

基于图卷积网络的电机故障诊断研究

吕栋腾[†], 胡春龙

(陕西国防工业职业技术学院, 陕西 西安 710300)

摘要: 振动信号常被用来分析和判断电机在运行中出现的故障类型。当前故障数据的处理方法主要是通过专家判断分析, 处理成本较高。近年来, 基于深度学习的电机故障数据处理方法已开始用于电机的故障诊断任务。本文提出了一种多头注意力增强的图卷积神经网络深度学习模型, 对输入的故障节点数据进行编码, 通过递归神经网络和非线性层实现电机的故障类型分类, 利用时序故障数据判断电机的故障类型。实验结果表明, 提出的模型在电机故障诊断任务中取得了良好效果。

关键词: 振动信号; 电机; 故障诊断; 图卷积神经网络

中图分类号: TM34

文献标识码: A

Research on Motor Fault Diagnosis Based on Graph Convolutional Network

LV Dongteng^{1†}, HU Chunlong

(Shaanxi Institute of Technology, Xi'an, Shaanxi 710300, China)

Abstract: Vibration signals are often used to analyze and judge the types of faults that occur in the operation of induction motors. At present, the processing method of fault data is mainly through expert judgment and analysis, and the processing cost is high. In recent years, deep learning-based induction motor fault data processing methods have been used in induction motor fault diagnosis tasks. In this paper, a deep learning model of graph convolutional neural network based on multi-head attention enhancement is proposed to determine the fault type of induction motor by using time series fault data. The proposed method firstly obtains the extended time sequence feature representation, encodes the input fault node data based on graph convolutional network, and finally realizes the fault type classification of induction motor through recurrent neural network and nonlinear layer. Experimental results show that the proposed model can achieve the better performance.

Key words: vibration signals; motors; fault diagnosis; graph convolutional network

电机是电力系统中的重要设备, 一旦出现故障或停止运行, 不仅会对电机自身造成损害, 更可能会带来巨大的经济损失, 并有可能导致人员伤亡。因此, 对电机可能的故障进行提前预测和诊断, 具有极其重要的意义^[1]。当前, 针对电机故障诊断的方法可以划分为基于传统诊断的方法和当前流行的基于深度学习的方法两种。基于传统诊断的方

法需要人工从记录的电机运行数据中提取有效的信号特征, 再通过一些浅层机器学习算法, 如支持向量机、随机森林和模糊推理算法等, 从人工构建的特征中学习电机的运行和故障模式, 从而实现电机的故障分类和诊断。在处理特征的过程中, 需要处理者在特征构建时具有大量的专业领域知识, 且时间成本较高。近年来, 随着传感器技术的迅速发

展,能够检测到的电机的运行信号,无论是在精度还是在信号类别方面都呈现出一种海量化、多样化和高速率的趋势。传统人工设计和提取特征的方法将变得十分困难,时间和人力成本也变得难以承受^[2]。近年来,深度学习方法已经开始被用于故障诊断领域中。相较于传统基于浅层机器学习的方法,基于深度学习的电机故障诊断方法能够自动地从海量的电机运行数据中抽取有效的特征信号,通过编码对信号中的有效信息进行提取,最终通过回归或分类的方式有效实现电机故障类型的精确匹配和分类^[3]。

深度学习的兴起,为电机故障诊断研究在传统方法之外提供了一些新的思路和方法。当前,在电机故障诊断中主要应用的是基于卷积神经网络 CNN^[4]和递归神经网络 RNN^[5]的方法。作为经典的深度学习模块,CNN 通过设置不同的卷积核,从不同的视角抽取电机数据内的故障特征。RNN 能够通过递归的方式,在上下文中学到有效的序列信息,因而具有强大的特征提取能力和任务分类能力。例如,在故障检测领域,孙洁娣等提出一种基于卷积神经网络和长短期神经网络 LSTM 等故障检测方法,通过 CNN 自适应提取底层等故障特征,并通过 LSTM 学习时序视角上不同时刻特征之间等依赖关系,在轴承诊断任务中取得来较高等任务性能^[6]。Fatima 等通过基于 CNN 和 LSTM 对电机故障进行分类^[7]。Prashant 等首先对原始信号进行处理,将一维的电机电流信号转化为二维对图像信号作为深度学习对输入,并基于 CNN 和迁移学习设计了一种新的故障诊断框架,从而提高了故行诊断的预测效果^[4]。虽然基于 CNN 的方法已经在电机故障诊断的相关领域中取得了不俗的效果。

