计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
62(4): 821−851, 2025

基于变分量子电路的量子机器学习算法综述

于瑞祺　张鑫云　任　爽
（北京交通大学计算机科学与技术学院　北京　100044）
（yuruiqi@bjtu.edu.cn）

A Review of Quantum Machine Learning Algorithms Based on Variational Quantum Circuit

Yu Ruiqi, Zhang Xinyun, and Ren Shuang
（*School of Computer Science & Technology*, *Beijing Jiaotong University*, *Beijing* 100044）

Abstract　As the scale of available data increases, the importance and impact of machine learning grows. It has been found that quantum computing can be realized with the help of the principles of quantum mechanics, and the quantum machine learning algorithm formed by combining quantum computing and machine learning can theoretically produce exponential acceleration advantages over classical machine learning algorithms. Quantum versions of many classical algorithms have been proposed and they may solve problems that are difficult to classical computers. At present, limited by the quantum computing hardware, the number of controllable qubits, noise, and other factors restrict the development of quantum computers. Quantum computing hardware is unlikely to reach the level needed for universal quantum computers in the short term, and current research focuses on the algorithms that can run on noisy intermediate-scale quantum (NISQ) computers. Variational quantum algorithms (VQAs) are hybrid quantum classical algorithms which are suitable for current quantum computing devices. Related research is one of the research hotspots in the field of quantum machine learning. Variational quantum circuits (VQCs) are parameterized quantum circuits (PQCs) used in variational quantum algorithms to solve quantum machine learning tasks. It is also be called Ansatz and quantum neural networks (QNNs). The framework of variational quantum algorithm mainly contains five steps: 1) Designing the loss function according to the task. Designing parameterized quantum circuits as model and initializing parameters. 2) Embedding classical data. The classical data is pre-processed and encoded to the quantum state. If quantum data is used as input, it only needs to be pre-processed without encoding. 3) Calculating the loss function through parameterized quantum circuit. This step is where quantum advantage comes in. 4) Measuring and post-processing. Through quantum measurement operation, the quantum superposition state wave packet collapses into classical state. The classical data can be obtained after post-processing. 5) Optimizing the parameters. Updating parameters and optimizing the model with classical optimization algorithms and then returning to step 3 until the loss function converges after several iterations. We can obtain a set of optimal parameters. The final result is the output of the optimal model. This paper reviews the basic theory of quantum computing and the basic framework of variational quantum algorithm, and further introduces the application and progress of variational quantum algorithm in the field of quantum machine learning, then reviews supervised quantum machine learning including quantum classifiers, unsupervised quantum machine learning including quantum circuit born machine, variational quantum Boltzmann

收稿日期：2023-12-05；**修回日期**：2024-11-19
基金项目：国家自然科学基金项目（62072025）
This work was supported by the National Natural Science Foundation of China (62072025).
通信作者：任爽（sren@bjtu.edu.cn）

machine and quantum autoencoder, semi-supervised quantum learning including quantum generative adversarial network, quantum reinforcement learning, and quantum circuit architecture search in detail. Next, this paper compares the models and analyses their advantages and disadvantages, and briefly discusses and summarizes the related datasets and simulation platforms that can reproduce the introduced models. Finally, this paper puts forward the challenges and future research trends of quantum machine learning algorithms based on variational quantum circuit.

Key words quantum computing; quantum machine learning; variational quantum algorithm; quantum neural network; quantum deep learning; quantum reinforcement learning

摘 要 随着数据规模的增加,机器学习的重要性与影响力随之增大.借助量子力学的原理能够实现量子计算,结合量子计算和机器学习形成的量子机器学习算法对经典机器学习算法理论上能够产生指数级的加速优势.部分经典算法的量子版本已经被提出,有望解决使用经典计算机难以解决的问题.当前受量子计算硬件所限,可操控的量子比特数目和噪声等因素制约着量子计算机的发展.短期内量子计算硬件难以达到通用量子计算机需要的程度,当前研究重点是获得能够在中等规模含噪声量子(noisy intermediate-scale quantum, NISQ)计算设备上运行的算法.变分量子算法是一种混合量子-经典算法,适合应用于当前量子计算设备,是量子机器学习领域的研究热点之一.变分量子电路是一种参数化量子电路,变分量子算法利用其完成量子机器学习任务.变分量子电路也被称为拟设或量子神经网络.变分量子算法框架主要由5个步骤组成:1)根据任务设计损失函数和量子电路结构;2)将经典数据预处理后编码到量子态上,量子数据可以省略编码;3)计算损失函数;4)测量和后处理;5)优化器优化参数.在此背景下,综述了量子计算基础理论与变分量子算法的基础框架,详细介绍了变分量子算法在量子机器学习领域的应用及进展,分别对量子有监督学习、量子无监督学习、量子半监督学习、量子强化学习以及量子电路结构搜索相关模型进行了介绍与对比,对相关数据集及相关模拟平台进行了简要介绍和汇总,最后提出了基于变分量子电路量子机器学习算法所面临的挑战及今后的研究趋势.

