

基于 GRU 与迁移学习的滚动轴承故障诊断

曹梦婷, 谷玉海, 王红军, 徐小力

(北京信息科技大学现代测控技术教育部重点实验室, 北京 100192)

摘要:现阶段基于深度学习的故障诊断需要大量的数据,而制作数据集是一项耗时耗力的工作。针对这一缺点,提出一种基于门控循环单元(Gate Recurrent Unit, GRU)与迁移学习的滚动轴承故障诊断方法。该方法利用与目标域特征相似且易获得源域数据的特点训练网络,确定网络结构和参数,冻结经过训练的卷积神经网络(Convolutional Neural Networks, CNN)和 GRU,用小样本目标域数据训练该网络,微调全连接层和分类层,达到迁移的目的。实验对比分析表明,基于 GRU 与迁移学习的滚动轴承故障诊断方法明显优于基于 BP 神经网络和基于概率神经网络(Probabilistic Neural Network, PNN)方法的故障诊断,能够更加准确地进行故障分类,为小样本数据集下的故障诊断提出了新思路。

关键词:迁移学习;故障诊断;滚动轴承;一维卷积神经网络;门控循环单元

中图分类号:TH66;TG506 **文献标志码:**A **文章编号:**1671-3133(2022)01-0143-05

DOI:10.16731/j.cnki.1671-3133.2022.01.021

Rolling bearing fault diagnosis based on GRU and transfer learning

CAO Mengting, GU Yuhai, WANG Hongjun, XU Xiaoli

(Key Laboratory of Modern Measurement & Control Technology Ministry of Education,
Beijing Information Science & Technology University, Beijing 100192, China)

Abstract: At present, fault diagnosis based on deep learning requires a lot of data, and making data sets is a time-consuming and labor-consuming work. Aiming at this shortcoming, a rolling bearing fault diagnosis method based on Gate Recurrent Unit (GRU) and migration learning was proposed. The method use the characteristics which similar to the target domain and easy to obtain the source domain data, the network trained, the network structure and parameters determined, the trained Convolutional Neural Networks (CNN) and GRU were froze, the network with small sample target domain data trained, the full connection layer and classification layer was fine tuned, and classification layer to achieve the purpose of migration was achieved. The experimental comparative analysis shows that the rolling bearing fault diagnosis method based on GRU and transfer learning is obviously superior to the fault diagnosis method based on BP neural network and Probabilistic Neural Network (PNN), which can classify the fault more accurately, and put forward a new idea for fault diagnosis under small sample data set.

Keywords: transfer learning; fault diagnosis; rolling bearing; 1D-convolution neural networks; Gate Recurrent Unit (GRU)

0 引言

滚动轴承是机械设备中常见的组成元件,从简单的电风扇到复杂的机床都有滚动轴承的应用。事实上,超过 50 % 的机械缺陷与轴承故障有关,从而导致机器停产、停机,甚至造成人员的伤亡^[1-2];因此,滚动轴承故障诊断是机械故障诊断的一个重要方面,也是近些年来研究热点。近些年来,随着学者们的不断深入研究,提出了各种故障诊断的方法。传统的故障诊断研究方法,如 BP 神经网络^[3]等得到了广泛的应用。

目前基于机器学习的故障诊断方法主要有支持

向量机(Support Vector Machine, SVM)^[4]、人工神经网络(Artificial Neural Network, ANN)^[5]及模糊推断^[6]。

近些年来,将深度学习运用到故障诊断领域也成为研究的热点。LU 等人^[7]研究了分级 CNN 对滚动轴承故障的智能诊断,结果证明了 CNN 模型对滚动轴承故障分类的有效性。WEN 等人^[8]采用将信号转化为二维图像后输入 LeNet-5 网络结构进行训练的方式对故障进行分类。将深度学习运用到故障诊断领域已经相对成熟和广泛,深度学习需要大量的数据集,然而大多数情况下,数据集样本较少,标签数据获取成本较高,很难同时满足需求。

迁移学习在近些年得到广泛研究,其主要解决通

过已知的领域知识(源域)求解相关领域(目标域)的问题。作为一种机器学习方法,迁移学习能够在数据集不足的情况下,取得令人满意的结果;因此本文提出一种基于 GRU 与迁移学习的滚动轴承故障诊断方法,CNN 用来提取特征,随后的 GRU 神经网络层有助于模型捕捉特征图中存在的时间动态,故障的分类取决于通过全连接层的最终输出。

首先利用源域数据集进行对比实验,确定 GRU 和 CNN 的网络结构和相关参数,然后采用目标域数据集对确定好的网络结构进行微调,从而确定整个目标域的网络架构及参数。

