

量子-经典混合神经网络及其故障诊断应用

窦 通¹, 周振威^{2†}, 刘 涛³, 汪凯蔚², 汪 皓¹, 崔 巍¹

(1. 华南理工大学 自动化科学与工程学院, 广东 广州 510640;

2. 工业和信息化部电子第五研究所, 广东 广州 511370; 3. 鸿富锦精密电子(烟台)有限公司, 山东 烟台 264000)

摘要: 量子神经网络由于结合了量子计算和神经网络的优点, 近年来受到了广泛的关注. 然而由于目前量子计算资源受限(如量子比特数、量子逻辑门的保真度等)以及贫瘠高原现象(量子神经网络优化过程中解空间变得平坦时出现的训练困难)的存在, 量子神经网络当前还难以大规模训练. 针对上述问题, 本文面向量子-经典混合神经网络模型提出了一种基于无监督学习的特征提取方法. 所采用的无监督学习方法结合了量子自编码器和K-medoids聚类方法, 可用于多层次结构的特征学习. 该方法创新地利用了K-medoids方法对训练得到的量子自编码器进行聚类, 以最大化量子自编码器性质的差异. 进一步, 本文在轴承异常检测问题上, 通过实验验证了所提出的无监督特征提取方法的有效性和实用性, 测试集准确率在二分类、四分类和十分类分别达到100%, 89.6%和81.6%.

关键词: 量子计算; 变分量子算法; 量子自编码器; 无监督学习; 异常检测

引用格式: 窦通, 周振威, 刘涛, 等. 量子-经典混合神经网络及其故障诊断应用. 控制理论与应用, 2021, 38(11): 1785 – 1792

DOI: 10.7641/CTA.2021.10881

Quantum-classical hybrid neural network and its application in fault diagnosis

DOU Tong¹, ZHOU Zhen-wei^{2†}, LIU Tao³, WANG Kai-wei², WANG Hao¹, CUI Wei¹

(1. College of Automation Science and Technology, South China University of Technology, Guangzhou Guangdong 510640, China;

2. The Fifth Institute of Electronics, Ministry of Industry and Information Technology, Guangzhou Guangdong 511370, China;

3. Hongfujin Precision Electronics (Yantai) Limited Company, Yantai Shandong 264000, China)

Abstract: Combining the merits of quantum computation and neural networks, quantum neural networks (QNNs) have gained considerable attention in recent years. However, because of the limitation of quantum resource (the number of qubits, quantum logic gate fidelity, et al.) and the barren plateau phenomenon (the trainability problem that occurs in quantum neural networks when the landscape turns flat as the optimization is run), it is costly to train QNNs. In this paper, an unsupervised feature learning is proposed for quantum-classical hybrid neural networks to alleviate the problems. The unsupervised feature learning method, which can learn a hierarchy of feature extractors, is introduced by combining quantum autoencoders and K-medoids clustering algorithm. Key to our approach is to use K-medoids clustering to maximize the difference of properties of the trained quantum autoencoders. The effectiveness and practicability of the proposed approach is verified on bearing anomaly detection using numerical simulation on binary, four-class and ten-class classification, achieving 100%, 89.6%, 81.6% on test set, respectively.

Key words: quantum computation; variational quantum algorithm; quantum autoencoder; unsupervised learning; anomaly detection

Citation: DOU Tong, ZHOU Zhenwei, LIU Tao, et al. Quantum-classical hybrid neural network and its application in fault diagnosis. *Control Theory & Applications*, 2021, 38(11): 1785 – 1792

1 引言

量子计算是一种基于量子力学规则的新型计算范式, 有望在人工智能、组合优化、量子化学等领域解决

经典计算机难以解决的问题. 理论上, 未来大规模的容错量子计算机可以在质因数分解^[1]、无序结构搜索^[2]、求解线性方程组^[3]等问题上实现加速, 然而噪

收稿日期: 2021-09-18; 录用日期: 2021-11-25.

[†]通信作者. E-mail: zhouzw_ceppei@163.com; Tel.: +86 20-37231385.

本文责任编辑: 赵千川.

国家自然科学基金项目(61801124, 61873317), 电子元器件可靠性物理及其应用技术重点实验室开放基金项目(19D08)资助.

Supported by the National Natural Science Foundation of China (61801124, 61873317) and the Opening Project of Science and Technology on Reliability Physics and Application Technology of Electronic Component Laboratory (19D08).

声和量子比特数的限制阻碍了上述量子算法的实现. 当前量子技术离通用量子计算机要求的保真度、相干时间还有很长的距离. 在未来的一段时间, 人们能使用的量子计算设备只有几十到几百的量子比特, 同时这些量子比特是没有经过纠错的物理量子比特, 而不是逻辑量子比特, 且只能进行相干时间有限的不完美的量子操作. 这些量子计算设备被称为带噪声中等规模量子(noisy intermediate-scale quantum, NISQ)计算设备^[4]. 因此在带噪声中等规模量子计算设备上进行计算必须考虑如下因素:

- 1) 量子比特的数量有限;
- 2) 量子比特之间的连通性有限;
- 3) 限制量子电路深度的相干和非相干错误.