关键词 量子计算;量子机器学习;变分量子算法;量子神经网络;量子深度学习;量子强化学习

中图法分类号 TP18

DOI: 10.7544/issn1000-1239.202330979 **CSTR:** 32373.14.issn1000-1239.202330979

近年来,机器学习已经成为人们日常生活中重要的部分之一,许多应用是基于机器学习算法实现的,数据集的质量是影响机器学习算法性能的决定性因素之一.由于全球数据规模的迅速增长,大规模数据的存储和处理成为一个重要的问题[1].随着数据集规模的增大,数据集可用性也在随之增强,机器学习的重要性也在增大,但经典机器学习算法所需要的训练时间也随之激增.特别是在计算复杂性较高的问题上,对于大规模数据经典算法无法在可接受的时间内解决问题,量子机器学习(quantum machine learning, QML)在其中展现出长足的优势.最初,Google公司完成了一项量子超越传统实验[2],在对伪随机量子电路输出的采样任务上,最先进的Summit超级计算机上需要约10 000年才能完成,而使用53个量子比特(qubit)的NISQ设备仅需200 s即可完成.这项实验标志着NISQ设备存在超越经典计算机的潜力.然而,经典模拟技术的发展也在对量子霸权提出挑战.Huang等人[3]提出一种基于张量网络的经典模拟算法,能够在不到20天的时间内使用Summit完成53量子比特20周期的随机量子电路采样任务.Pan等人[4]提出了精确计算200万比特串的精确幅度的方法,获得比特串所需计算复杂度远低于Google公司使用的方法,获得振幅的计算复杂度也低于最先进的张量网络方法.在实验中仅使用60个GPU,完成53量子比特和20个周期的Sycamore电路采样任务的经典模拟只用了5天.因此,量子计算的优势是需要在长期发展中不断对比经典模拟和量子硬件才能得以证明的.为了保持量子优势,构建更加复杂、更加能够体现量子优势的任务并实现,也需要支持更多量子比特、具有更高保真度的量子处理器.目前最能体现量子优越性的实验是由中国科学技术大学Zhu等人[5]提出,实现了一种66量子比特2维超导量子处理器,命名“祖冲之2.1”,具有97.74%的保真度.在该处理器上设置了60量子比特、20周期的随机电路采样任务,模拟该采样任务所需的经典计算成本大约是Sycamore处理器上所能完成的最难任务的

500 倍. “祖冲之 2.1”使用了 4.2 h 来完成该任务，而使用目前最先进的经典模拟算法在超算 Summit 上完成该任务大约需要 48 000 年，量子优势的表现得到了显著增强.

量子机器学习是综合量子计算(quantum computing, QC)、量子信息(quantum information, QI)和机器学习形成的一门学科，主要研究如何利用量子计算机的特性完成机器学习任务，从而获得比经典算法更快的速度，这被称为量子优势. 当前量子机器学习方向可以分为 3 类：第 1 类是在量子实验环境下实现机器学习算法，利用量子特性改进经典算法，抑或是寻找可以在量子计算机上实现的更有效的新方法，以期提高计算效率、降低计算复杂度，这类研究被称为量子计算学习(quantum computational learning, QCL)；第 2 类是通过量子计算过程启发得到新经典机器学习算法；第 3 类是将经典机器学习算法应用在量子实验中，包括物理化学模拟等，在物理研究领域应用较多. 本文综述的基于变分量子电路的量子机器学习算法属于量子计算学习.

由于量子硬件的限制，可以操控的量子比特数目较少以及量子纠缠和退相干造成的噪声是当前量子设备所面临的 2 个重大问题，物理学在原理的探究上进展缓慢，大规模量子比特、容错量子计算机在未来或许能够实现，但是在短期内进展缓慢. 当前多数研究集中于如何利用现有的 NISQ 设备来实现量子优势，根据是否使用 NISQ 设备可以将量子机器学习算法分为第 1 代量子机器学习算法[6-7]和第 2 代量子机器学习算法[8-16].

变分量子算法(variational quantum algorithms, VQAs)是为应用于 NISQ 设备而设计的算法，已经成为获取量子优势的主要手段之一. 变分量子算法是一种混合量子-经典算法，将任务拆分为 2 部分分别交给 NISQ 设备和经典计算机运行，以期在分配给 NISQ 设备的任务部分利用量子设备并行计算的特性获得量子加速. 变分量子算法已经被用于解决多种任务，包括变分量子本征求解器(variational quantum state eigensolver, VQE)[8]、用于组合优化问题求解的量子近似优化算法(quantum approximate optimization algorithm, QAOA)[9,7-19]、分子模拟[20]等，本文主要综述其在机器学习领域的应用.

在机器学习的领域，变分量子算法可以根据使用设备及实现方法简单分为 2 类：基于量子门电路的模型和基于绝热量子计算的模型，前者使用量子电路计算机，而后者使用量子退火机. 基于门电路的模型使用变分量子电路(variational quantum circuits, VQCs)作为学习任务的主体，数据被编码到量子比特上，经过一系列具有可学习参数的量子门离散化地操作量子比特以完成学习任务，最后通过测量操作得到结果，因此也称参数化量子电路(parameterized quantum circuits, PQCs)，其具有深度浅、使用量子比特较少的特点. 经典计算机执行优化算法更新其参数. 这种学习方式类似于经典机器学习中的神经网络，故参数化量子电路也称量子神经网络(quantum neural networks, QNNs). 基于门的模型衍生出了基于脉冲的模型，使用含可训练参数的脉冲电路作为完成机器学习任务的电路主体. 基于绝热量子计算的算法源于绝热定理，通过将问题编码到初始哈密顿量基态上，依赖系统的绝热演化得到最终的哈密顿量，在演化结束后对系统测量得到问题的解. 相较而言，基于门和脉冲的模型更接近经典范式，是一种更通用的框架，能够提供更高的精度，解决问题更全面，但更容易受量子退相干影响，对纠错的要求更严格. 基于绝热量子计算的模型优势在于电路简单，可以减少量子退相干的问题，无需严格的纠错方法，适合于解决特定的组合优化问题，在机器学习任务中仅被用于生成模型中的量子玻尔兹曼机和量子自编码器. 在硬件设备的发展中，基于门电路的量子计算机以 IBM、微软、谷歌等公司为主，量子退火机则以 D-Wave 公司为主.

本文重点综述基于门模型的变分量子算法在机器学习领域的应用，对于基于绝热量子计算和基于脉冲的模型仅对其开创性工作和近期进展作简要介绍. 表 1 总结了本文所讨论的模型及其可以用于实现的设备.

Table 1 Device Used for Quantum Machine Learning Model

表 1 用于量子机器学习模型的设备

模型	门电路	量子退火机	脉冲电路
分类器	√	×	√
量子卷积神经网络	√	×	×
影子电路	√	×	×
量子玻恩机	√	×	×
量子玻尔兹曼机	√	√	×
量子自编码器	√	√	×
量子生成对抗网络	√	×	×
量子强化学习	√	×	×
量子电路结构搜索	√	×	√

注：“√”表示可以实现的设备，“×”表示不能实现的设备.

