第 35 卷 第 2 期 石家庄铁道大学学报(自然科学版) Vol. 35 No. 2
2022 年 6 月 Journal of Shijiazhuang Tiedao University(Natural Science Edition) Jun. 2022

旋转机械故障诊断模型的泛化性能研究与改进

张晓锋， 郝如江， 夏晗铎， 段泽森， 程 旺

(石家庄铁道大学 机械工程学院,河北 石家庄 050043)

摘要:泛化能力是机器学习的重要评价指标,由于现实中的工业应用中外部环境十分复杂,在旋转机械工作时,其负载是不断变化的。如何使用一种负载下训练好的网络模型对另一种负载的数据集进行故障诊断是目前众多学者的研究重点。针对网络模型训练与测试时的数据特征分布不同来进行网络泛化性能的改进研究,通过改进网络模型的首层卷积核尺寸与随机失活来加强网络特征提取能力,采用 AdaBN 优化算法使得网络模型在训练与测试时的参数进行调整,提高网络的泛化能力。

关键词:故障诊断;特征提取;卷积核尺度;泛化能力

中图分类号: TH17 **文献标志码**: A **文章编号**: 2095-0373(2022)02-0108-05

0 引言

旋转机械的负载变化在工作中很常见,在负载变化时传感器采集到的信号也会发生改变,使得信号的特征分布变得更加分散。故障类型在不同负载下的信号特征分布不同,幅值大小与周期波动也不同[1],这样的特征分布问题会使网络模型在故障分类时产生分类误差,影响网络模型的故障诊断性能。而负载变化在工业条件下是不可避免的,所以优秀的故障诊断模型需要在负载不断变化的情况下也能够保持良好的故障诊断性能[2]。这要求网络模型具有良好的泛化能力[3],能够在不同负载造成的差异数据集中保持优秀的学习能力与测试能力[4]。

1 网络结构改进

1.1 首层卷积核尺度

卷积神经网络的首层卷积核尺寸对网络的性能影响较大,因为振动信号是周期信号,每个输入到网络的数据信号的相位值不同。首层卷积层相当于一组带通滤波器,滤波器的频带范围直观地影响故障特征提取,所以神经网络对首层卷积层十分敏感。首层卷积核的尺寸越大,频带范围越大,此时提取的信号特征较多,但是包含负载变化产生的冗余信息也多;首层卷积和尺寸越小,此时提取的信号特征较少,但是包含的负载变化产生的冗余信息也少。所以为了寻求到最合适的首层卷积核尺寸,拟采用多种尺度大小的首层卷积核来测试相同的实验数据,通过对比找到最适合变负载下齿轮箱故障诊断的首层卷积核尺寸。选取 1×3、1×9、1×64、1×128、1×256、1×512 共 6 种不同的首层卷积核尺寸。

使用美国凯斯西储大学(Case Western Reserve University, CWRU)[5]的数据来验证不同负载下的故障诊断准确率。数据负载分别为 750、1 500、2 250 W 3 种状态,每种状态都包含 10 种故障类型。分别在 750、1 500、2 250 W 以及无负载状态下测试上述 6 种首层卷积核尺度。为了形象直观,计算每种类型卷积核的平均准确率,绘成柱状图,见图 1。

可以发现当首层卷积核尺度为 1×128 的时候,平均故障分类准确率最高。说明在数据集下,此尺度下提取的信号特征最为合适,因此,选用尺度为 1×128 作为改进网络首层卷积核尺度。

收稿日期:2022-03-31 责任编辑:车轩玉 DOI:10.13319/j.cnki.sjztddxxbzrb.20220079
基金项目:国家自然科学基金(11872256)
作者简介:张晓锋(1997—),男,硕士研究生,研究方向为机电系统故障诊断。E-mail:1318274182@qq.com
张晓锋,郝如江,夏晗铎,等. 旋转机械故障诊断模型的泛化性能研究与改进[J]. 石家庄铁道大学学报(自然科学版),2022,35(2):108-112.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1.2　随机失活