如何最大限度地利用现有的量子资源, 为带噪声中等规模量子计算设备寻找可以发挥其优势的应用问题, 是当前量子计算领域的研究重点. 现阶段, 变分量子算法(variational quantum algorithm, VQA)已经成为在NISQ设备上获得量子优势的主要策略, 因为变分量子算法采用的线路拟设(ansatz)在低深度时可以有效实现. 其旨在借助经典计算机的力量尽可能发挥有噪声中等规模量子计算设备的计算能力去解决具体的问题. 变分量子算法的一部分任务由量子计算设备完成, 然后通过经典计算调整量子计算部分的可调参数, 反复迭代最后输出结果.

基于变分量子算法, 学术界和工业界在机器学习领域进行了许多应用的探索, 并产生了新兴的量子机器学习(quantum machine learning, QML)^[5-8]研究领域. 量子机器学习尝试利用量子计算的优势从而增强机器学习模型的性能, 并致力于回答下面的开放性问题:

- 1) 量子信息如何影响机器学习模型的行为? 如何利用它?
- 2) 量子计算机是否更适合高度相关的数据集(或具有量子相关的经典数据集)?
- 3) 量子机器学习模型能比经典机器学习模型学习得更快吗?
- 4) 如何降低量子机器学习模型的复杂性, 以及是否能从更小的数据样本中学习?

本文拟在当前已实现的量子计算硬件条件下, 研究如何优化特征提取方法从而提升量子神经网络的性能. 进一步, 本文将所研究的混合量子-经典神经网络模型应用到故障数据异构以及是否发生故障未知的故障诊断问题中, 从而进一步验证所提出的混合神经网络的有效性. 该量子-经典混合卷积神经网络(convolutional neural networks, CNN)^[9-13]可以通过局部连接和参数共享学习数据的上下文关联性. 本文异常检测的基本思想是, 将检测问题转化为分类问题, 通过卷积神经网络强大的特征提取能力提取相关特

征, 然后经过分类器进行分类. 文献[14]提出了一种量子-经典混合卷积神经模型, 该模型通过随机量子线路代替经典的卷积操作, 从而构造经典计算难以产生的复杂的映射关系, 并在手写数字识别任务上取得较好的结果. 但由于选取的量子线路具有随机性, 与问题或数据无关; 并且线路的参数固定, 会导致最后结果具有很大的随机性. 因此, 本文在文献[14]的模型基础上, 提出一种基于量子自编码器^[15-16]和K-medoids聚类算法的无监督特征学习方法, 可用于多层次结构的特征学习. 该方法的关键是通过量子自编码器将数据与量子线路的参数关联起来, 并通过K-medoids聚类算法选取在“质点”的量子自编码器作为量子-经典混合卷积神经模型的量子“滤波”器, 然后再将选取好的“量子”滤波器通过卷积形式进行特征提取. 实验说明, 该模型在轴承异常检测问题上具有较好的效果.

本文结构如下: 第2节介绍量子变分算法的基本概念与设计思想; 第3节简单介绍基于量子变分算法的量子自编码器; 第4节介绍本文提出的无监督特征学习方法; 第5节为实验及分析; 最后以第6节作为论文的总结.

2 变分量子算法

本节将简单介绍变分量子算法的基础概念, 使用的主要符号如表1所示. 变分量子算法的核心在于将计算任务转化为优化问题, 由量子计算机和经典计算机协同完成任务. 量子计算机执行困难的部分, 而经典计算机完成相对简单的优化任务. 可以表达为下式:

$$C = \min_{\vec{\theta}} \text{Tr}(\hat{O}U(\vec{\theta})\rho_{\text{in}}U(\vec{\theta})^\dagger) = \min_{\vec{\theta}} f(\vec{\theta}), \quad (1)$$

其中 C 为问题的最优解.

表1 主要符号表

Table 1 Notations

名称	符号
厄密算符	$\hat{O}$
含参数酉变换	$U(\vec{\theta}), V(\vec{\theta})$
不含参数酉变换	W
可调参数	$\vec{\theta}$
量子态	ρ, σ
量子比特数	n
线路层数	L
代价函数	$f(\cdot)$
能量	E
泡利算符	P

变分量子算法典型的算法流程如图1所示, 通常包括状态准备、参数化酉变换、测量和经典优化器. 变分量子算法可分为如下步骤:

- 步骤1** 在量子计算设备上准备量子态 ρ_{in} ;
- 步骤2** 施加由参数 $\vec{\theta}$ 表征的酉变化 $U(\vec{\theta})$;

步骤 3 在厄密算符 $\hat{O}$ 上对量子线路末态进行测量, 求取代价函数 $f(\vec{\theta})$;

步骤 4 经典计算机根据损失函数的值更新参数 $\vec{\theta}$;

步骤 5 重复步骤1-4, 直至收敛.

