

基于多尺度卷积神经网络的轴承剩余寿命预测^{*}

孙鑫^{1,2}, 孙维堂¹

(1. 中国科学院 沈阳计算技术研究所, 沈阳 110168; 2. 中国科学院大学, 北京 100049)

摘要: 针对滚动轴承剩余寿命难预测的情况, 在分析了轴承原始信号特征提取困难的基础上, 提出了基于多尺度卷积神经网络的轴承剩余寿命预测方法。该方法将原始振动加速度信号作为输入, 依次经过浅层特征提取模块、深层特征提取模块、数据融合模块和输出模块这4部分进行处理, 最后输出预测的剩余寿命。同时提出了一种新型的改进均方误差作为网络的损失函数, 取得了较好的效果。通过对轴承寿命预测实验的测试数据进行预测分析, 该方法能够有效的预测轴承的剩余寿命。

关键词: 多尺度; 卷积神经网络; 轴承; 剩余寿命; 损失函数

中图分类号: TH113; TG502

文献标识码: A

Research on Remaining Life Prediction of Bearing Based on Multi-scale Convolution Neural Network

SUN Xin^{1,2}, SUN Wei-tang¹

(1. Shenyang Institute of Computing Technology, Chinese Academy of Science, Shenyang 110168, China; 2. China University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: It is difficult to predict the remaining life of rolling bearings, because the original signal characteristics of rolling bearings are not obvious, a method for predicting the remaining life of bearings based on multi-scale convolutional neural networks has been proposed. This method takes the original vibration acceleration signal as input, and then processes it through four parts: shallow feature extraction module, deep feature extraction module, data fusion module and output module, finally outputs the predicted remaining life. At the same time, an improved mean square error was proposed as a loss function, which achieved good results. By predicting and analyzing the test data of the bearing life prediction experiment, this method can effectively predict the remaining life of the bearing.

Key words: multiscale; convolutional neural network; bearing; remaining life; loss function

0 引言

当今社会科技不断进步, 数控行业快速发展, 为了满足生产的需求, 机械设备需要不停的运转, 然而设备的长时间运行会出现老化的现象, 滚动轴承作为机械设备最重要的零件之一, 老化现象尤其显著, 所以建立滚动轴承剩余寿命预测模型, 实时的评估轴承的损坏程度, 根据机械的使用要求, 及时制定轴承维修策略, 对保证机械设备健康稳定运行具有重要的意义。

目前, 许多专家学者对滚动轴承的剩余寿命预测技术进行了研究, 雷亚国等^[1]从轴承原始振动信号中提取大量特征, 然后经过特征选择和融合获得新的特征指标, 采用粒子滤波的方式预测轴承的剩余寿命。马海龙^[2]对采集的振动加速度信号数据样本提取有效的峰值、小波熵等特征指标, 采用主元分析融合多个特征指标后进行 SVM 的轴承剩余寿命预测。邱晓梅等^[3]对采集的原始振动加速度信号进行处理, 提取时

域和频域特征, 对特征求取相关系数, 进行特征简约后利用 BP 神经网络来进行寿命预测。上述的轴承寿命预测方式通常需要人为手动的提取时域、频域和统计等特征, 然后进行筛选和融合, 最后经过机器学习等方式进行回归预测, 在这个过程中, 提取特征的操作较繁琐, 需要研究人员投入大量的精力, 其次使用机器学习等方式进行回归预测的精度并不高。

针对上述问题, 本文提出了基于多尺度卷积神经网络的轴承剩余寿命预测方法, 利用多尺度卷积神经网络进行轴承剩余寿命预测, 可以免去传统方法中的特征提取、筛选和融合等步骤; 同时提出了一种效果更好的模型损失函数。实验结果表明, 利用该方法对轴承进行寿命预测具有更高的精度。

