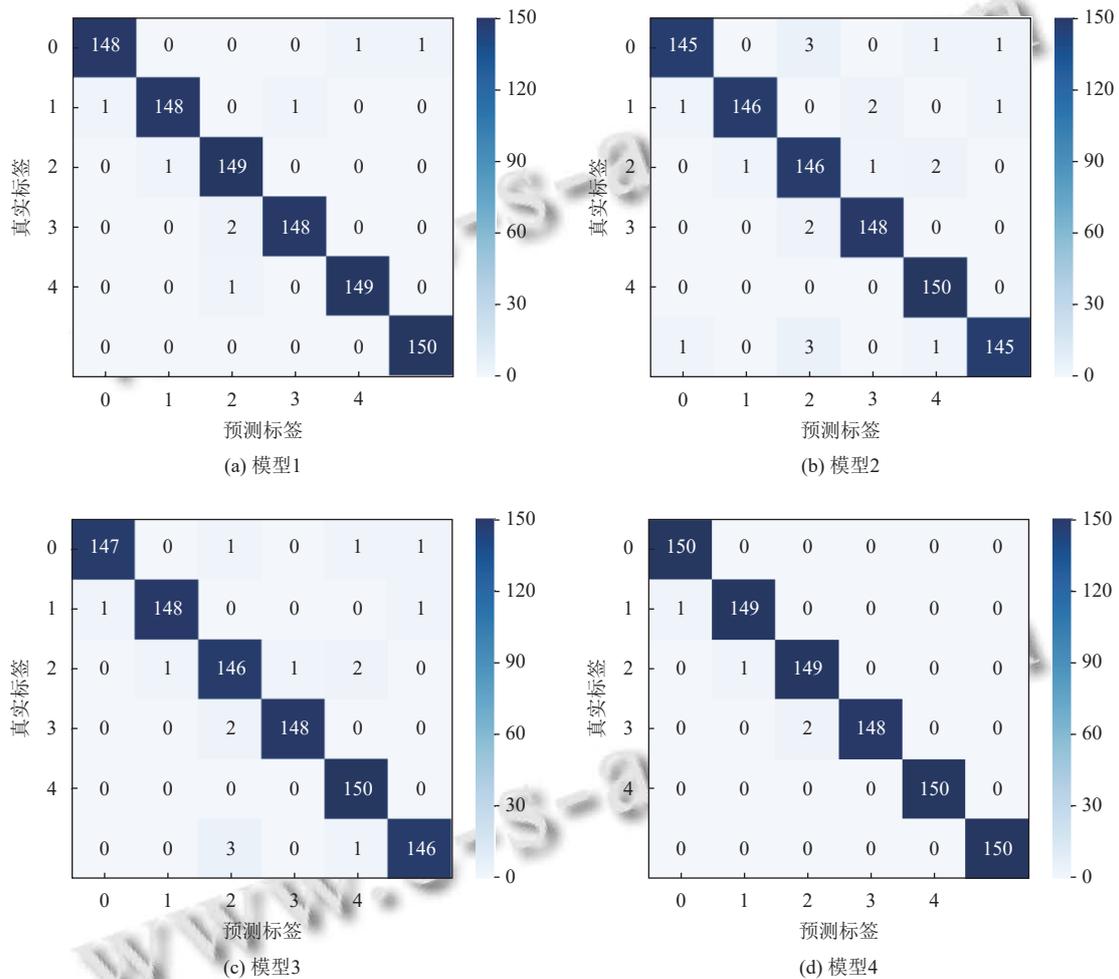


图 10 各模型混淆矩阵图

5 结论

本研究构建 MobileNetV3-SKNet-Stacked LSTM 融合模型, 在齿轮箱故障识别任务中展现出较好性能. 模型通过 MobileNetV3 轻量化骨干网络提取齿轮箱振动信号的空间特征, SKNet 的动态卷积机制增强了特征选择的灵活性, 结合 Stacked LSTM 对时序特征的深度挖掘能力, 实现了空间与时间维度的多尺度特征融合. 实验结果表明, 该模型在多工况、多类型故障数据

集上, 识别准确率较传统方法提升显著, 有效降低了漏检与误判风险, 在齿轮箱早期故障识别中具备高灵敏度与可靠性. 未来研究可进一步探索模型结构优化、引入迁移学习技术, 或结合更多传感器数据构建多模态融合模型, 以应对更复杂的工程应用需求.