本文的主要贡献有3个：

1)针对变分量子算法应用和量子机器学习算法交叉领域，对在变分量子算法框架下实现的量子机器学习算法进行综述.

2)综述了变分量子算法框架下的量子强化学习算法以及量子神经架构搜索算法. 以此分类，按时间顺序对相关算法进行了总结，并对同类量子机器学习算法进行了对比分析.

3)对所总结的量子机器学习算法相关的经典优化算法、数据集以及模拟实验平台进行了总结与对比分析.

本综述介绍了变分量子算法相关的量子计算基础概念，介绍了实验数据集及模拟实验平台，详细介绍了变分量子算法的原理和基于变分量子电路实现的量子机器学习算法的原理和研究进展，总结了其在未来的挑战和展望.

1 量子计算

量子计算是一种借助量子力学进行计算的方式. 在研究量子力学的过程中，人们发现量子系统具有模拟线性代数运算的功能，同时证明了量子算法在某些具体问题上远远优于经典算法，甚至可以实现指数级的加速. 量子计算有着广泛的应用，有望用于基础科学研究、化工、能源、材料、人工智能、信息安全、加密通信等领域，中短期有望在量子优化、量子模拟、量子增强人工智能等方向出现突破性进展.

1.1 量子比特

经典计算机操作的经典比特(bit)只有0和1这2种状态. 在量子计算机中，量子计算机操作的基本单元被称作量子比特或量子位. 与经典比特相对应，量子比特存在$|0\rangle$和$|1\rangle$共2种基态，与经典比特的区别是量子计算机中比特的状态可以是基态的叠加态. 叠加态也称量子态，用$|\psi\rangle$表示，是以2个基态为基底的2维复空间上的向量，故可以表示为基态的线性组合：

$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle. \tag{1}$$

对于n位量子比特构成的量子态可以表示为2^n个基态的复合态：

$$|\psi\rangle=\underbrace{\alpha_{0\cdots00}}_{n位}\underbrace{|0\cdots00\rangle}_{n位}+\underbrace{\alpha_{0\cdots01}}_{n位}\underbrace{|0\cdots01\rangle}_{n位}+\cdots+\underbrace{\alpha_{1\cdots11}}_{n位}\underbrace{|1\cdots11\rangle}_{n位}. \tag{2}$$

式(2)也可以写为

$$|\psi\rangle=\sum_{i=0}^{2^n-1}\alpha_i|i\rangle, \tag{3}$$

其中$|i\rangle$表示n位量子比特构成的基态，α_i表示该基态对应的概率幅. 量子态$|\psi\rangle$被测量时塌缩到基态$|i\rangle$的概率为$|\alpha_i|^2$. 故概率幅满足归一化条件：

$$\sum_{i=0}^{2^n-1}\alpha_i=1. \tag{4}$$

1.2 量子门

在经典计算机中，使用或、与、非等逻辑门能够对经典比特的状态进行变换. 相应地，在量子计算机中使用量子门可以操作量子比特的状态. 对量子比特进行逻辑运算可以视为乘一个$2^n\times2^n$的矩阵. 量子门可以根据是否还有可学习的参数分为基本量子逻辑门和含参量子逻辑门，后者是构成参数化量子电路和实现变分量子算法的关键.

1.2.1 基本量子逻辑门

基本量子逻辑门也被称为固定门. 对于固定门，根据其操作的量子比特数目可以简单分为2类：单量子比特门与多量子比特门，多量子比特门中常用的包括双量子比特门与三量子比特门.

单量子比特门中，最基础的门是单位门，经过其变换的量子比特不发生变化，单位门常用于和其他门组合使用. 常用的单量子比特门主要包括Pauli-X门、Pauli-Y门、Pauli-Z门，其作用是对量子比特的概率振幅或相位进行翻转，在布洛赫球上表现为分别绕对应轴旋转π的角度. S门和T门与Pauli-Z门类似，区别在于绕Z轴的旋转度数分别为$\frac{\pi}{2}$和$\frac{\pi}{4}$. 常见的还包括Hadamard门，其作用为制备叠加态，例如作用于状态为0的量子比特$|0\rangle$，可以得到概率均等的$|0\rangle$和$|1\rangle$.

多量子比特门包括2种类型的量子比特：控制比特和目标比特. 故按照是否含有控制比特可以将其分类为带控制量子门与不带控制量子门. 带控制量子门根据控制比特的状态决定对目标比特的不同操作. 常用的带控制的双量子比特门包括受控非门：***CNOT***门(也称***CX***门)、***CY***门、***CZ***门、***CS***门等，其触发控制时作用与相应的单量子比特门一致. 交换门(***SWAP***)是常用的不带控制量子门，用于交换量子比特的状态. 常用的三量子比特门包括双控制式三量子非门：Toffoli门(***CCNOT***)，以及单控制式量子交换门：Fredkin门(***CSWAP***)等.

量子门作用于量子态等同于对量子比特进行一个酉变换，因此量子门与酉算子是对应的，每个量子门可以用一个酉矩阵表示，满足$\boldsymbol{U}\boldsymbol{U}^{\dagger}=\boldsymbol{U}^{\dagger}\boldsymbol{U}=\boldsymbol{I}$. 其中$\boldsymbol{U}$表示酉算子，对应1个酉矩阵，$\boldsymbol{U}^{\dagger}$是$\boldsymbol{U}$对应的复共轭转置. 酉变换属于可逆变换，量子门对量子比特的

变换是可逆的.

以 Hardmard 门为例，将其作用于基本量子态 $|0\rangle$ 可以表示为 $\boldsymbol{H}|0\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$，将其作用于量子态 $|1\rangle$ 可以表示为 $\boldsymbol{H}|1\rangle=\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$，也可以作用于量子叠加态，如 $\boldsymbol{H}\left[\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\right]=|1\rangle$.