1 基础知识

1.1 一维卷积神经网络

卷积神经网络是受生物学影响发展而来的前反馈神经网络,具有对输入信息进行位置感知的能力,从而实现特征提取。根据卷积神经网络输入类型不同,可以将其分为一维卷积神经网络和二维卷积神经网络,本文采用一维卷积神经网络对数据集(长序列)进行特征提取。

1.2 门控循环单元

门控循环单元由经典的循环神经网络——长短期记忆(Long Short-Term Memory, LSTM)网络优化所得,GRU 神经网络是 LSTM 网络的一种效果更佳变体,具有更加简单的结构、更少的参数和更好的收敛性。GRU 神经网络的整体结构见图 1,图 1 中主要包括更新门和重置门,更新门主要用于控制前一时状态信息被代入到当前时刻状态的程度,更新门的值越大说明前一时状态信息代入越多;重置门用于控制忽略前一时状态信息的程度,重置门的值越小说明忽略的越多^[9]。图 1 中: x_t 为当前时刻网络的输入值; h_{t-1} 为前一时 GRU 神经网络的输出值; h_t 为当前时刻 GRU 神经网络的输出值; S 为 Logistic 函数。

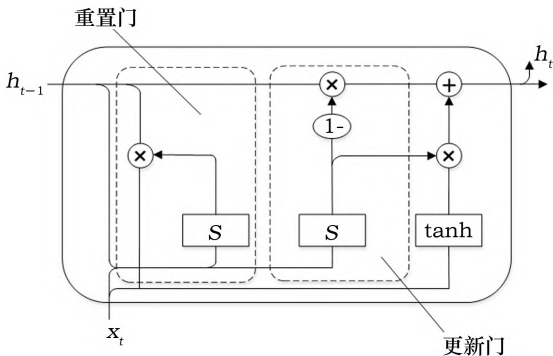


图 1 GRU 神经网络的整体结构

1.3 迁移学习

虽然深度学习被广泛应用在各种领域,但要想得到令人满意的结果仍需要大量的数据。为了解决制作数据集耗时耗力的问题,迁移学习得到越来越多的关注。迁移学习主要实现将某个领域或者任务上学习得到的知识运用到其他不同但相关的领域中,其重要思想就是通过迁移源域的标注数据或者知识结构,完成或改善目标域的学习效果。

2 故障诊断模型

2.1 数据集描述

为了验证本文方法的可行性,采用凯斯西储大学的公开实验数据集^[10]作为实验数据进行训练与测试。将不同的工况数据作为迁移学习的源域及目标域,并且将每种工况的数据分成 4 种状态:正常(N)、内圈故障(IF)、外圈故障(OF)和滚珠故障(BF);将故障部分划分为 3 种尺寸:小尺寸(0.18 mm)、中尺寸(0.36 mm)和大尺寸(0.54 mm)。根据不同轴承状态和故障尺寸,每个域中有 10 种轴承状态,轴承状态分类见表 1,表 1 中 IF18 为滚动轴承发生小尺寸内圈故障;BF36 为滚动轴承发生中尺寸滚珠故障;OF54 为滚动轴承发生大尺寸外圈故障;其他符号以此类推。

表 1 轴承状态分类

故障状态 \ 故障尺寸	轴承状态		
	小尺寸 (0.18 mm)	中尺寸 (0.36 mm)	大尺寸 (0.54 mm)
正常(N)	N	N	N
内圈故障(IF)	IF18	IF36	IF54
滚珠故障(BF)	BF18	BF36	BF54
外圈故障(OF)	OF18	OF36	OF54

由于数据集中提供的原始振动信号是一条很长的一维数据,每种数据样本数较少。为获得尽可能多的样本,采用滑动窗口进行重叠采样;为获得更好的训练效果,将获得的样本随机打乱。每一个域中有 10 000 个样本,每个样本中有 2 048 个数据点,并且每一个域的数据按照 90 % 作为训练集用于训练 CNN-GRU 网络,剩下的 10 % 作为测试集用于测试系统的性能。将负载为 0 HP、1 HP 的 2 种情况分别作为工况 I 和工况 II。工况 I 作为源域,工况 II 作为目标域,各工况具体数据如表 2 所示。

2.2 网络结构确定

本文提出 CNN 和 GRU 相结合的迁移学习网络模型,为了提高网络模型的准确率,利用源域数据进行

表 2 各工况具体数据

域(工况)	转速/(r·min ⁻¹)	样本数	数据点数	轴承状态种类数
I	1 797	10 000	2 048	10
II	1 772	10 000	2 048	10