上述基于 LSTM 的方法忽略不同信号之间的远距离关联,无法获得更加精细化的时序信号的编码信息,针对上述问题,本文提出了一种基于多头注意力机制的图卷积网络 GCN^[8]的电机故障检测模型,首先通过多头注意力机制将低维度的信号特征映射到不同的映射空间,通过信号合并获得最终的时序节点的增强信号,对于不同时间节点的编码增强信号,通过对于两个时序节点嵌入对计算余弦相似性得分并进行排序,获得其中 Top-K 个有潜在关联的节点对并构建 0/1 权重的邻接矩阵,通过图卷积网络进行时序节点编码,从而获得远距离的电机运行信息,最后通过 LSTM 网络和全连阶层实现时序数据的上下文学习和当前时间节点的电

机运行状态分类,即电机故障类型的分类。最终的实验结果表明,提出的故障检测模型能够实现高精度的故障类型预测,对电机运行监控和预警具有一定的积极意义。

1 电机故障诊断的模型设计

电机故障信号通常是一种单一的时序振动特征信号,人工处理和分析数据代价较高。针对上述问题,本文提出一种基于多头注意力增强的图卷积神经网络 (Graph Convolutional Network Enhanced by Multi-head Attention Mechanism, GCN-MHAM) 方法,用于发掘电机中原始的故障振动信号,提出的模型的整体框架如图 1 所示,主要包括三个模块:(1)多头注意力层:多头注意力机制^[9]被用来将原始的差分故障信号映射到不同的向量空间,经过向量拼接之后将原始的故障特征信号拓展到更高的特征维度;(2)图卷积神经网络层:图卷积网络层通过计算不同时刻拓展特征向量之间的余弦相似度来获得特征向量之间的相似度打分,选取其中 Top-K 的特征向量对,视为图中存在潜在边联系的节点二元组,从而建立起由 0/1 权值构成的邻接矩阵,通过节点之间的聚合操作,而实现时序序列不同距离,特别是远距离节点间的特征编码;(3)有向学习模块:通过引入 LSMT 网络进一步学习序列上的时序特征最终获得一定范围内的综合时序表示;(4)分类器网络:通过使用一个基于 Softmax 的分类激活函数实现故障类型的有效预测。各模块具体的功能和公式如图 1 所示。

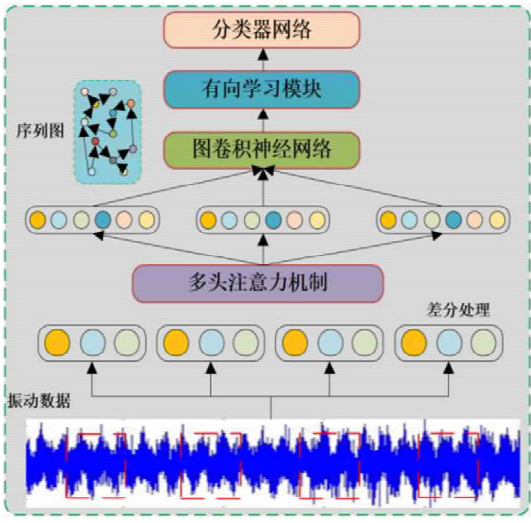


图 1 提出的 GCN-MHAM 模型结构图

由于原始的振动信号在每个时刻都是一系列

的离散信号,为了给时序序列中每个时刻的振动信号构建一个基础的映射向量,本文采用差分处理的方法,经过差分处理后的 t 时刻的信号如公式(1)所示:

$$\mathbf{S}_t = [s_t - s_{t-1}, s_t, s_{t+1} - s_t] \in \mathbb{R}^3 \tag{1}$$

其中, $1 \leq t \leq T-1$, T 表示的是输入模型的一段振动信号的时序长度。在获得 t 时刻的初始特征之后,本文将基于多头注意力机制对上述构建的初始信号特征进行扩展。多头注意力机制可以首先通过三个线性层分别得到对应的 Query、Key 和 Value 向量,随后通过 Query 和 Key 计算注意力权重并与最终的 Value 表示进行相乘,从而得到在某个向量空间,即某个多头空间中对对应的高维特征表示。由于多头注意力机制中上述操作可以在每个头中进行,因而可以得到多个高维特征表示,即多头表示。最终通过多头表示拼接获得对应的高维特征嵌入。因此,对于第 i 个映射头,在经过注意力机制后,获得的映射表示如公式(2)所示:

$$\begin{aligned} \mathbf{Q}_t &= \mathbf{W}_q \cdot \mathbf{S}_t, \mathbf{K}_t = \mathbf{W}_k \cdot \mathbf{S}_t, \mathbf{V}_t = \mathbf{W}_v \cdot \mathbf{S}_t, \\ x_t^i &= \text{softmax}\left(\frac{\mathbf{Q}_t \times \mathbf{K}_t^T}{\sqrt{3}}\right) \mathbf{V}_t \in \mathbb{R}^{d_k}, \end{aligned} \tag{2}$$