1 传统的卷积神经网络

卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络, 本质是一个多层感知机, 采用局部

收稿日期: 2019-12-13; 修回日期: 2020-01-22

^{*} 基金项目: “高档数控机床与基础制造装备”国家科技重大专项课题: 航空发动机典型零件加工设备国产数控系统换脑工(2017ZX04011004)

作者简介: 孙鑫(1993—), 男, 山东菏泽人, 中国科技大学硕士研究生, 研究方向为人工智能和数控机床 (E-mail: 84996429@qq.com); 通讯作者: 孙维堂(1975—), 男, 山东德州人, 中科院沈阳计算技术研究所研究员, 硕士研究生导师, 博士, 研究方向为数控技术及先进制造与自动化的研究与应用。

连接和权值共享的方式,一方面这种结构可以自动提取特征,方便下一步的计算,另一方面也减少了模型的复杂度,使得模型更容易学习,减少了过拟合的风险^[4-6]。传统的卷积神经网络包括输入层、卷积层、激励层、池化层、全连接层,其中卷积层、激励层、池化层构成特征提取模块,可以从原始信号中提取训练特征。本文中用作对比参考的传统卷积神经网络的结构共 16 层,具体包括 1 个输入层、3 个卷积层、3 个激励层、3 个池化层、1 个扁平化层、4 个全连接层、1 个输出层。

2 多尺度卷积神经网络

传统的卷积神经网络可以直接处理传感器提取的振动加速度信号,提取隐藏在原始信号中的特征,但是原始信号的数据量比较大,使用单一大小的卷积核会导致网络模型在自适应筛选特征的过程中遗漏局部重要特征,导致模型的预测精度低下。

因此本文为改进传统卷积神经网络的不足,在其基础上,提出一种多尺度卷积神经网络模型,其中多尺度体现在:该网络模型可以利用多个不同尺度的卷积核进行卷积操作,以达到提取详细特征,增加模型预测精度的效果。

多尺度卷积神经网络通过对原始信号内隐藏的特征进行多尺度的提取,深度学习原始信号和剩余寿命之间的映射关系,达到预测寿命的效果。具体的网络结构如图 1 所示。

该网络具体包括 5 个部分:输入模块、浅层特征提取模块、深层特征提取模块、数据融合模块、输出模块。

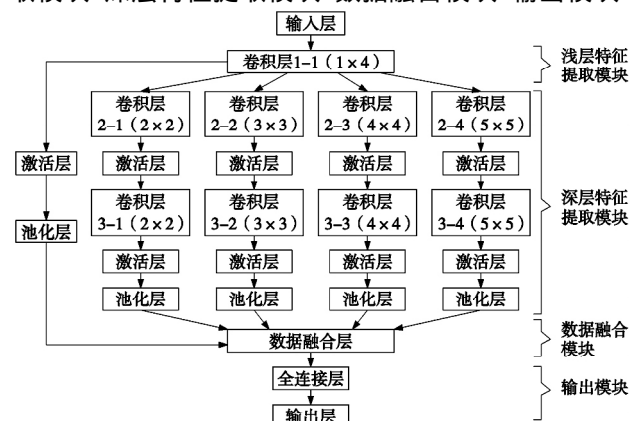


图 1 多尺度卷积神经网络结构图

2.1 浅层特征提取模块

浅层提取模块可对输入模块的数据进行处理,其中输入模块的数据为:轴承全寿命周期的原始振动信号。

浅层特征提取模块包含一个卷积层(命名为卷积层 1-1),作用是提取传感器测量的原始信号的浅层特征,该卷积层使用大小为 1×4 的卷积核,卷积核的数量设置为 64,卷积步长设置为 1,并通过补零的措施保留边界的卷积结果。卷积层 1-1 的 64 个卷积核通过权值共享和局部连接的方式与输入模块的数据进行运算,以实现浅层特征的学习,其中卷积公式^[7]为:

$$M_j = \sum_{i=1}^k Z_i * W_{ij} + b_j \quad (1)$$

其中 M_j 表示经过卷积操作后的第 j 个输出特征图; Z_i 表示与卷积核相连的输入单元的第 i 个特征图; W_{ij} 为

卷积核,表示为与第 j 个特征图相连的所有输入单元中第 i 个单元的权重阵; $*$ 为卷积操作; b_j 表示每个输出浅层特征图的偏置项。

通过浅层特征提取模块中的卷积操作,提取原始信号浅层特征,此外原始信号中还包含很多深层的细节信息,但仅仅靠简单的单一尺度卷积核进行卷积操作,这些深层信息是无法提取出来的,所以该神经网络模型引入深层特征提取模块。