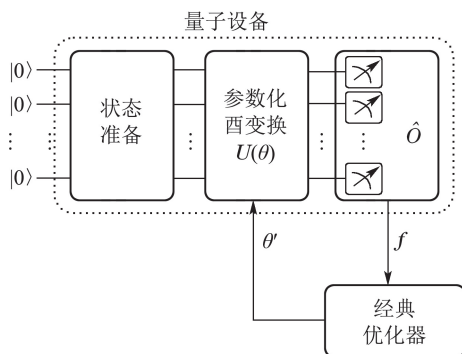


图 1 变分量子算法流程图

Fig. 1 Framework for variational quantum algorithm

值得注意的是, 变分量子算法通常包含一个精心设计的代价函数 f , 使得当 f 取最小值时可以实现计算任务. 由于代价函数实现了从实数(可调参数)到实数(测量结果)的映射, 变分量子算法可以使用经典优化方法来优化可调参数 $\vec{\theta}$. 可见, 该算法结合了量子设备的计算能力与经典设备的优化方法, 故又称作混合量子-经典算法.

实现变分量子算法必须考虑如何设计代价函数将计算任务转化为优化问题, 以及如何构造一个参数化的酉变换(如果算法处理的任务涉及到经典数据, 则需要额外的编码过程将数据转换成量子数据). 此外, 变分量子算法中采用的经典优化方法往往会对算法的效率和准确度产生影响. 特别的, 相比于不基于梯度的优化方法, 解析梯度以及解析高阶导数因为需要运行更少的量子电路来决定优化方向已经成为当前变分量子算法所关注的重点. 本节接下来将简要介绍变分量子算法中代价函数和参数化酉变换两者基本概念.

2.1 代价函数

变分量子算法中的代价函数一般具有如下形式:

$$f(\vec{\theta}) = \sum_j f_j(\text{Tr}(H_j U(\vec{\theta}) \rho_j U(\vec{\theta})^\dagger)) = \sum_j f_j(\langle H_j \rangle), \quad (2)$$

其中: ρ_j 是一组输入量子态; H_j 是一组厄密算符; $\langle H_j \rangle = \text{Tr}(H_j U(\vec{\theta}) \rho_j U(\vec{\theta})^\dagger)$ 表示在线路末端进行 H_j 测量的结果; f_j 是一组经典函数, 表示对测量结果的经典后处理. 由变分量子算法的框架可知, 代价函数的设计必须满足: 当且仅当代价函数达到全局最小时, 计算任务求得解. 这被称为变分量子算法的忠实性(faithfulness)条件^[17].

值得注意的是, 考虑到在带噪声中等规模量子设备上的适用性, 变分量子算法的代价函数必须能够高效地计算. 在实际中, H_j 一般为泡利算符. 此时损失函数可以直接通过对量子态的泡利测量得到. 特别的, 在许多常见的物理模型(如伊辛模型)中, H_j 写成局部泡利算符的线性组合

$$H_j = \sum_k c_{j,k} P_{j,k}, \quad (3)$$

其中 $P_{j,k}$ 是泡利算符, 且只不平凡地作用在局部的量子比特上. 若每一个 $P_{j,k}$ 最多只作用于 m 个量子比特, 则称损失函数为 m -局部的(m -local). 局部损失函数相较于全局损失函数在梯度消失问题^[18]上被认为更具有优势.

2.2 参数化酉变换

在带噪声中等规模量子计算设备实现参数化的酉变换是变分量子算法的一个关键. 参数化量子电路(parameterized quantum circuit, PQC)是一种通用的参数化酉变换实现方式. 在参数化量子电路中, 电路一般采用分层结构:

$$U(\vec{\theta}) = U_L(\vec{\theta}_L) \cdots U_2(\vec{\theta}_2) U_1(\vec{\theta}_1) = V_L(\vec{\theta}_L) W_L \cdots V_2(\vec{\theta}_2) W_2 V_1(\vec{\theta}_1) W_1, \quad (4)$$

式中 L 被称为电路的层数. 由于量子计算设备噪声的存在, 近期设备所能提供的有效的电路宽度(量子比特数 n)和深度(层数 L)将受到限制. 同时, 为使电路能够在近期设备上高效实现, 在电路的每一层 $U_i(\vec{\theta}_i) = V_i(\vec{\theta}_i) W_i$ 中, 一般采用单比特旋转门或含参数的双比特门(ZZ门)来表示参数化的酉门

$$V_i(\vec{\theta}_i) = \prod_k e^{-i\theta_k A_k/2}.$$

而不含参数的酉变换 W_i 作为连接层, 一般由邻近量子比特间的CNOT或CZ双比特门构成, 为电路提供纠缠能力. 参数化电路的结构也常被称为拟设(ansatz). 从PQC的结构逻辑来看, 变分量子算法可以说是经典机器学习方法(如神经网络)的量子扩展, 故参数化量子线路又可以认为是量子神经网络(quantum neural network, QNN)^[19].

在针对特定问题的算法中, 可以通过问题的解的先验信息针对地设计电路拟设(如组合优化问题中的量子交替算符拟设^[20]等). 而在针对一般问题的电路结构设计中, 或当问题的解不存在先验信息时, 电路结构的每一层 $U_i(\vec{\theta}_i)$ 往往采用相同的结构, 即 $U_i(\vec{\theta}_i) = U(\vec{\theta}_i) = V(\vec{\theta}_i) W$. 文献[21]列举了若干含噪声中等规模量子计算机上可实现的量子线路结构拟设.