对比实验,通过比较不同 CNN 层数和 GRU 神经网络层数的组合,选择出最佳的深度学习模型,测试集准确率对比、测试集预测时间对比分别如表 3、表 4 所示。

表 3 测试集准确率对比

GRU 神经网络层数	CNN 层数		
	2	3	4
1	93.8	99.7	99.3
2	98.5	99.8	99.7
3	99.3	99.8	99.8

表 3 为不同层数的 CNN 与 GRU 神经网络组合时,在测试集中的类间准确率。由表 3 可知,当 CNN 层数大于 2 层时,准确率较高,均高于 99 %,这是由于随着网络层数的加深,深层的 CNN 能够提取到较为抽象的特征,更有助于随后连接的全连接层进行分类。表 4 为对 1 000 个测试样本进行分类的预测时间,由表 4 可知,当 GRU 神经网络层数较少时,分类所需的

预测时间较短。按照分类预测时间短,准确率高这 2 个条件对网络层数进行选取,最后选出 3 层 CNN+1 层 GRU 神经网络以及 3 层 CNN+2 层 GRU 神经网络这 2 种网络结构。将 2 种结构进行对比发现,将 GRU 神经网络的层数从 1 层提升至 2 层时,类间平均准确率上升了 0.1 %,但所需时间却成倍增加,因此,综合分析各种情况,选取 3 层 CNN+1 层 GRU 神经网络作为迁移前的网络结构。

表 4 测试集预测时间对比

GRU 神经网络层数	CNN 层数		
	2	3	4
1	2.000 298	2.000 352	3.000 474
2	4.000 586	5.000 754	8.001
3	10.001	14.002	25.004

根据上述对比确立最终模型,初始层(0 层)作为滚动轴承振动信号输入层;1~6 层为卷积层和池化层,用于提取振动信号特征;第 7 层为 GRU 神经网络层,用于提取特征中的时间动态;第 8 层为全连接层,用于连接 GRU 神经网络层与分类层,为数据分类做准备;第 9 层为分类层,用于对信号进行分类,故障诊断模型如图 2 所示,故障诊断网络参数如表 5 所示。

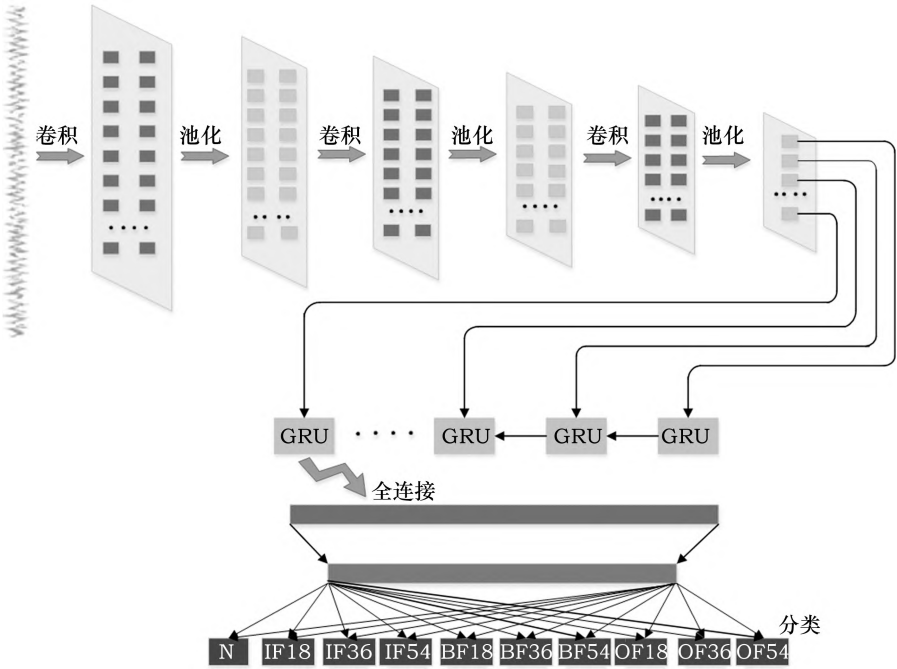


图 2 故障诊断模型

2.3 迁移方式确立

根据上述源域对比实验确定 CNN 和 GRU 神经网络的网络结构。在迁移学习的过程中,将模型中前面

的 CNN 和 GRU 神经网络层的参数冻结,将目标域的故障数据输入到网络中,只对后面网络的全连接层和分类层进行 fine-tune 微调,达到迁移的目的,解冻 CNN