表 2 中列举了一些常见的基本量子逻辑门及其符号和酉矩阵表示.

Table 2 Frequently Used Basic Quantum Gates

表 2 常用基本量子逻辑门

量子逻辑门	符号表示	酉矩阵表示
单位门	$\boldsymbol{I}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$
Pauli-X 门	$\boldsymbol{X}$	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$
Pauli-Y 门	$\boldsymbol{Y}$	$\begin{pmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{pmatrix}$
Pauli-Z 门	$\boldsymbol{Z}$	$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$
Hadmard 门	$\boldsymbol{H}$	$\frac{\sqrt{2}}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$
Phase 门	$\boldsymbol{S}$	$\begin{pmatrix}1&0\\0&\mathrm{i}\end{pmatrix}$
交换门	***SWAP***	$\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}$
受控非门	***CNOT***，***CX***	$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}$
受控 Y 门	***CY***	$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&-\mathrm{i}\\0&0&\mathrm{i}&0\end{pmatrix}$
受控 Z 门	***CZ***	$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}$
Toffoli 门	***Toffoli***，***CCNOT***	$\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\\0&0&0&1&0&0&1&0\end{pmatrix}$
Fredkin 门	***Fredkin***，***CSWAP***	$\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&0&0\\0&0&0&1&0&0&0&1\end{pmatrix}$

1.2.2 含参量子逻辑门

变分量子电路与第 1 代量子电路的不同之处在于其使用了参数化量子电路，参数化量子电路的门具有可训练的参数，这些门被称为旋转门或可调门. 其中最重要的门主要包括 RX, RY, RZ. 相对于固定门中的 Pauli-X, Pauli-Y, Pauli-Z, 这 3 个可调门可操作量子比特沿对应轴逆时针旋转，该门中可训练的参数 θ 并非 π 旋转门，可通过固定门组合构成，其矩阵表示如下：

$$\boldsymbol{R}_{\boldsymbol{X}}(\theta)=\cos\left(\frac{1}{2}\theta\right)\boldsymbol{I}-\mathrm{i}\sin\left(\frac{1}{2}\theta\right)\boldsymbol{X}=\begin{pmatrix}\cos\left(\frac{1}{2}\theta\right)&-\mathrm{i}\sin\left(\frac{1}{2}\theta\right)\\-\mathrm{i}\sin\left(\frac{1}{2}\theta\right)&\cos\left(\frac{1}{2}\theta\right)\end{pmatrix},\tag{5}$$

$$\boldsymbol{R}_{\boldsymbol{Y}}(\theta)=\cos\left(\frac{1}{2}\theta\right)\boldsymbol{I}-\mathrm{i}\sin\left(\frac{1}{2}\theta\right)\boldsymbol{Y}=\begin{pmatrix}\cos\left(\frac{1}{2}\theta\right)&-\sin\left(\frac{1}{2}\theta\right)\\\sin\left(\frac{1}{2}\theta\right)&\cos\left(\frac{1}{2}\theta\right)\end{pmatrix},\tag{6}$$

$$\boldsymbol{R}_{\boldsymbol{Z}}(\theta)=\cos\left(\frac{1}{2}\theta\right)\boldsymbol{I}-\mathrm{i}\sin\left(\frac{1}{2}\theta\right)\boldsymbol{Z}=\begin{pmatrix}\mathrm{e}^{-\frac{1}{2}\mathrm{i}\theta}&0\\0&\mathrm{e}^{\frac{1}{2}\mathrm{i}\theta}\end{pmatrix}.\tag{7}$$

在模型中旋转门通常以组合的方式出现，而任意作用于单量子比特的酉变换都可以被分解为绕任意 2 个相互垂直的轴进行旋转，例如可以分解为绕 Y 轴和 Z 轴旋转：$\boldsymbol{R}(\alpha,\beta,\gamma)=\boldsymbol{R}_{\boldsymbol{Z}}(\gamma)\boldsymbol{R}_{\boldsymbol{Y}}(\beta)\boldsymbol{R}_{\boldsymbol{Z}}(\alpha)$，其中 $\boldsymbol{\omega}=(\alpha,\beta,\gamma)$ 这组参数是可学习的.

2 变分量子算法

变分量子算法作为混合量子-经典算法，核心是使用经典优化算法训练参数化量子电路. Bang 等人[21]描述了一种由经典计算机控制量子器件实现的统一操作的方法. 变分量子算法所提供框架能够用于解决各种量子机器学习问题. 算法流程如图 1 所示.

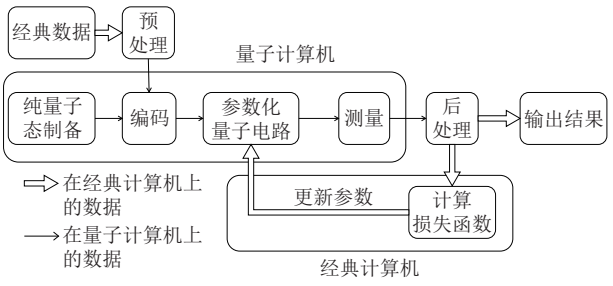


Fig. 1 Illustration of variational quantum algorithms

图 1 变分量子算法示意图

1)根据希望解决的任务设计损失函数，使用一组参数化量子电路作为模型并对其参数初始化；

2)将经典数据进行预处理后编码到量子态上，

若使用量子数据作为输入则仅需预处理，无需编码；

3)利用参数化量子电路进行损失函数的计算；

4)通过量子测量操作使量子叠加态波包塌缩为经典态，经过后处理得到经典数据；

5)使用经典优化算法更新参数和优化模型，回到3)，直到经过多次迭代后收敛，得到一组在损失函数上最优的模型参数.