随机失活(Dropout)是 Krizhevsky et al[6]提出的,在网络层数较深时,训练后会出现过拟合现象,可以通过影响特征检测器来抑制过拟合现象。Dropout 可以在神经网络的前向传播过程中,在 dense 层之前使每个批次以一定的概率 P 来忽略一部分特征检测器,让某个神经元的激活值以 P 的概率停止工作,从而使网络不会太过于依赖某些局部的特征,同时减少了网络的计算参量,使模型泛化性更强[7]。

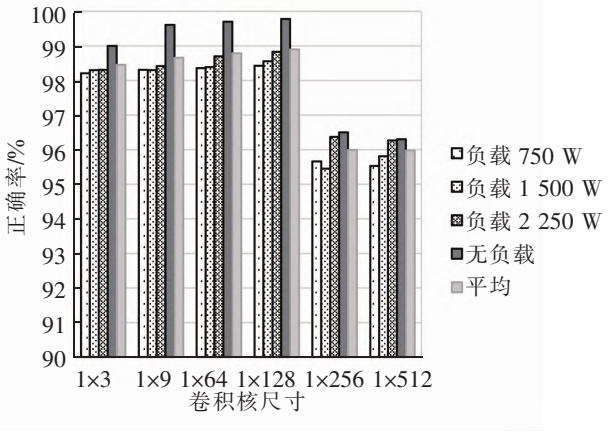

图 1　不同首层卷积核尺寸验证对比

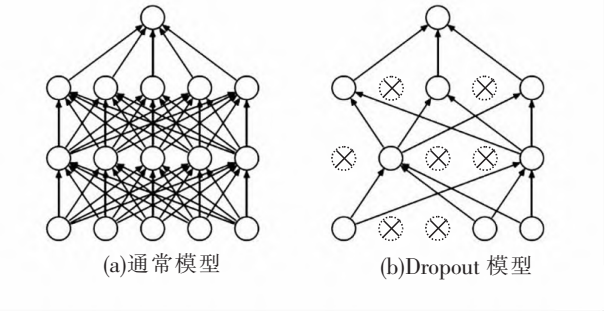

图 2　网络模型对比

使用 Dropout 的神经网络训练过程如图 2 所示。首先,临时随机删除网络中一定数目的隐藏神经元,据 Srivastava et al[8]所提出的理论与多次实验论证,失活概率采用 0.5。图 2 中虚线为临时删除的神经元,同时输入输出的神经元保持不变;然后,把输入 x 通过随机失活后的网络进行前向传播,把得到的损失结果通过随机失活后的网络进行反向传播,一小批训练样本执行完这个过程后,在没有被删除的神经元上按照随机梯度下降法更新对应的参数(w,b);最后,恢复被删掉的神经元,此时被删除的神经元保持原样,而没有被删除的神经元已经有所更新,不断更新没有被删除的那一部分参数,删除的神经元参数保持不变;不断重复这一过程。

整个 Dropout 过程就相当于对很多个不同的神经网络取平均,同时减少权重使得网络对丢失特定神经元连接的鲁棒性也得到了提高,从而使得网络在整体上达到减少过拟合的目的。

1.3　测试分析

为了验证基于多尺度特征融合与深度残差收缩网络的泛化能力,现采用 3 种负载下的数据来进行训练验证实验。具体操作为使用其中一种负载训练网络模型,将网络模型保存后对其余 2 种负载进行准确率测试,以此来验证网络的泛化能力。使用负载为 750 W 的数据集来训练网络模型,将模型保存后对负载为 1 500、2 250 W 的数据集进行测试,将其准确率分别记录下来,然后依次进行负载为 1 500、2 250 W 的模型训练与之后的交叉测试。其测试结果如表 1 所示。

表 1　泛化能力测试结果

训练集数据/W	测试集数据/W	识别准确率/%
750	1 500	97.47
	2 250	91.70
1 500	750	98.06
	2 250	92.94
2 250	750	88.90
	1 500	93.70