表 5 故障诊断网络参数

层数	网络层	激活函数	输出大小	卷积核大小	滤波器数量	步长
0	输入层	—	2 048×1	—	—	—
1	一维卷积层	ReLU	128×16	64×1	16	16
2	一维池化层	—	64×16	2×1	16	2
3	一维卷积层	ReLU	64×32	3×1	32	1
4	一维池化层	—	32×32	2×1	32	2
5	一维卷积层	ReLU	16×64	3×1	64	1
6	一维池化层	—	16×64	2×1	64	2
7	GRU 神经网络层	—	16×128	—	—	—
8	全连接层	ReLU	3	—	—	—
9	分类层	Softmax	10	—	—	—

和 GRU 神经网络层的参数并和全连接层相连接,确立最终的故障诊断迁移模型,整个迁移过程如图 3 所示。

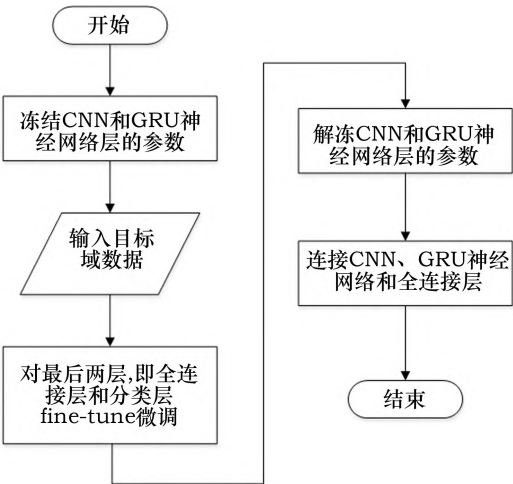


图 3 迁移过程

表 7 中的准确率从图 4 中可得。

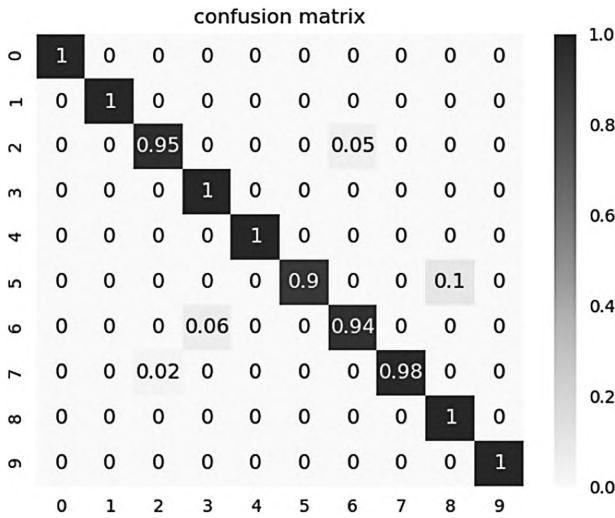


图 4 各类混淆矩阵

3 实验分析

3.1 实验结果

将工况 I 数据集训练所得网络中 CNN 与 GRU 神经网络中的权重参数固定,仅对最后两层,即全连接层与分类层的权重进行 fine-tune 训练。从工况 II 中随机选择 1 000 条数据作为迁移学习的输入数据,其中 70 %作为训练集,20 %作为验证集,10 %作为测试集。实验结果表明,直接将 1 000 条目标域数据与源域数据同时输入网络,在测试集中的类间平均准确率为 87.66 %,使用提出的迁移学习的诊断方式在测试集中的类间平均准确率为 98.04 %。

图 4 所示为各类混淆矩阵,体现出真实值和预测值的验证情况。图 4 中横、纵坐标的数值 0~9 分别表示不同的轴承状态类别,混淆矩阵中轴承状态类别的含义如表 6 所示。迁移学习的实验结果如表 7 所示,

表 6 混淆矩阵中轴承状态类别的含义

故障尺寸 混淆矩阵参数 故障状态	小尺寸 (0.18 mm)	中尺寸 (0.36 mm)	大尺寸 (0.54 mm)
	故障尺寸	故障尺寸	故障尺寸
正常(N)	0	0	0
内圈故障(IF)	1	4	7
滚珠故障(BF)	2	5	8
外圈故障(OF)	3	6	9

表 7 迁移学习的实验结果

故障尺寸 准确率/% 故障状态	小尺寸 (0.18 mm)	中尺寸 (0.36 mm)	大尺寸 (0.54 mm)
	故障尺寸	故障尺寸	故障尺寸
正常(N)	100	100	100
内圈故障(IF)	100	100	98
滚珠故障(BF)	95	90	100
外圈故障(OF)	100	94	100