可以将变分量子算法与深度学习算法类比，将参数化量子电路类比于神经网络，参数化量子电路中旋转门能够变化的相位即为其可训练参数，多个相位参数构成模型参数. 使用经典优化器对模型进行训练，更新这些相位参数. 基于变分量子电路的算法在机器学习任务上良好的表现得益于其与神经网络结构上的相似性.

变分量子算法的3个核心部分分别是损失函数、参数化量子电路以及经典优化器. 下面围绕3个核心部分展开介绍.

2.1 损失函数

损失函数根据所需解决的机器学习任务进行设计，用于描述解决的问题，损失函数的设计直接影响算法的性能. 损失函数一般通过量子电路的测量期望值表示：

$$C(\boldsymbol{\theta})=\sum_k f_k(\mathrm{tr}[\boldsymbol{O}_k\boldsymbol{U}(\boldsymbol{\theta})\boldsymbol{\rho}_k\boldsymbol{U}^{\dagger}(\boldsymbol{\theta})]), \tag{8}$$

其中$\boldsymbol{\rho}_k$表示训练集中的一组输入量子态，$\boldsymbol{U}(\boldsymbol{\theta})$表示含参的酉变换，是参数化量子电路对输入量子态的操控，$\boldsymbol{O}_k$表示一组易测量的可观察值，f_k表示编码函数.

Cerezo等人[22]提出损失函数必须满足3个条件：

1)损失函数最小值对应问题的解，并且较小损失函数值对应着较好的解；

2)损失函数的计算必须通过量子计算机的测量以及经典的后处理得到，否则不具有量子优势；

3)损失函数的参数必须是可训练的.

保真度用于度量量子系统中2个量子态之间的相似程度，保真度越高表示2个量子态之间的相似程度越高. 在量子机器学习任务中，通常基于量子态之间的保真度设计损失函数，具体量子态则需要结合具体任务需要和电路设计进行选择.

2.2 参数化量子电路

在用于构造机器学习模型的参数化量子电路中，一般的框架如图2所示.

首先从数据集中获得数据向量，对其进行经典预处理，得到用于编码的向量$\phi(\boldsymbol{x})$，它决定了编码电路的参数. 对于经过编码后的量子态数据，使用参数化量子电路作为模型进行运算，该部分被称为拟设电路(Ansatz)，在设计机器学习模型时重点考虑Ansatz的结构. 最后测量各个观察量的期望值，经典后处理的功能是对之前设计的损失函数进行相应的构造. 其中预处理、后处理都是可以被训练的.

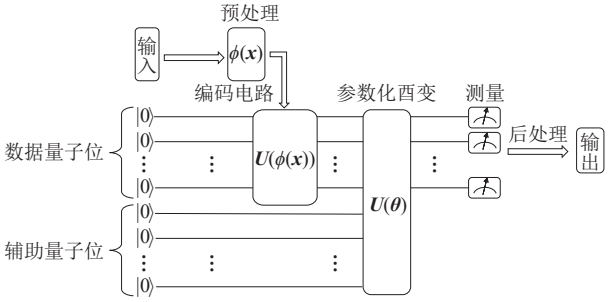


Fig. 2 Parameterized quantum circuits frame

图2 参数化量子电路框架

2.2.1 编 码

对于经典数据，编码电路可以看作是一个非线性特征映射F，它将经典数据$\boldsymbol{x}$映射到具有n位量子比特的量子态上，数学上表示为将数据向量映射到高维希尔伯特空间. 而该特征映射可以通过参数化量子电路实现，经典数据决定其参数，如式(9)所示.

$$|\psi(\boldsymbol{x})\rangle=F(\phi(\boldsymbol{x}))|0\rangle^{\otimes n}, \tag{9}$$

其中$|\psi(\boldsymbol{x})\rangle$表示编码到$n$位量子比特上的初态，作为电路模型的输入态.$\phi(\boldsymbol{x})$是预处理函数，$F(\boldsymbol{x})$是编码函数. 已经提出了一些复杂的特征映射[23-24]，多种编码方式[25]包括概率幅编码[26]、量子抽样编码[26-27]、哈密顿量编码[28]等已被提出.

2.2.2 拟设电路

Ansatz设计目的是在控制参数数量和深度的前提下，设计量子门组合结构使电路变换尽可能地近似目标函数.

Ansatz的选择极大程度影响着变分量子算法的收敛速度、最终解与最优解的近似程度等性能，同时电路深度所带来的噪声影响、门的构建成本等问题也对Ansatz造成一定的限制. 目前Ansatz主要分为2种：硬件高效电路和问题启发电路. 量子机器学习研究以使用硬件高效电路为主.

1)硬件高效电路. 这种电路最初由Kandala等人[29]提出，目的是在量子位较少等硬件设备限制的情况下尽可能地高效完成任务. 硬件高效电路交替使用纠缠门和单量子旋转层，具有损失函数不依赖于所解决的问题、在硬件上对纠缠的保真度要求不高、使用的门在电路上较好实现等特点. 图3中所示的交错分层形式的电路一般称为多层参数化量子线路

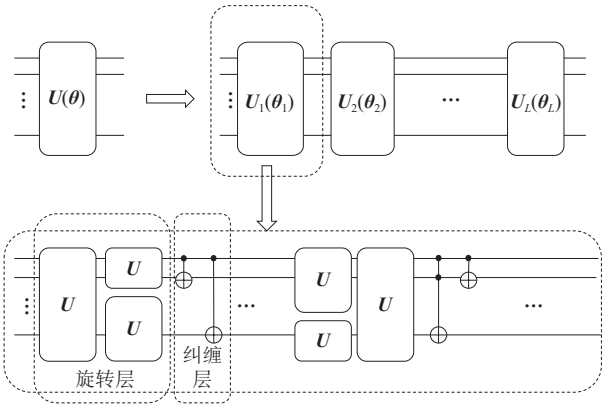


Fig. 3 Multilayer parameterized quantum circuits

图 3 多层参数化量子电路

(multiple parameterized quantum circuit, MPQC).