2.3 挑战

尽管变分量子算法的发展迅速, 但该领域仍有许多挑战需要解决. 这里简单介绍变分量子算法的可训练性(trainability)问题.

变分量子算法的可训练性是指代价函数的“贫瘠高原”(barren plateau)^[22]现象对优化过程造成的困难。这是当前变分量子算法的主要瓶颈。“贫瘠高原”现象指出当变分量子算法所选取的非结构化的拟设电路 $U(\vec{\theta})$ 进行随机参数初始化时,算法对应损失函数 $f(\vec{\theta})$ 的梯度方差 $\text{Var}(\frac{\partial f(\vec{\theta})}{\partial \theta_k})$ 在采样的平均意义下会随着问题规模 n 的扩大呈指数递减。直观来看,这种现象会使得问题的优化曲面变得非常平坦(故称贫瘠),从而使得为了达到确定优化方向的计算精度所需要的测量数会变得无法承担。文献[22]中的结果部分表明了电路表达能力(可以理解为线路遍历所有参数时能表示的酉变换的范围)和可训练性之间的权衡关系。当电路的深度增加时相应的表达能力增强,但同时梯度的方差(可训练性)也会逐渐减少。

3 量子自编码器

经典的机器学习中,自编码器(autoencoder, AE)可以看成一种神经网络,可以学习高维空间中数据有效的低维表示。其任务是给定输入 x ,将 x 映射到低维表示 h ,而 x 可以从 h 中重构。该网络可以分为两部分:一个由映射 $h = e(x)$ 表示的编码器和一个生成重构的解码器,由映射 $r = d(h)$ 表示。图2展示了自编码器的一般结构。自编码器的学习过程可以表述为最小化一个损失函数

$$g(x, d(e(x))), \quad (5)$$

其中 $g(\cdot, \cdot)$ 是一个损失函数,表示惩罚 $d(e(x))$ 与 x 之间的差异。

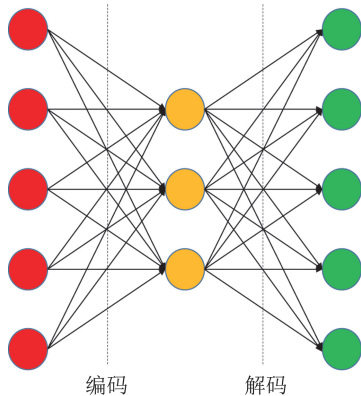


图2 经典自编码器的一般结构,通过内部表示 h 将输入 x 映射到输出(重构) r

Fig. 2 The general structure of a classical autoencoder, mapping an input x to an output (reconstruction) r through an internal representation h

3.1 基于VQA的量子自编码器

基于经典自编码器的思想,图3为一种基于量子变分算法的量子自编码器模型,可以用来执行量子数据的自编码任务。与经典的自编码器模型相似,量子自编码器由一组节点表示输入,每个节点代表一个量子

比特,网络的第1层代表输入寄存器,最后一层代表输出寄存器。对于量子自编码器中的量子线路,输入态包含的部分信息在经过编码部分的量子线路后被丢弃(被测量的量子比特称为“丢弃比特”),然后初始化为一些参考态(通常为 $|0\rangle$ 态),并使用它实现解码操作,并将其与初态进行比较。最后,计算输入态与重构输出态的偏差,本文用两量子态之间的保真度进行度量,如下式所示:

$$F(\rho_i^{\text{in}}, \rho_i^{\text{out}}) = \text{Tr}(\sqrt{\sqrt{\rho_i^{\text{in}}} \rho_i^{\text{out}} \sqrt{\rho_i^{\text{in}}}})^2, \quad (6)$$

其中: ρ_i^{in} 为量子自编码器输入量子态, ρ_i^{out} 为输出的重构量子态。可以看出,当对所有的输入态 ρ_i^{in} 都有

$$F(\rho_i^{\text{in}}, \rho_i^{\text{out}}) \approx 1$$

时,量子自编码器训练成功。

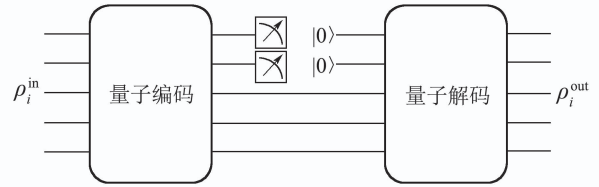


图3 一种量子自编码器的结构

Fig. 3 The structure of a quantum autoencoder

量子自编码器的编码部分可以由参数化量子线路组成,由 $U(\theta)$ 表示。简单地,解码部分可以用 $U(\theta)^\dagger$ 表示。这时候当“丢弃比特”的测量结果等于参考态,则被编码的量子态可以完美地重构为输出态。本文可以采用“丢弃比特”与参考态之间的保真度作为损失函数,这样可以使用基于梯度下降的算法对参数进行优化,并且在优化时忽略解码部分的量子线路。