2.2 深层特征提取模块

深层特征提取模块的核心部分是卷积层采用多尺度的卷积核,多尺度卷积核的优势为:让卷积层的滤波器具有多样性,从而使权重和偏置的学习更加多样,进而可以充分有效的提取并学习振动加速度信号的有用信息。

该深层特征提取模块首先使用了 4 组不同尺度卷积核的卷积层(命名为 2-1 ~ 2-4)进行并行的卷积滤波操作,卷积核的大小分别为 2×2 、 3×3 、 4×4 、 5×5 ,然后使用激活函数分别对卷积层的结果进行激活操作,接着把激活结果分别输入下一层多尺度卷积核的卷积层(命名为 3-1 ~ 3-4),卷积核的大小和上层卷积层相同,经过并行卷积运算提取的深层特征再次进行激活操作,最后把激活结果并行输入池化层进行池化操作。多尺度卷积操作的公式为:

$$M_{n,l+1,q} = \sum_{i=1}^k Z_{i,l} * W_{i,n,l+1,q} + b_{n,l+1,q} \quad (2)$$

其中, $M_{n,l+1,q}$ 表示为第 $l+1$ 卷积层通过不同尺度卷积运算后输出的 n 个深层特征图; $W_{i,n,l+1,q}$ ($q=2,3,4,5$) 表示多尺度卷积的 4 组不同大小的卷积核。其中卷积核的个数设置为 64,这样每组卷积核运算后得到 64 个特征图。

由于上述的多尺度卷积操作为线性变化,而线性操作的表达能力有所欠缺,所以为了更好的学习模型参数,需要引入激活层,通过激活层的非线性激活函数来处理多尺度卷积后的结果。本文选择 Relu 函数作为激活函数,其定义如公式(3)所示:

$$f(x) = \max(0, x) \quad (3)$$

池化层可对上层的输出起到降采样的作用,不但可以有效的加速模型的训练,而且还能减少过拟合现象。本文池化层采用全局最大池化,可对输入形如 (samples, steps, features) 的 3D 张量,输出形如 (samples, features) 的 2D 张量。

2.3 数据融合模块

要想准确的预测寿命,就需要综合原始信号的浅层特征和深层特征,因此数据融合模块变得十分重要。

本论文的数据融合模块借助 Resnet 网络^[8]中残差块跳跃连接的思想,把浅层特征提取模块中卷积层的结果经过激励和最大池化操作后,与深层特征提取模块中最后形成的 4 组并行结果跳跃连接合并。合并过程即为深层特征和浅层特征的拼接 (Concat):

$$Y = \text{Concat}(f(a), b, c, d, e) \quad (4)$$

上述公式中 $f(a)$ 为浅层提取模块中的输出进行激励和池化操作, b, c, d, e 为深层提出模块的输出结果, Y 为深层特征和浅层特征拼接的结果。

本文这种跳跃连接的方式不但能够整合深层和浅层特征,而且能解决因网络层数较深而导致的梯度消

失的问题。

2.4 输出模块

输出模块包括 4 个全连接层和一个输出层,其中全连接层的节点个数分别为 1024、512、256、128,激活函数为 ReLU,输出层的节点个数为 1,考虑到全连接层的参数较多,为了增强模型的泛化能力,防止过拟合现象的发生,因此在训练学习的过程中使用 Dropout 方法,随机丢弃神经元的概率设置为 0.5。