Ansatz 是一个作用于输入 n 位量子比特上的参数化酉变换:

$$|\psi(\boldsymbol{\theta})\rangle = \boldsymbol{U}(\boldsymbol{\theta})|\psi_0\rangle, \tag{10}$$

其中$|\psi_0\rangle$为输入的量子态, $|\psi(\boldsymbol{\theta})\rangle$为参数化量子电路输出的量子态. 该酉变换可以由多个酉变的乘积构成:

$$\boldsymbol{U}(\boldsymbol{\theta}) = \boldsymbol{U}_1(\boldsymbol{\theta}_1)\boldsymbol{U}_2(\boldsymbol{\theta}_2)\cdots\boldsymbol{U}_L(\boldsymbol{\theta}_L), \tag{11}$$

其中 L 称为该参数化量子电路的层数或深度. 每个酉变由多个基本量子门和含参量子门组成. 如图 3 所示.

还有一些 Ansatz 结合张量网络进行设计, 能够有效地表示高维空间中的数据, 矩阵乘积态(matrix product state, MPS)[30-31]、树型张量网络(tree tensor network, TTN)[10,32-33]、多尺度纠缠重整化 Ansatz(multiscale entanglement renormalization Ansatz, MERA)[10,32]、量子卷积神经网络(QCNN)[11]、影子电路[34]等模型的量子版本已经被广泛地应用于分类和生成模型之中.

2)问题启发电路. 这种电路的设计对具体问题进行特定的设计, 最著名的包括用于求解化学基态的 UCC Ansatz[20], ADAPT-VQE[35] 等.

一般情况下, 量子硬件设备产生的噪声影响会随着量子电路的深度增加而逐渐增大, 变分量子电路的优势在于在保持良好精度的前提下, 能够使用较浅的电路完成相应任务, 具有应用于 NISQ 设备的理想条件. 需要注意的是, 在设计电路时, 可训练的参数不能过多也不能过少, 以避免出现过拟合和偏差过大的情况. 如何在保持较浅的深度下设计能够有效完成任务的模型是目前变分量子算法设计的关键.

2.3 优化算法

在变分量子算法中, 优化部分是使用经典计算机完成的, 许多经典优化算法都可以用于该部分. 优化过程的收敛速度是决定算法整体性能的重要因素之一, 对于大规模数据训练过程, 经典优化部分容易成为限制算法的瓶颈, 加上量子硬件的限制, 其产生的噪声也会对优化算法产生一定的影响, 因此如何选择和设计优化算法是一项非常重要的研究. Bittel 等人[36]证明了训练参数化量子电路是一个 NP-Hard 问题, 即使用经典计算机进行优化是 NP-Hard 的, 经典优化部分对变分量子算法造成的限制使得其更需要寻找性能更好的优化方法.

优化算法根据通用与否分为 2 类. 部分经典优化算法可以用于训练大多数参数化量子电路, 部分优化算法针对特定电路结构设计, 能够在特定的量子机器学习任务上获得优势. 表 3 对经典优化算法及量

Table 3 Classical Optimization Algorithms and Usage Preference of Models

表 3 经典优化算法及模型使用偏好

经典优化器	基于梯度	有监督模型	无监督模型	半监督模型	强化学习模型	电路结构搜索
Adam[37]	是	TTNs[10,32-33] QCCNN[38-39] VSQL[34]	QCBM[31,40-42], QBM[14] QAE[15,43]	QGAN[19,44-50]	QDQN[16,51] QActor-critic[52] QMARL[53-54]	QuantumNAS[12] QAS[55] MQNE[56]
(mini-batch)SGD[57]	是	HNN[58]		QGAN[59]	QDDQN[60]	
AMSGRAD[61]	是			QGAN[50,62]		
BFGS/L-BFGS-B[63-64]	是		QCBM[40] QAE[65-66]			
Nesterov moment[67]	是	QCNN[39]				
RMSProp[68]	是	MPS-VQC[30]			VQ-DQN[69]	
基于梯度优化的其他模型	是	QCNN[11,70-71]	QCBM[41] QBM[72]	QGAN[73-74]		
PSO[75]	否		QCBM[13,76]			
SPSA[77]	否			QGAN[62]		CRLQAS[78]
CMA-ES[79]	否		QCBM[40,80]			
GA[81]	否		QCBM[82] QAE[83]			

子机器学习模型使用偏好进行了统计与对比.

2.3.1 经典优化算法

一般情况下，可以根据优化算法是否使用了梯度下降的方法将算法分为 2 类，即基于梯度的优化算法和基于非梯度的优化算法.

参数平移法是用于计算变分量子电路测量期望值对 Ansatz 各参数的梯度的方法，通过 2 个参数发生偏移的相似的量子电路来获得梯度，Mitarai 等人[84]将这种求解梯度的方法引入到了机器学习领域中的有监督学习任务之中. 损失函数$C(\boldsymbol{\theta})$对参数θ_i的梯度可以表示为

$$\frac{\partial C}{\partial \theta_i}=C\left(\theta_i+\frac{\pi}{4}\right)-C\left(\theta_i-\frac{\pi}{4}\right). \tag{12}$$

对参数的梯度值的估计可以通过同一电路不同参数的运算结果进行计算，不需要额外的量子比特和量子电路，适合在 NISQ 计算机上使用. 参数平移法对于每个可学习的参数进行梯度计算都需要计算 2 次损失函数，并且每次计算参数梯度都要调整电路，存在参数量过大以及电路深度增加时，计算和调整次数过多、运行时间过长的缺陷. He[85]提出了一种仅使用 1 个量子电路计算梯度的方法，能够有效地减小电路深度和缩短编译时间.

Li 等人[86]提出了一种对参数化量子电路求解偏导数的解析方法，能够通过厄米算子的期望值直接得到损失函数的梯度. 这使得基于梯度下降的经典优化算法可以被用于变分量子算法的优化部分，经典机器学习任务中常用的 Adam 优化等[37]同样在量子机器学习中被广泛使用.