4 基于无监督学习的特征学习

文献[14]基于卷积神经网络提出了一种量子-经典混合神经网络,主要是利用随机量子线路代替卷积神经网络中的卷积操作,可以把量子线路看作量子“滤波器”。动机是使用随机的非线性特征可以提高机器学习算法的性能和减少训练时间,而量子线路可以映射经典计算机难以高效计算复杂的函数关系。然而因为“贫瘠高原”现象的存在,许多随机挑选的参数化量子线路在很大的参数空间内,其输出态的量子性质是相似的,这会使得随机量子线路产生的非线性特征相近。本节将介绍一种结合量子自编码器和K-medoids聚类方法的量子“滤波器”选取方法,并用于量子特征提取。

总体上看,无监督学习过程可以分为以下3个步骤:

1) 随机选取多个数据块和参数化量子线路的结构,对每个数据块训练一个量子自编码器;

2) 通过K-medoids算法对训练得到的量子自编码器进行聚类;

3) 对训练数据进行预处理并分类.

4.1 参数化量子线路结构的选取

正如第2.2节所提及的, 文献[21]列举了若干能够在含噪声中等规模量子计算设备(NISQ)上实现的量子线路结构(ansatz). 本文将采取表达能力较强的4个量子线路, 即circuit 5, 13, 4, 6进行实验.

4.2 特征学习

选取好量子线路结构的宽度和深度, 初始化好线路的参数之后, 对选取的数据块训练一个量子自编码器. 然后运行编码部分的线路, 得到末态. 可知可以通过量子编码线路的末态表征该量子自编码器, 所以对量子自编码器进行聚类本质上是对量子编码线路的末态进行聚类. 由于需要聚类的量子态众多, 为了方便, 本文对所有量子态在计算基上进行多次测量, 得到了每个量子态的概率分布描述 x^i . 通过对量子态概率分布进行聚类从而实现对量子线路的聚类. 假设有 m 个不同量子线路, 所得样本集为

$$X = \{x^1, x^2, \dots, x^m\},$$

其中每一个样本为以向量表示的概率分布. 若希望将 X 划分为 k 类, 则算法如表2所示.

表 2 用于量子线路的K-medoids聚类算法

Table 2 K-medoids for clustering quantum circuit

Algorithm 1: 使用K-medoids算法对量子线路聚类

Data: 样本集 $X = \{x^1, x^2, \dots, x^m\}$; 聚类数 k

Result: 簇划分 $C = \{c_1, c_2, \dots, c_k\}$

```

1 从 $D$ 中随机选取 $k$ 个样本作为初始“质点” $\{\mu_1, \mu_2, \dots, \mu_k\}$ ;
2 repeat
3   令 $c_i = \emptyset (1 \leq i \leq k)$ ;
4   for  $j = 1, 2, \dots, m$  do
5     计算样本 $x_j$ 与各“质点” $\mu_i (1 \leq i \leq k)$ 的距离:  $d_{ji}$ ;
6     根据距离最近的“质点”确定 $x_j$ 的簇标记:
        $\lambda_j = \arg \min_{i \in \{1, 2, \dots, k\}} d_{ji}$ ;
7     将样本 $x_j$ 划入相应的簇:  $c_{\lambda_j} = c_{\lambda_j} \cup \{x_j\}$ ;
8   end
9   for  $i = 1, 2, \dots, k$  do
10    对每个簇 $c_{\lambda_i}$ 的样本, 计算其与簇内其他样本的距离之和 $d_l^{\lambda_i}, \forall l \in c_{\lambda_i}$ ;
11     $\mu'_i$ 满足 $d_{\mu'_i}^{\lambda_i} = \min d_l^{\lambda_i}, \forall l \in c_{\lambda_i}$ ;
12    if  $\mu'_i \neq \mu_i$  then
13      将当前“质点” $\mu_i$ 更新为 $\mu'_i$ 
14    else
15      保持当前“质点”不变
16    end
17  end
18 until 当前“质点”均未更新;
```

需要注意的是, 由于输入数据本质上是一个概率分布, 距离度量函数 $d(x^j, x^i) = d_{ji}$ 选取两个概率分布之间的JS散度(jensen-shannon divergence, JSD), 如下式:

$$d(x^j, x^i) = \text{JS}(x^j \| x^i) = \frac{1}{2} \text{KL}(x^j \| \frac{x^j + x^i}{2}) + \frac{1}{2} \text{KL}(x^i \| \frac{x^j + x^i}{2}), \quad (7)$$

其中 $\text{KL}(\cdot \| \cdot)$ 为两个概率分布之间的KL散度(kullback-leibler divergence, KLD), 并定义为

$$\text{KL}(P \| Q) = \sum_{i=1} Q_i \log \frac{P_i}{Q_i}, \quad (8)$$

其中 P, Q 为两个概率分布.

得到聚类的结果之后, 选取每个簇类中心的概率分布所代表的量子线路作为一个“量子”滤波器, 这样便可以得到 k 个“量子”滤波器. 可以根据“量子”滤波器的需求选取 k , 为了更好的收敛也可通过手肘法确定 k .