实验结果表明,本文所提出的故障诊断方法在小样本数据集中,不借助专家经验的情况下,也能够起

到理想的效果。

3.2 对比实验

为了能够进一步突出本文方法的优越性,将本文方法的准确率与文献[11]中采用的基于 BP 神经网络和基于 PNN 方法的准确率,以及文献[12]中采用的基于 LSTM+迁移学习方法的准确率进行对比。由于文献[11]故障分类只有 4 种状态,没有故障尺寸大小分类,所以取本次实验在 4 种状态下故障尺寸大小分类的准确率平均值与基于 BP 神经网络和基于 PNN 方法的准确率进行对比,不同方法的准确率如表 8 所示。

表 8 不同方法的准确率

故障状态	不同方法的准确率/%			
	本文方法	基于 BP 神经网络	基于 PNN	基于 LSTM+迁移学习
正常(N)	100	80	100	100
内圈故障(IF)	99.33	70	90	96.67
外圈故障(OF)	95	75	90	97.78
滚珠故障(BF)	98	70	95	96
平均准确率/%	98.0	73.7	93.7	96.69

根据表 8 的对比可以明显看出,本文提出的基于 GRU 与迁移学习的滚动轴承故障诊断方法无论是单项还是整体上都优于采用基于 BP 神经网络和基于 PNN 方法的故障诊断,而在整体上是优于采用基于 LSTM+迁移学习方法的故障诊断。其主要原因是基于 BP 神经网络和基于 PNN 方法的故障诊断需要大量的训练样本,且随着网络层数的增加,训练样本需求成指数上升的同时也会产生梯度消失的风险,从而导致其结果不太理想;而 GRU 是 LSTM 网络优化所得,所需参数较少,在小样本的条件下训练速度更快,导致基于 LSTM+迁移学习的诊断效果稍逊色。本文采用的迁移学习利用与目标域有相似特征的源域数据进行训练,通过迁移学习只需要小样本的目标域数据,对特定的网络层微调,从而可以获得更高的准确率。

4 结语

针对深度学习的故障诊断方法需要大量训练样本的问题,本文提出一种基于 GRU 与迁移学习的滚动轴承故障诊断方法,通过一维卷积神经网络提取特征,GRU 捕捉时间动态,以迁移学习的方式进行训练。根据实验分析可以证明基于 GRU 与迁移学习的滚动轴承故障诊断方法整体上能够达到 98.0 % 的准确率,说明了本文方法的可行性,为在小样本数据集的情况

下利用深度学习进行故障诊断提出了新的思路。

参考文献:

- [1] WANG W, LEE H. An energy kurtosis demodulation technique for signal denoising and bearing fault detection[J]. Measurement Science and Technology, 2013, 24(2): 025601.
- [2] OSMAN S, WANG W. An enhanced Hilbert-Huang transform technique for bearing condition monitoring[J]. Measurement Science and Technology, 2013, 24(8): 085004.
- [3] 冯玉芳, 卢厚清, 殷宏, 等. 基于 BP 神经网络的故障诊断模型研究[J]. 计算机工程与应用, 2019, 55(6): 24-30.
- [4] WANG D, PETER W T, GUO W, et al. Support vector data description for fusion of multiple health indicators for enhancing gearbox fault diagnosis and prognosis[J]. Measurement Science and Technology, 2010, 22(2): 025102.
- [5] SHEN C, WANG D, KONG F, et al. Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier[J]. Measurement Journal of the International Measurement Confederation, 2013, 46(4): 1551-1564.
- [6] WIDODO A, YANG B S. Support vector machine in machine condition monitoring and fault diagnosis[J]. Mechanical Systems and Signal Processing, 2007, 21(6): 2560-2574.
- [7] LU C, LI T, XU L, et al. Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification[J]. Advanced Engineering Informatics, 2017, 32: 139-151.
- [8] WEN Long, LIU Xinyu, GAO Liang, et al. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.
- [9] 张毛毛. 基于深度双向加权 GRU 的回转窑烧结温度预测[D]. 长沙: 湖南大学, 2018.
- [10] SMITH W, RANDALL R. Rolling element bearing diagnostics using the case western reserve university data: a benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64/65: 100.
- [11] 刘乐, 孙虎儿, 谢志谦. 基于 SVD-LMD 模糊熵与 PNN 的滚动轴承故障诊断[J]. 机械传动, 2017, 41(3): 172-176, 18.
- [12] 修嘉芸, 谷玉海, 任斌, 等. 基于 LSTM 与迁移学习的滚动轴承故障诊断[J]. 重庆理工大学学报(自然科学), 2021, 35(1): 83-88.

作者简介: 曹梦婷, 硕士研究生, 主要研究方向为滚动轴承故障诊断及路径规划算法研究。

E-mail: 13661221867@163.com

收稿日期: 2021-03-12