2.5 损失函数

对于传统的回归问题,当网络模型中最后的输出层的节点个数为 1 时,其输出值就是网络的预测值。通常使用均方误差(MSE)作为损失函数,具体定义如公式(5):

$$MSE(y, y') = \frac{\sum_{i=1}^n (y_i - y_i')^2}{n} \quad (5)$$

其中 y_i 为第 i 个样本的真实值, y_i' 为第 i 个样本的预测值。

但是对于本文中滚动轴承剩余寿命的预测问题,由于预测结果可能高于真实值,也可能低于真实值,而当预测结果小于真实值时,可及时的维修或更换零件,即提前预知损坏相比于延迟预知损坏有着更低的风险和代价,因此在此基础上对均方误差(MSE)进行修正,提出改进的均方误差损失函数,定义如下:

$$MSE(y, y') = \frac{\sum_{i=1}^n \alpha (y_i - y_i')^2}{n} \quad (6)$$

其中, y_i 为第 i 个样本的真实值, y_i' 为第 i 个样本的预测值, α 为修正参数,当 $y_i - y_i'$ 大于 0 时 α 为 1,当 $y_i - y_i'$ 小于等于 0 时 α 为 2。

从公式中的修正参数可以看出:对于预测值大于真实值,将有更大的代价,因此使用修正的均方误差作为损失函数,使得预测的结果更加准确,更能做到提前预知损坏,保证机械设备健康运行。

3 实验结果及分析

3.1 数据来源

本文的实验数据来源于 IEEE 2012 年的 PHM 数据挑战赛^[9],它由法国弗朗什孔泰大学的 PRONOSTIA 实验平台测量得到。该数据集中包含水平和垂直两个方向上的振动加速度信号,根据文献[10]的相关说明,对于轴承退化程度的研究,水平振动信号比垂直振动信号提供更多的有用信息,因此本文仅采用水平振动信号进行实验。该数据集包括如表 1 所示的三种不同工况下的 17 个滚动轴承的全寿命周期的振动加速度数据,平台中传感器每 10s 采集一次数据,采样频率为 25.6kHz,一次采样的时间为 0.1s,所以每次采集到 2560 个振动加速度,如图 1 所示。

表 1 工况详细信息

工况	转速/(rpm)	载荷/(N)
工况 1	1800	4000
工况 2	1650	4200
工况 3	1500	5000

本文模型使用第一组工况中的前 6 个轴承测量的数据作为训练集,最后 1 个轴承的数据作为测试集,对于每个实验样本(X_i, Y_i), X_i 作为特征输入,为第 i 次

采集到的 2560 个振动加速度数值。 Y_i 作为标签,为当前时刻和失效时刻之间的时间间隔与起始时刻和失效时刻之间的时间间隔的比值(具体数值在 0~1 之间)。

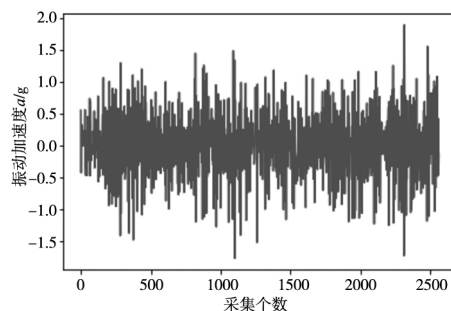


图 2 原始振动加速度信号

3.2 实验结果

首先,根据本文第 2 小节的说明对多尺度卷积神经网络设置参数,损失函数选为改进的均方误差,其中因为标签的值在 0~1 之间,所以最后输出层的激活函数为 sigmoid 函数,接着,使用训练集训练模型,其中,输入信号为 2560×1 的一维振动加速度信号,经过浅层特征提取模块中的卷积层输出后,维度变为 2560×64 ,以此类推,最后输出预测的剩余寿命。最终,使用测试数据进行测试验证,最终结果如图 3 所示。