4.3 卷积预处理及分类

通过上述无监督学习, 可以得到 k 个“量子”滤波器, 这样对于输入 $x \in \mathbb{R}^N$ 可以学习到新的表示 $y = g(x) \in \mathbb{R}^M$.

4.3.1 特征提取

对于映射 $g: \mathbb{R}^N \rightarrow \mathbb{R}^M$ 和 $N = h \times w$ 大小的输入数据块, 可以学习得到 $y \in \mathbb{R}^M$ 的表示. 结合类似卷积神经网络对数据进行窗口滑动操作, 对于 $n \times n$ 的输入, 可以得到 $(n - h + 1) \times (n - w + 1) \times k$ 的表示 y (k 个通道). y 的第 (i, j) 个元素为一个 k 维的向量. 更一般地, 可以通过增加步长 s 学习大小为 $(\frac{n-h}{s} + 1) \times (\frac{n-w}{s} + 1) \times k$ 的表示.

整个特征提取的过程可以看成“量子”预处理的过程, 在之后训练分类过程中可以完全不使用“量子”滤波器, 某种程度可以节省现时珍贵的量子计算资源.

4.3.2 分类

通过“量子”预训练之后, 得到了输入数据的特征表示. 如果特征的维度 $(\frac{n-h}{s} + 1) \times (\frac{n-w}{s} + 1) \times k$ 较大, 在进行分类前, 还可以对特征做池化(pooling)或者主成分分析(principal components analysis, PCA)进行降维. 然后再通过全连接网络或者支持向量机进行最后的分类.

5 实验

本节将通过数值实验验证所提方法的可行性, 该实验主要关注轴承的异常检测问题. 主要因为异常检测是一个复杂的问题, 在相当长的一段时间内, 一直是机器学习领域的一个具有挑战性的问题^[23-24]. 而轴承是在机械设备中具有广泛应用的关键部件之一. 由

于过载、疲劳、磨损、腐蚀等原因,轴承在机器操作过程中容易损坏,造成设备剧烈摇晃、设备停机、停止生产、甚至人员伤亡,所以轴承状态的监测和分析非常重要. 如果可以发现轴承的早期弱故障,则防止故障造成损失. 一般来说,早期的轴承弱故障是复杂的,难以检测.

5.1 数据集

本实验使用一个小型的真实轴承振动信号数据集. 该数据集中,轴承有3种故障: 外圈故障、内圈故障、滚珠故障,外加正常的工作状态. 如表3所示,结合轴承的3种直径(直径1、直径2、直径3),轴承的工作状态有10类,见表1所示.

表3 故障类别

Table 3 The fault categories

故障类别	标签
正常	0
外圈故障, 直径1	1
内圈故障, 直径1	2
滚珠故障, 直径1	3
外圈故障, 直径2	4
内圈故障, 直径2	5
滚珠故障, 直径2	6
外圈故障, 直径3	7
内圈故障, 直径3	8
滚珠故障, 直径3	9

数据集中,训练集样本数为629,测试集个数为163. 每一个样本由6000个时间序列的采样点组成,标签使用独热(one-hot)编码.

5.2 结果

本实验训练一层基于量子自编码器的量子“卷积”层. 对于每种量子线路的拟设,需要选择的参数有两个: 线路的宽度,即量子比特的数目 n ; 线路的深度,即拟设的层数 L . 对于线路的量子比特数目,统一选取 $n = 6$. 对于线路的深度,对于每种结构的线路,选取 $L = 2$. 而量子线路参数的选取则从区间 $[0, 2\pi]$ 中随机均匀选取. 总共训练600个量子自编码器.

经过K-medoids聚类之后,得到9个聚类中心,即9个“量子”滤波器. 在“量子”卷积操作中,步长 $s = 32$,每个样本对应的输出数据为9个通道,每个通道数据长度为188,展开之后得到1692维的特征. 通过PCA降维之后,由SVM进行分类,核函数为高斯核(Gaussian kernel)函数. 最终效果如下.

图4为测试集准确率(十分类)与经过PCA后的特征数个数的关系. 在每种特征数情况下,采用十折交叉验证方法,取最好结果. 由图可知,测试集准确率随着PCA后的特征数个数的增大,先上升后下降. 存在

一个合适的特征数(大约在32左右),使得测试集的准确率最高.

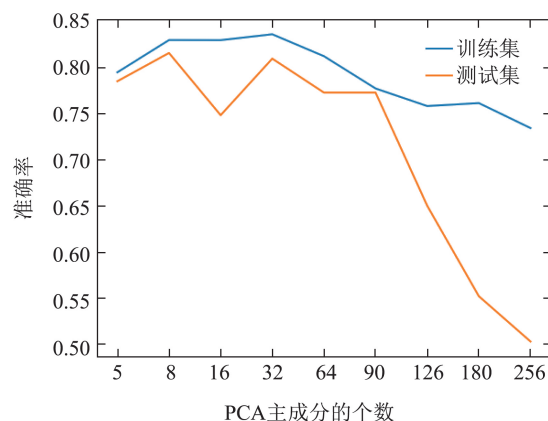


图4 经过PCA后的特征数与测试集准确率的关系

Fig. 4 The relationship between the number of components after PCA and the accuracy on test set

图5为特征数为32时测试集的混淆矩阵. 可以看到,在判断数据是否异常时(二分类),模型在测试集上的准确率为100%;在判断数据为外圈故障、内圈故障、滚珠故障还是正常数据时(四分类),模型在测试集上的准确率为89.6%;而到进一步的具体故障(十分类)时,模型在测试集上的准确率为81.6%.