如表 3 所示，大部分变分量子算法偏好采用 Adam 作为优化器. Adam 算法是一种在经典机器学习中极为常用的优化器，是一种将动量法和 RMSProp[68]相结合的自适应学习率优化算法，综合考虑了梯度一阶矩估计与二阶矩估计，每次迭代中通过梯度的指数加权平均m_t与梯度平方的指数加权平均s_t进行参数更新，首先计算m_t和s_t：

$$\begin{aligned} m_t&=\beta_1 m_{t-1}+(1-\beta_1)g_t,\\ s_t&=\beta_2 s_{t-1}+(1-\beta_2)g_t^2, \end{aligned} \tag{13}$$

其中g_t表示第 t 步迭代的梯度，β_1，β_2是 2 个取值为$(0, 1)$范围内的衰减率超参数. m_0，s_0初始化为 0，迭代初始化为 0 会导致偏差较大，需要进行偏差修正：

$$\begin{aligned} \hat{m}_t&=\frac{m_t}{1-\beta_1^t},\\ \hat{s}_t&=\frac{s_t}{1-\beta_2^t}. \end{aligned} \tag{14}$$

最后进行参数更新：

$$\theta_t=\theta_{t-1}-\frac{\alpha}{\sqrt{\hat{s}_t}+\varepsilon}\hat{m}_t, \tag{15}$$

α为学习率相关超参数，算法建议取 α=0.001，ε是 1 个大于 0 的实数，一般$\varepsilon=10^{-8}$.

除了 Adam 算法，随机梯度下降(SGD)[57]、Adam 算法改进版 AMSGRAD[61]、拟牛顿法(包括 BFGS, L-BFGS, L-BFGS-B 等[63-64])、Nesterov 动量算法[67]也可以作为经典优化器.

对于一些参数化量子电路，使用基于梯度下降的优化算法会面临很多问题. 例如在有噪声的实验环境中损失函数不光滑，不适合使用基于梯度下降的优化算法；还有可能损失函数的某些部分是无法确定的，无法计算梯度，同样无法使用基于梯度下降的优化算法. 对于这些情况，可以使用非梯度优化方法.

SPSA 算法采用估计目标函数梯度的方法得到近似最优解. 每次使用 2 个目标函数估计值近似梯度，梯度的运算与优化参数的维度无关，从而极大程度上减少了对目标函数梯度的计算量. 对于计算梯度复杂度较高的变分量子电路，SPSA 是变分量子算法的一种有效优化算法，具有更快的收敛速度.

除此之外，粒子群算法(PSO)[75]、协方差矩阵自适应进化策略(CMA-ES)[79]、遗传算法(GA)[81]等无梯度优化算法也可以作为经典优化器.

2.3.2 针对性优化算法

由于参数化量子电路的特殊性，可以针对参数化量子电路模型的结构设计合适的优化算法.

对于损失函数是关于算子期望值的函数问题，Nakanishi 等人[87]提出了量子序列最小优化方法. 利用参数化量子电路的特征结构，通过将损失函数分解为三角函数的和并将函数与一组参数进行匹配，损失函数每次迭代选择一组参数进行优化，直接将选定的参数更新至最优，在所有参数上顺序执行这种局部更新，在所有参数上进行迭代，直至收敛. 这种方法对统计误差具有更强的鲁棒性，收敛速度快，在给定参数序列的条件下无需超参数等特点. 对于这种方法，Parrish 等人[88]提出通过 Anderson 加速算法来加速收敛.

Grimsley 等人[35]提出一种迭代地添加参数化量子门的方法来优化电路的结构，并在电路结构变化后使用梯度下降的优化算法重新优化电路，这种方法在 VQE 中比传统梯度下降法性能提升更大，在量子机器学习任务中也可能通过类似的方法提高算法的性能. Ostaszewski 等人[89]提出一种同时优化参数化量子电路的结构和损失函数的优化算法，这种优

化算法同样可能被迁移到量子机器学习任务中.

2.4 复杂度

量子计算在计算复杂度上展现出相对经典算法更大的优势，最著名的是 Shor[90] 的算法，能够在多项式时间内解决大数质因子分解问题. 衡量量子算法相对于经典算法的优势一般从 5 个复杂度进行探讨：查询复杂度、交互计算复杂度、时间复杂度、样本复杂度、电路复杂度. 目前对于量子核方法的复杂度研究比较深入[91]，而对于量子神经网络的复杂度研究尚浅，有研究表明，预测性能与参数量紧密相关，用少量数据训练就可能获得较好的泛化性[92]. 在量子机器学习领域，我们一般更关注量子算法的查询复杂性、时间复杂性、样本复杂性以及电路复杂性. 而对于基于变分量子电路的模型，我们更关注时间复杂性与电路复杂性.

时间复杂度关注于在计算过程中的时间开销. Shor[90] 的算法是在时间复杂性上展现出优势的最佳示例，而量子计算相对于经典计算优势的严格证明尚未出现，这与著名计算复杂性问题“P=NP”相关. 对经典数据上的机器学习任务，目前尚未出现已知的指数级量子优势[93]，但设计出多项式级优势的量子算法是比较有希望的. 而对于量子数据在时间复杂度上的优势比较明显，原因是用经典方法去模拟量子过程是比较困难的. 然而，数据的可用性也会使经典机器学习算法的计算能力得到提高，例如对于几何区域哈密顿量的基态性质预测任务，最坏情况下不存在指数级的量子优势[94].

样本复杂度经常与时间复杂度共同讨论，指数级的样本复杂度往往意味着指数级的时间复杂度，而反过来却不一定. 在某些场景下使用量子存储器或量子计算机存取或处理量子信息时，存在可以被严格证明的指数级样本复杂度，即不存在可以使其不存在指数级优势的改进版经典算法[95-96].