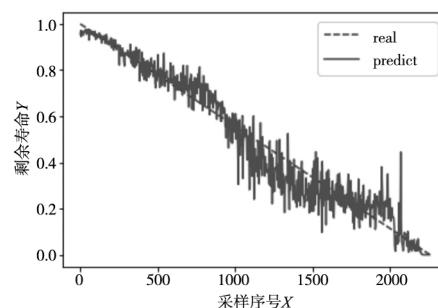


图 3 应用改进损失函数的多尺度卷积神经网络预测结果

3.3 结果对比分析

将本文中使用改进均方误差作为损失函数的模型与使用传统均方误差作为损失函数的模型进行对比,传统均方误差作为损失函数的模型参数细节和本文模型一致,使用模型对测试集预测的结果如图 4 所示。

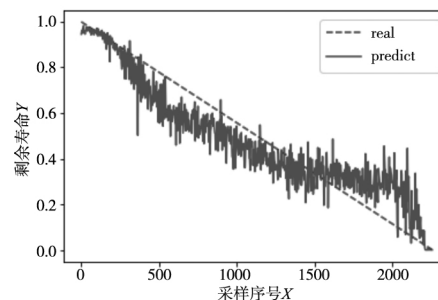


图 4 应用传统损失函数的多尺度卷积神经网络预测结果

对比上小节的图像,发现改进损失函数的模型曲线的拟合程度更好,在采样序号大于 1200 时,改进损失函数模型拟合的曲线相较于原始损失函数模型的曲线,整体有下移的趋势,说明预测的寿命相对较小,更能超前预知轴承的损坏,体现了改进损失函数的优越性。

将本文中的多尺度卷积神经网络模型和传统的卷积神经网络模型进行对比,传统的卷积神经网络模型的细节参考本文第 1 小节,使用传统卷积神经网络对测试数据预测的结果如图 5 所示。

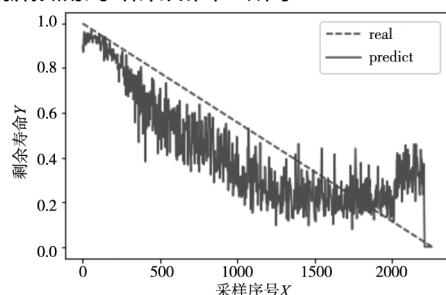


图 5 传统卷积神经网络预测结果

该图像对比多尺度卷积神经网络模型的图像,发现该图像曲线的拟合效果较差,整体波动明显,而多尺度卷积神经网络模型拟合的曲线更趋向于真实值,说明多尺度卷积神经网络模型自适应提取到的特征比传统卷积神经网络提取的特征更详细、精准,多尺度卷积神经网络的模型更加复杂,预测结果更加精确,优于传统卷积神经网络模型。

为了更加客观的对比上述的三个模型以及传统的支持向量回归(support vector regression, SVR)预测模型^[11](SVR 方法的性能退化指标设置为峭度和均方根)。本文采用预测结果相对于总寿命的均方误差以及平均绝对误差作为评价标准^[12],均方误差(MSE)和平均绝对误差(MAE)的公式分别如下:

$$MSE(y, \hat{y}) = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (7)$$

$$MAE(y, \hat{y}) = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (8)$$

4 种不同模型对测试集预测的误差具体如表 2 所示。

表 2 3 种模型预测误差

模型名称	MSE	MAE
SVR	0.078	0.071
传统卷积神经网络	0.055	0.045
基于传统损失函数的多尺度卷积神经网络	0.032	0.027
基于改进损失函数的多尺度卷积神经网络	0.028	0.024