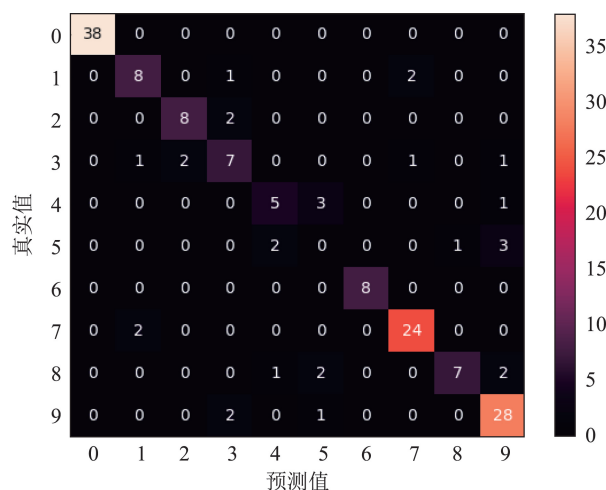


图5 特征数为32时,测试集的混淆矩阵

Fig. 5 The confusion matrix of test set when the number of components is 32

5.2.1 对比

基于同样的数据集,使用PCA和SVM相结合的分类模型作为对比,效果如下.

图6为使用PCA和SVM相结合的分类模型时,测试集准确率(十分类)与经过PCA后的特征数个数的关系. 在每种特征数情况下,采用十折交叉验证方法,取最好结果. 由图可知,当经过PCA后的特征数为128时,测试集的准确率最高,为55.2%.

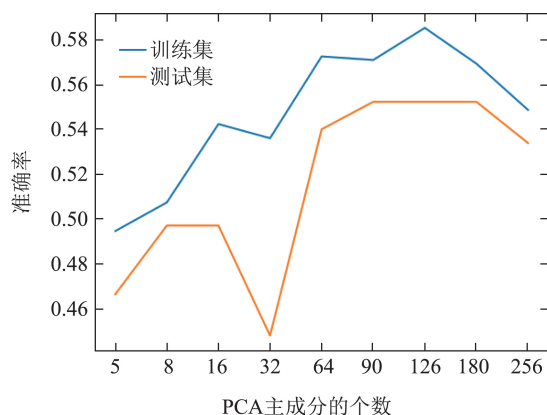


图 6 经过PCA后的特征数与测试集准确率的关系 (PCA+SVM)

Fig. 6 The relationship between the number of components after PCA and the accuracy on test set (PCA+SVM)

图7为使用PCA和SVM相结合的分类模型时,特征数为128时测试集的混淆矩阵。可以看到,在判断数据是否异常时(二分类),模型在测试集上的准确率为99.4%;在判断数据为外圈故障、内圈故障、滚珠故障还是正常数据时(四分类),模型在测试集上的准确率为57.7%;而到进一步的具体故障(十分类)时,模型在测试集上的准确率为55.2%。对比可知,经过无监督特征学习所得量子“滤波器”的特征提取有助于模型对异常检测数据集分类性能的提高。

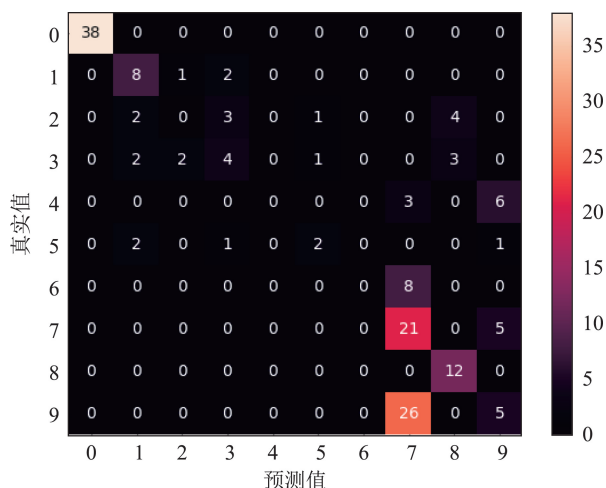


图 7 特征数为128时, 测试集的混淆矩阵(PCA+SVM)

Fig. 7 The confusion matrix of test set when the number of components is 128 (PCA+SVM)

6 总结

基于量子-经典混合卷积神经网络模型, 本文给出了一种无监督学习的特征学习方法。该方法的关键是利用K-medoids算法对训练得到的量子自编码器进行聚类, 将得到的“质点”作为量子-经典混合神经网络的“滤波器”对数据进行特征提取。由于每次只需要训练一个小型的量子自编码器, 可缓解当前量子-经典混合神经网络当前难以大规模训练的问题。进一步,