其中, $\mathbf{W}_q \in \mathbb{R}^{3 \times d_q}$, $\mathbf{W}_k \in \mathbb{R}^{3 \times d_k}$ 和 $\mathbf{W}_v \in \mathbb{R}^{3 \times d_v}$ 均为映射矩阵,最终多头拼接获得特征向量为 $\bar{\mathbf{x}}_t = [x_t^1, x_t^2, \dots, x_t^n] \in \mathbb{R}^d$, $d = nd_v$ 。

根据上述获得的拼接向量通过计算时序序列上两两节点间的余弦相似度,从而获得节点之间的相似度排序,从而可以建立不同节点之间的潜在关联,取排序分中的 Top- K 个节点对作为构图中的节点二元组,从而在长度为 T 的时序序列上构建由 0/1 权值构成的二元邻接矩阵 $\mathbf{A} \in \mathbb{R}^{T \times T}$ 。基于构建的邻接矩阵,图卷积神经网络可以根据其中蕴含的不同节点间的连接边权重(即 0 和 1),通过简单加和或注意力加权求和的方式将关联邻居节点的表示合并到当前的目标节点中,从而完成节点聚合并实现对结构化数据的节点编码操作。因此,基于对上述图卷积神经网络原理的描述,根据获得的节点表示 $\bar{\mathbf{x}}_T = [\bar{x}_0, \bar{x}_1, \dots, \bar{x}_T] \in \mathbb{R}^{T \times d}$,模型在第 l 层卷积神经网络(GCN)中执行的卷积操作时如公式(3)所示:

$$\mathbf{x}_l^{(g,l)} = \sigma\left[\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \bar{\mathbf{x}}_l^{(g,l-1)} \mathbf{W}_g^{(l)}\right], \tag{3}$$

其中, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, $\tilde{\mathbf{D}}$ 是 $\tilde{\mathbf{A}}$ 的度矩阵, $\mathbf{W}_g^{(l)} \in \mathbb{R}^{d \times d}$ 表示映射向量, $\sigma(\cdot)$ 表示 sigmoid 激活函数。最终通过 L 层 GCN 编层后,模型将获得的编码在有向学习模块中输入一个 LSTM 网络学习序列上的上下文信息,如公式(4)所示:

$$x_{\text{final}} = \overrightarrow{\text{LSTM}}(\mathbf{x}_{-1}^{(g,L)}, \mathbf{x}_t^{(g,L)}), \tag{4}$$

最终将学习的时间序列 T 内的隐藏状态 x_{final} 输入到一个非线性层中进行分类,分类公式如公式(5)所示:

$$\hat{\mathbf{y}} = \text{softmax}(\mathbf{W}_{\text{final}} \cdot \mathbf{x}_{\text{final}} + \mathbf{b}_{\text{final}}) \in \mathbb{R}^8 \tag{5}$$

2 实验及讨论

实验部分使用了一份开源的电机故障诊断数据集^[10]进行实验验证。该数据集中包含了涵盖正常工况在内的 8 种电机运转类型。在实验过程中,本文将按照 20%、10% 和 70% 的比例对故障数据集进行分割得到训练集、开发集和测试集,采用精确率(Accuracy)指标对模型性能进行评估。最终,在训练过程中,模型通过网格搜索,获得的最优实验超参数如表 1 所示。

表 1 设计模型中的超参数设置

名称	超参数值
训练轮数	100
批大小 T	32
学习率	1e-3
GCN 层数 L	2
d_q, d_k 和 d_v 的值	{3,3,5}

此外,在实验讨论中,本文提出的方法还将与一些由递归神经网络(LSTM)、卷积神经网络(CNN)和单头注意力机制(Attention)等方法的组合模型进行比较。在引入的基准模型训练过程中,训练轮数、批大小以及学习律与表 1 保持一致,其余参数设置如表 2 所示。