从表 2 的数据可以看出,使用改进损失函数的多尺度卷积神经网络进行预测的均方误差和平均绝对误差是最小的。

基于改进损失函数的多尺度卷积神经网络比基于传统损失函数的多尺度卷积神经网络的预测均方误差减少了 12.5%,平均绝对误差减少了 11.1%;比基于传统卷积神经网络的预测均方误差减少了 49.1%,平均绝对误差减少了 46.7%;比 SVR 模型的预测均方误差减少了 64.1%,平均绝对误差减少了 66.1%。这一结果证明:传统的 SVR 方法与神经网络方法相比,在预测精度上有比较大的差距,证明了深度学习方法的优越性;多尺度卷积神经网络比较适用于滚动轴承剩余寿命预测问题;改进的损失函数在寿命预测模型中能发挥更好的作用。

为了验证多尺度卷积神经网络的鲁棒性,本文同

时预测了工况 2 和工况 3 下的轴承的剩余寿命,其预测结果接近真实剩余寿命值,预测值和真实值的均方误差和平均绝对误差较小,结果证明:多尺度卷积神经网络模型的稳定性较高,鲁棒性较好。

4 结论

本文提出的基于多尺度的卷积神经网络应用于滚动轴承剩余寿命预测,取得了非常好的效果,主要结论如下:

(1) 多尺度卷积神经网络模型利用多尺度卷积核和跳跃连接的方式可以直接从原始信号中提取深层和浅层特征,省去人工提取特征的步骤。

(2) 本文提出的改进损失函数使模型取得了比较好的效果。

(3) 多尺度卷积神经网络模型的结构比较复杂,预测精度更高,均方误差和平均绝对误差比较小。

[参考文献]

- [1] 雷亚国. 基于粒子滤波的滚动轴承寿命预测方法[A]. 中国机械工程学会可靠性工程分会 2014 年全国机械行业可靠性技术学术交流会暨可靠性工程分会第五届委员会成立大会论文集[C]. 中国机械工程学会可靠性工程分会: 中国机械工程学会, 2014.
- [2] 马海龙. 基于主元特征融合和 SVM 的轴承剩余寿命预测[J]. 工矿自动化, 2019, 45(8): 74-78.
- [3] 邱晓梅, 隋文涛, 王峰, 等. 基于相关系数和 BP 神经网络的轴承剩余寿命预测[J]. 组合机床与自动化加工技术, 2019(4): 63-65.
- [4] 王海龙, 夏筱筠, 孙维堂. 基于 EMD 与卷积神经网络的滚动轴承故障诊断[J]. 组合机床与自动化加工技术, 2019(10): 46-48, 52.
- [5] 宁永杰. 基于机器学习的滚动轴承状态评估与剩余寿命预测[D]. 徐州: 中国矿业大学, 2019.
- [6] 胡晓依, 荆云建, 宋志坤, 等. 基于 CNN-SVM 的深度卷积神经网络轴承故障识别研究[J]. 振动与冲击, 2019, 38(18): 173-178.
- [7] 李少鹏. 结合 CNN 和 LSTM 的滚动轴承剩余使用寿命预测方法研究[D]. 哈尔滨: 哈尔滨理工大学, 2019.
- [8] David McNeely-White, J Ross Beveridge, Bruce A Draper. Inception and ResNet features are (almost) equivalent[J]. Cognitive Systems Research, 2020, 59: 312-318.
- [9] Patrick Nectoux, Rafael Gouriveau, Kamal Medjaher, et al. PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test [C]//IEEE International Conference on Prognostics and Health Management, Denver, CO, USA, 2012.
- [10] Soualhi A, Medjaher K, Zerouni N. Bearing health monitoring based on Hilbert-Huang transform, support vector machine and regression[J]. IEEE TransInstrum Meas, 2015; 64(1): 52-62.
- [11] 申中杰, 陈雪峰, 何正嘉, 等. 基于相对特征和多变量支持向量机的滚动轴承剩余寿命预测[J]. 机械工程学报, 2013, 49(2): 183-189.
- [12] 张继冬, 邹益胜, 邓佳林, 等. 基于全卷积层神经网络的轴承剩余寿命预测[J]. 中国机械工程, 2019, 30(18): 2231-2235.

(编辑 李秀敏)