本文通过轴承异常检测问题验证了该方法的有效性。与PCA和SVM结合的模型相比, 实验结果表明, 基于无监督学习的特征学习方法有助于提高分类模型的性能。需要说明的是, 本文实验只用了一层“量子”卷积层, 具有两层及多层“量子”卷积层的混合卷积神经网络模型的性能有待进一步研究。实际上该过程可以看作一个预训练的过程, 可用于全监督学习精调之前, 从而节约量子计算资源。

参考文献:

- [1] SHOR P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 1999, 41(2): 303 – 332.
- [2] GROVER L K. Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 1997, 79(2): 325 – 328.
- [3] HARROW A W, HASSIDIM A, LLOYD S. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 2009, 103(15): 150502.
- [4] PRESKILL J. Quantum computing in the nisy era and beyond. *Quantum*, 2018, 2(79): 1 – 20.
- [5] BIAMONTE J, WITTEK P, PANCOTTI N, et al. Quantum machine learning. *Nature*, 2017, 549(7671): 195 – 202.
- [6] LU Sicong, ZHENG Yu, WANG Xiaoting, et al. Quantum machine learning. *Control Theory & Applications*, 2017, 34(11): 1429 – 1436. (陆思聪, 郑昱, 王晓霆, 等. 量子机器学习. 控制理论与应用, 2017, 34(11): 1429 – 1436.)
- [7] WU R, CAO X, XIE P, et al. End-to-end quantum machine learning implemented with controlled quantum dynamics. *Physical Review Applied*, 2020, 14(6): 064020.
- [8] LV Yanxuan, CAO Qing, LÜ Jinhui, et al. Recent advances of quantum neural networks on the near term quantum processor. *Scientia Sinica Technologica*, 2021, DOI: 10.1360/SST-2020-0459. (吕颜轩, 高庆, 吕金虎, 等. 面向近期量子处理器的量子神经网络研究进展. 中国科学: 技术科学, 2021, DOI: 10.1360/SST-2020-0459.)
- [9] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, 1(4): 541 – 551.
- [10] CHEN Z Q, LI C, SANCHEZ R V. Gearbox fault identification and classification with convolutional neural networks. *Shock and Vibration*, 2015, DOI: 10.1155/2015/390134.
- [11] BABU G S, ZHAO P L, LI X L. Deep convolutional neural network based regression approach for estimation of remaining useful life. *International Conference on Database Systems for Advanced Applications*. Cham: Springer, 2016, 9642: 214 – 228.
- [12] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration*, 2016, 377: 331 – 345.
- [13] JEONG H, PARK S, WOO S, et al. Rotating machinery diagnostics using deep learning on orbit plot images. *Procedia Manufacturing*, 2016, 5: 1107 – 1118.
- [14] HENDERSON M, SHAKYA S, PRADHAN S, et al. Quantum machine learning: powering image recognition with quantum circuits. *Quantum Machine Intelligence*, 2020, 2(1): 1 – 9.
- [15] ROMERO J, OLSON J P, GUZIK A A. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2017, 2(4): 045001.
- [16] PRIETO C B. Quantum autoencoders with enhanced data encoding. *Machine Learning: Science and Technology*, 2021, 2(3): 035028.

- [17] CEREZO M, SHARMA K, ARRASMITH A, et al. *Variational quantum state eigensolver*. arXiv preprint arXiv: 2004.01372, 2020.
- [18] CEREZO M, SONE A, VOLKOFF T, et al. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 2021, 12(1): 1 – 12.
- [19] FARHI E, NEVEN H. *Classification with quantum neural networks on near term processors*. arXiv preprint arXiv: 1802.06002, 2018.
- [20] HADFIELD S, WANG Z, O’GORMAN B, et al. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms*, 2019, 12(2): 34.
- [21] SIM S, JOHNSON P D, GUZIK A A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2019, 2(12): 1900070.
- [22] MCCLEAN J R, BOIXO S, SMELYANSKIY V N, et al. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 2018, 9(1): 1 – 6.
- [23] PANG G, SHEN C, CAO L, et al. Deep learning for anomaly detection: a review. *ACM Computing Surveys (CSUR)*, 2021, 54(2): 1 – 38.
- [24] CHALAPATHY R, CHAWLA S. *Deep learning for anomaly detection: A survey*. arXiv preprint arXiv:1901.03407, 2019.

作者简介:

窦 通 硕士研究生, 目前研究方向为量子神经网络, E-mail: 893767081@qq.com;

周振威 高级工程师, 目前研究方向为装备故障预测与健康管理及可靠性技术研究, E-mail: zhouzw_ceppei@163.com;

刘 涛 助理工程师, 目前研究方向为机械装备与管理技术, E-mail: liutao_fsk@163.com;

汪凯蔚 高级工程师, 目前研究方向为装备可靠性及健康管理技术, E-mail: wangkaiwei82@163.com;

汪 皓 硕士研究生, 目前研究方向为量子机器学习, E-mail: 202020116491@mail.scut.edu.cn;

崔 巍 教授, 目前研究方向为量子人工智能算法、区块链数据分析, E-mail: aucuiwei@scut.edu.cn.