可以发现在 2 250 W 数据集下训练的网络模型对 750 W 数据集的故障识别正确率最低,为 88.91%,当使用 1 500 W 下数据集来训练网络模型时,对其余 2 种数据集进行故障识别,此时的故障识别准确率最为稳定且精度较高。以上结果表明,改进网络的跨负载领域自适应能力很强,为适应多变的工作环境提供了合理性。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

2 优化算法改进

网络模型在训练过程中,学习初期迭代很快,训练效果很好,而且没有发生训练变化异常等现象;结合表 1 发现,虽然训练的模型故障诊断效果良好,但是在测试时的故障诊断准确率不太理想,为此使用数据降维来分析此想象的具体原因并采用优化算法来增强泛化能力。

2.1 *t*-SNE 数据降维

为了查出使用 750 W 数据集训练模型来测试 2 250 W 数据集与使用 2 250 W 数据集训练模型来测试 750 W 数据集正确率偏低的原因,采用 *t*-SNE[9] 进行数据降维,*t*-SNE 图中不同的颜色代表不同的故障类型(见图 3)。选用 750 W 数据集训练模型来测试 2 250 W 数据集来进行降维,其最后一层隐含层的分类结果如图 3(a)所示,可以发现在测试 2 250 W 数据集有 3 种故障存在明显的混叠现象[10],因为此 3 种故障分布具有相似性,使得网络模型在分类故障时造成错误分类,这是导致准确率较低的主要原因。而训练网络模型过程中的降维分布如图 3(b)所示。

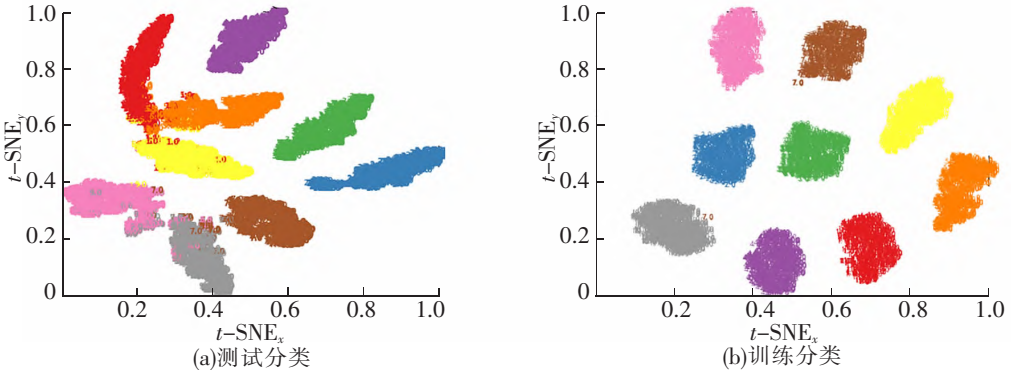

图 3 特征分布对比图

可以发现,网络模型虽然在 750 W 训练数据集中具有良好的可分性,但是将模型应用到 2 250 W 时却发生少许的分布不一致现象,说明当训练模型的数据集与测试模型的数据集相差较大时会使得网络性能下降,故障分类正确率受到影响。为此拟采用 AdaBN 算法来解决这一问题,提高网络的泛化能力。

2.2 AdaBN 优化算法

AdaBN 算法是 Lu et al[11] 针对训练集与测试集特征分布相差较多时对 BN(Batch Normalization)进行改进的一种算法,以此来缓解神经网络性能退化这一问题,一般用于图像识别领域[12]。该算法在网络的每一次批量归一化过程中将训练集(目标域)在 BN 层的均值与方差一并记录,在遇到训练集的特征分布与测试集的特征分布不一致时,将测试集数据输入到改进网络中,此时网络不再进行反向传播只进行正向传播。此时可以用之前记录下来的训练集方差与均值来代替测试集在进行 BN 时的均值与方差,其工作过程如图 4 所示。