参考文献

1 周舟, 陈捷, 吴明明. 基于 CWT 和优化 Swin Transformer

- 的风电齿轮箱故障诊断方法. 振动与冲击, 2024, 43(15): 200–208. [doi: [10.13465/j.cnki.jvs.2024.15.023](https://doi.org/10.13465/j.cnki.jvs.2024.15.023)]
- 2 Zhang SY, He CX, Miao ZH. Intelligent fault diagnosis system based on vibration signal edge computing. Proceedings of the 2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai). Shanghai: IEEE, 2020. 1–5.
 - 3 陈超, 许琦, 宋正华, 等. 结合深度学习和迁移学习的行星齿轮箱故障诊断算法研究. 机床与液压, 2025, 53(10): 40–49.
 - 4 吴启航, 丁晓喜, 何清波, 等. 齿轮箱故障边缘智能诊断方法及应用研究. 仪器仪表学报, 2024, 45(1): 70–80.
 - 5 Shi JC, Peng DK, Peng ZX, *et al.* Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks. Mechanical Systems and Signal Processing, 2022, 162: 107996. [doi: [10.1016/j.ymsp.2021.107996](https://doi.org/10.1016/j.ymsp.2021.107996)]
 - 6 杨青松, 郝如江, 范亚飞, 等. 一种改进轻量化神经网络的齿轮箱故障诊断方法. 科学技术与工程, 2024, 24(7): 2699–2705. [doi: [10.12404/j.issn.1671-1815.2302418](https://doi.org/10.12404/j.issn.1671-1815.2302418)]
 - 7 Zhang ST, Shan TM, You ZJ, *et al.* A novel SVD-SDP-CNN fault diagnosis method. Proceedings of the 2024 IEEE International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC). Shijiazhuang: IEEE, 2024. 52–56. [doi: [10.1109/SDPC62810.2024.10707732](https://doi.org/10.1109/SDPC62810.2024.10707732)]
 - 8 Chen W, Wang S, Luo QS, *et al.* Research on coal and gangue recognition method based on improved MobileNetV3 network. Proceedings of the 3rd International Conference on Computer, Control and Robotics (ICCCR). Shanghai: IEEE, 2023. 56–60. [doi: [10.1109/ICCCR56747.2023.10194036](https://doi.org/10.1109/ICCCR56747.2023.10194036)]
 - 9 Singh G, Guleria K, Sharma S. DeepFungusDet: MobileNetV3 model in medical imaging for fungal disease detection. Proceedings of the 3rd International Conference for Innovation in Technology (INOCON). Bangalore: IEEE, 2024. 1–6. [doi: [10.1109/INOCON60754.2024.10511726](https://doi.org/10.1109/INOCON60754.2024.10511726)]
 - 10 Zhang S, Tong Q, Kong ZX, *et al.* TSM-MobileNetV3: A novel lightweight network model for video action recognition. Proceedings of the 4th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT). Nanjing: IEEE, 2023. 228–232. [doi: [10.1109/AINIT59027.2023.10212611](https://doi.org/10.1109/AINIT59027.2023.10212611)]
 - 11 Yang XT, Liu L, Song XF, *et al.* An efficient lightweight satellite image classification model with improved MobileNetV3. Proceedings of the 2025 IEEE INFOCOM-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Vancouver: IEEE, 2024. 1–6. [doi: [10.1109/INFOCOMWKSHPS61880.2024.10620744](https://doi.org/10.1109/INFOCOMWKSHPS61880.2024.10620744)]
 - 12 Zhao LQ, Wang LL. A new lightweight network based on MobileNetV3. KSII Transactions on Internet and Information Systems, 2022, 16(1): 1–15.
 - 13 Wang SQ, Wang YZ, Zheng BS, *et al.* Intrusion detection system for vehicular networks based on MobileNetV3. IEEE Access, 2024, 12: 106285–106302. [doi: [10.1109/ACCESS.2024.3437416](https://doi.org/10.1109/ACCESS.2024.3437416)]
 - 14 Choi K, Kim S, Kim J, *et al.* Hardware-friendly approximation for swish activation and its implementation. IEEE Transactions on Circuits and Systems II: Express Briefs, 2024, 71(10): 4516–4520. [doi: [10.1109/TCSII.2024.3394806](https://doi.org/10.1109/TCSII.2024.3394806)]
 - 15 刘哲, 许超, 熊栋栋. 基于融合注意力机制的盾构姿态组合预测模型研究. 隧道建设(中英文), 2025, 45(1): 139–150.
 - 16 Cai CZ, Guo XY, Xue YF, *et al.* Structure damage diagnosis of bleacher based on DSKNet model. The Journal of Supercomputing, 2024, 80(8): 10197–10222. [doi: [10.1007/s11227-023-05834-8](https://doi.org/10.1007/s11227-023-05834-8)]
 - 17 Hu BQ, Liu J, Xu Y, *et al.* An integrated bearing fault diagnosis method based on multibranch SKNet and enhanced inception-ResNet-v2. Shock and Vibration, 2024, 2024(1): 9071328. [doi: [10.1155/2024/9071328](https://doi.org/10.1155/2024/9071328)]
 - 18 Pan GL, Li J, Li ML. Multi-channel multi-step spectrum prediction using Transformer and stacked Bi-LSTM. China Communications, 2025, 22(5): 1–13.
 - 19 Wu W, Sun DS. Stack LSTM for Chinese image captioning. Proceedings of the 33rd Chinese Control and Decision Conference (CCDC). Kunming: IEEE, 2021. 1613–1617.
 - 20 Ren YW, Zhang RD, Gao FR. A network structure for industrial process fault diagnosis based on hyper feature extraction and stacked LSTM. Chemical Engineering Science, 2024, 287: 119745. [doi: [10.1016/j.ces.2024.119745](https://doi.org/10.1016/j.ces.2024.119745)]
 - 21 Zhen Y, Zhang A, Zhao XM, *et al.* Prediction of deep low permeability sandstone seismic reservoir based on CBAM-CNN. Geoenergy Science and Engineering, 2024, 242: 213241. [doi: [10.1016/j.geoen.2024.213241](https://doi.org/10.1016/j.geoen.2024.213241)]
 - 22 于洋, 马军, 王晓东, 等. 融合深度可分离小卷积核和CBAM的改进CNN故障诊断模型. 电子测量技术, 2022, 45(6): 171–178.

(校对责编: 李慧鑫)