表 2 比较模型中的超参数设置

名称	超参数值
LSTM 隐藏维度	32
LSTM 层数	2
CNN 多卷积核大小	[1,2,3,5]
CNN 层数	5

基于上述实验配置,本文方法与当前主流的几种神经网络方法组合模型的实验结果比较如表 3 和图 2 所示。

从表 3 中可以看出,基于简单 LSTM 的方法几乎无法从构造的差分特征中学到有效信息,这是因为输入的差分信号属于低维特征,直接使用

LSTM 很难从抽象的时序信号中学到有效特征。在模型中加入 CNN 后,结合注意力集中的权重计算,模型学会自适应的关注输入批数据中的部分关键信息,并结合注意力机制实现性能的显著提升。然而上述方法无法获取不同时间点时序特征间的关联性信息,因此,本文引入了 GCN 网络之后,模型实现了最好的故障识别精度。

此外,本文还验证了 GCN 层数对模型性能的影响,实验结果如图 3 所示。

表 3 不同配置模型的精确率得分

名称	准确率(%)
LSTM	36.54
Attention + LSTM	70.21
CNN+LSTM	85.46
Attention + CNN	92.22
GCN-MHAM	95.77

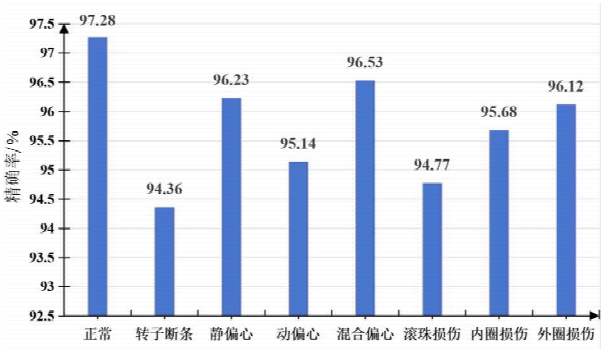


图 2 电机故障类型的精确率分布

由图 3 可得,模型的性能并不会随着 GCN 层数的增强而单调递增,而是由于 GCN 过平滑问题^[11]的存在而呈现出一种先增加后减小的状态,因此,在实验设置过程中,GCN 的层数不宜设置过大,否则会给模型带来性能恶化的风险。

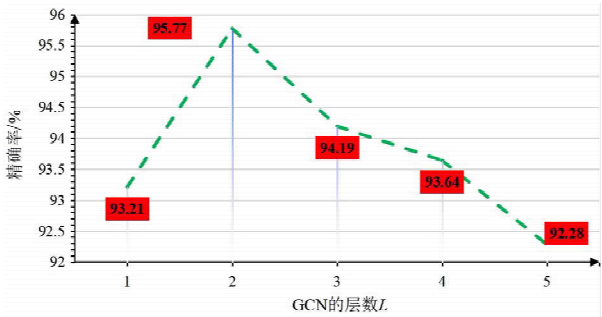


图 3 模型性能随 GCN 层数的变化

3 结 论

提出了一种基于多头注意力增强的图卷积神经网络故障电机分类模型。在对数据进行差分处理之后,使用的多头注意力机制能对低维的特征通过多头映射获得高维的信号特征表示,通过图卷积网络进行构图和节点聚合操作,获得更好的时序特征编码,从而改善了电机故障分类性能。

参考文献

[1] 丁石川,厉雪衣,杭俊,等.深度学习理论及其在电机故障诊断中的研究现状与展望[J].电力系统保护与控制,2020,48(8):172—187.

[2] 孙文珺.基于深度学习模型的电机故障诊断方法研究[D].南京:东南大学,2017.

[3] 孙洁娣,毛新茹,温江涛,等.深度卷积长短期记忆网络的轴承故障诊断[J].机械科学与技术,2021,40(7):1091—1099.

[4] KUMAR P, HATI A S. Transfer learning-based deep CNN model for multiple faults detection in SCIM[J]. Neural Computing and Applications, 2021, 33(22): 15851—15862.

[5] 田旭东,王兵,步跃跃,等.基于深度学习的电机匝间短路故障诊断方法研究[J].机械管理开发,2023,38(5),5—8.

[6] 孙洁娣,毛新茹,温江涛,等.深度卷积长短期记忆网络的轴承故障诊断[J].机械科学与技术,2021,40(7):1091—1099.

[7] HUSARI F, SESHADRINATH J. Sensitive inter-tum fault identification in induction motors using deep learning based methods[C]//2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020). IEEE, 2020: 1—6.

[8] 敖邦乾,曾丽娟,陈孝玉,等.融合注意力机制的电机故障检测系统设计[J].西南大学学报(自然科学版),2022,44(12),163—175.

[9] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30(3),25—30.

[10] 王跃龙.笼型异步电动机多故障智能诊断及分离方法的研究[D].太原:太原理工大学,2017.

[11] LI Q, HAN Z, WU X M. Deeper insights into graph convolutional networks for semi-supervised learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018: 32—36.