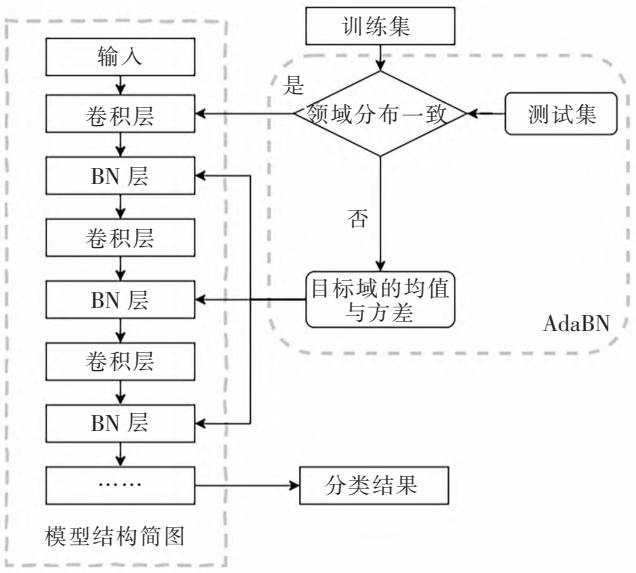

图 4 AdaBN 算法工作流程

在计算过程中,目标领域的信号 p,在改进网络模型的 BN 层的第 i 个神经元的表

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

达为 $x_t^{(i)}(p)\in x_t^{(i)}$,其中,$x_t^{(i)}=\{x_t^{(i)}(1),\cdots,x_t^{(i)}(n)\}$;$i$ 为目标领域内的每个神经元;p 为每个信号,计算目标领域中所有的样本均值和方差,其计算公式为

$$\mu_t^{(i)}\leftarrow E[x_t^{(i)}] \tag{1}$$

$$\sigma_t^{(i)}\leftarrow\sqrt{Var[x_t^{(i)}]} \tag{2}$$

之后网络将参数输入到后续网络,其他参数不再变化,由此,神经网络可以使源领域与目标领域的特征分布近似一致,从而达到缓解网络性能退化的目的,计算 BN 层的输出

$$\hat{x}_t^{(i)}(p)=\frac{x_t^{(i)}(p)-\mu_t^{(i)}}{\sigma_t^{(i)}} \tag{3}$$

$$y_t^{(i)}(p)=\gamma_s^{(i)}\hat{x}_t^{(i)}(p)+\beta_s^{(i)} \tag{4}$$

在神经网络中添加 AdaBN 优化算法,重复实验,结果如表 2 所示。

表 2　使用 AdaBN 优化算法测试结果

训练集数据/W	测试集数据/W	识别准确率/%
750	1 500	98.87
	2 250	95.10
1 500	750	98.36
	2 250	96.14
2 250	750	93.91
	1 500	94.70

其中在负载为 750 W 的情况下训练时的正确率与损失的变化如图 5 所示。

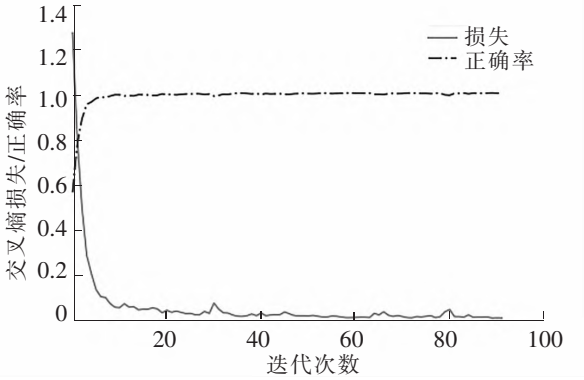

图 5　实验测试结果图

可以发现使用 AdaBN 优化算法后,将训练集在 BN 层的均值与方差代替测试集的均值与方差时,通过不同负载下的交叉训练验证,表明此种改进能提高网络模型泛化能力。可以发现网络模型的故障诊断正确率明显提高,在使用 750、2 250 W 的数据集进行交叉训练与验证时正确率提升最为明显,正确率分别上升 3.4%、5.0%。通过使用 AdaBN 优化算法,可以使网络模型在训练时 750 W 数据集与 2 250 W数据集的特征分布基本吻合,故而模型可以很好地进行分类。综上所述,AdaBN 优化算法可以很好地提升网络在不同负载下的故障诊断性能,使得基于多尺度特征融合与残差收缩网络具有优良的泛化能力。

3　结论

在凯斯西储大学的数据基础上进行测试之后得到以下结论:

(1)选择合适的首层卷积核的尺寸可以提高网络模型的故障诊断正确率,而且利用随机失活能够使得网络模型的泛化能力得到很大提升,这对提升网络的故障诊断性能提供了支持。

(2)AdaBN 优化算法可以很好地提升网络在不同负载下的故障诊断性能,使得网络模型具有优良的泛化能力,实现了一种负载下训练好的网络模型对另一种负载的数据集进行故障诊断这一需求,满足在不同负载下进行旋转机械的故障诊断研究。

(3)为了验证网络更广泛的适应能力,对网络模型及其算法参数进行改进,使其满足负载差距更大的交叉验证诊断要求,同时探究网络在其他类型故障诊断的应用及改进。

参　考　文　献

[1]武妍,张立明. 神经网络的泛化能力与结构优化算法研究[J]. 计算机应用研究,2002(6):21-25,84.

[2]魏海坤,徐嗣鑫,宋文忠. 神经网络的泛化理论和泛化方法[J]. 自动化学报,2001,27(6):806-815.

[3]Patel V M, Gopalan R, Li R, et al. Visual domain adaptation: a survey of recent advances[J]. IEEE Signal Processing Magazine, 2015, 32(3): 53-69.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

[4]Li Y, Wang N, Shi J, et al. Revisiting batch normalization for practical domain adaptation[J]. ArXiv Preprint ArXiv, 2016(3): 21-23.

[5]The Case Western Reserve University Bearing Data Center. Bearing data center fault test data [EB/OL]. [2012-03-01]. http://www.eecs.cwru.edu/laboratory/beaning.

[6]Krizhevsky A, Sutskever I, Hinton G E. Imagenet classfication with deep convolutional nearal networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[7]祝道强,周新志,宁芊. 变负载轴承故障诊断卷积神经网络模型[J]. 科学技术与工程, 2020, 20(15): 6054-6059.

[8]Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.

[9]葛阳,郭兰中,牛曙光,等. 基于 t-SNE 和 LSTM 的旋转机械剩余寿命预测[J]. 振动与冲击, 2020, 39(7): 223-231.

[10]郭艳萍,高云,吕丙东. 基于 t-SNE 降维的数据可视化分析研究[J]. 电子技术与软件工程, 2021(21): 166-170.

[11]Lu W, Liang B, Cheng Y, et al. Deep model based domain adaptation for fault diagnosis[J]. IEEE Transactions on Industrial Electronics, 2017, 64(3): 2296-2305.

[12]郑远攀,李广阳,李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55(12): 20-36.

Research and Improvement on Generalization Performance of Rotating Machinery Fault Diagnosis Model

Zhang Xiaofeng, Hao Rujiang, Xia Handuo, Duan Zesen, Cheng Wang

(School of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

Abstract: Generalization ability is an important evaluation index of machine learning. Due to the complex external environment in real industrial applications, when rotating machinery is working, its load is constantly changing. How to use a network model trained under one load to perform fault diagnosis on a data set under another load is the focus of many scholars' research. Therefore, this paper mainly studied the improvement of the generalization performance of the network based on the different distribution of data features during training and testing of the network model, and strengthened the network feature extraction ability by improving the size of the first-layer convolution kernel and random inactivation of the network model, using the AdaBN optimization algorithm to adjust the parameters of the network model during training and testing to improve the generalization ability of the network.

Key words: fault diagnosis; feature extraction; convolution kernel scale; generalization ability

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net