变分量子电路更注重电路复杂性，一般情况下指的是量子门复杂度，即完成任务所需要的最少量子门的个数. 能体现出较明显优势的是 Wei 等人[71] 提出的量子卷积神经网络，在分类任务中保持准确率的前提下对经典卷积神经网络在门复杂度上获得了指数级的加速，同时也减少了所需的参数量. 此外也存在电路深度复杂性的概念，关注完成任务所需的电路层数. 电路深度复杂性上的量子优势有实验证明：Bravyi 等人[97] 提出一种二次型计算问题，能够使用常数级的量子电路解决，而不能使用常数级经典电路解决. 对于变分量子电路，为了缓解噪声的影响，寻找能够完成任务的最浅电路是重要的研究方向之一，这也是电路结构搜索算法出现的需求之一.

3 数据集及模拟实验简介

3.1 数据集介绍

在量子机器学习实验中，使用经典数据集比较多，使用量子数据集较少，当前设计易制备的量子数据集是一个重要的任务. 下面对几个常见的量子机器学习中用到的数据集进行介绍.

对于分类任务，经典数据集中最常用的是 MNIST 数据集，常常被用于二分类和多分类任务. 该数据集由 55 000 个训练样本构成，每个样本是 28×28 像素的手写数字图像，其标签分别为 0~9 中的其中之一. 一般要对数据集进行降维和下采样处理[98] 或是扩展处理[71]，从而满足电路的输入需求. 许多经典数据集都被用于模拟实验或真实量子实验，例如 Iris 数据集、Fashion-MNIST 数据集、Tetri 数据集、CIFAR10 数据集、Cancer 数据集等. 量子态分类（quantum state discrimination, QSD）任务[99-100] 是一个基本问题，能够对量子态尤其是含噪声量子态进行分类的量子分类器才拥有应用 NISQ 设备的可能. 一般情况下，模型的量子态分类测试使用有限个互不正交的量子态[24,34,56].

对于生成模型，经典任务集中于对概率分布的学习，包括单峰分布、多峰分布、金融数据分布等. 数据集以使用 MNIST 数据集和 BAS（bars-and-stripes）数据集为主. 对于量子数据的学习同样不存在完善统一的数据集，多是对于某量子态进行学习，制备与目标量子态尽可能相似的量子态. 文献 [73] 分别学习了具有 1, 2, 4, 8 个量子比特的纯量子态和 1, 2, 3 个量子比特的混合态. 文献 [101] 学习了量子分布，即测量量子态的结果所构成的分布，这类似于学习影响一个量子态的参数，通过固定参数的酉门制备和采样量子态数据，这在其提出的 CVBM 模型中实际学习了制备量子态的酉门的参数，该文献中所使用的参数化量子态由文献 [102] 提出. 在文献 [56] 中则使用受对称性保护的拓扑相（symmetry protected topological, SPT）完成分类任务. 相对而言，量子数据集学习困难程度由低到高为纯量子态、由量子电路制备的混合量子态、具有复杂结构的物质拓扑相，同时也随量子比特数量增加而增大. 电路制备的纯量子态和混合量子态数据一般用于生成模型[101-102]，而拓扑相数据一般用于量子态分类器和电路结构搜索算法评估[56].

对于强化学习模型,一般都是在经典环境中进行交互,OpenAI Gym 中的环境基本都可以用于量子强化学习算法的测试.最常用的 2 个环境分别是冰湖(frozen-lake)和倒立摆(cart pole),这些环境比较简单,其优势在于易于使用小量子比特的电路实现算法,适合在量子模拟平台上进行模拟测试.

表 4 中统计与对比了量子机器学习任务常用数据集.

Table 4 Frequently Used Datasets of Quantum Machine Learning Tasks

表 4 量子机器学习任务常用数据集

任务	数据集/交互环境
有监督学习	MNIST[10,30,32-33,38,70-71,98]、Iris [10,32,58]、BAS[58]、量子数据[32]
无监督学习	MNIST[103]、BAS[13,15,31,40,42,72,76]、金融数据[41,80]、药物数据 QM9[43]、生成概率分布[15,40,82,101]、量子数据[101]
半监督学习	MNIST[19,49-59]、BAS[19,44-45]、QM9[47,104]、生成概率分布数据[46,62]、量子数据[73]
电路结构搜索	MNIST[12,55-56]、量子数据[56]
强化学习	frozen-lake[69,105]、cart pole[16,51-52,105-106]

3.2 模拟实验介绍

在量子机器学习中,由于 NISQ 量子计算机数量较少,同时又具有可操控量子比特数少和噪声随电路深度增加等特点,当前提出的多数基于变分量子电路实现的量子机器学习算法都是在模拟平台上完成的仿真实验.模拟平台使用经典计算机模拟量子算法的结果.表 5 对一些常见的模拟平台进行了总结.

Table 5 Frequently Used Simulation Platforms and Models

表 5 常用模拟平台及模型

模拟平台	机构	模型	语言
Qiskit	IBM	HNN[58]、QGAN[50]、QBM[72]	Python
TFQ	Google	TTNs[32]、MERAs[32]、VQTN[10]、QCNN[70]、QGAN[107]、QRL[51]	Python
Pennylane	Xanadu	QTN[33]、QCNN[39]、QCBM[31]、QGAN[47]、QVAE[43]、VQ−DQN[69]、QRL[52]、QAS[108]	Python
Torchquantum	MIT	QuantumNAS[12]、QAE[15]、QMARL[53]	Python
Yao	QuantumBFS	MQNE[56]、QCBM[40]、QGAN[44]	Julia
Paddle Quantum	百度	QCL[84]、VSQL[34]、QAE[65]、QGAN[74]	Python
VQNet	本源量子	VQM[109]、QCNN[11,110]、VSQL[34]、QAE[65]、QGAN[50]、VQ-DQN[69]	C++

4 量子机器学习应用

基于变分量子电路的量子机器学习算法适用于 NISQ 设备,当前一般被用于完成以下机器学习任务:对于有监督学习任务,研究一般关注分类任务,以及以多种拟设结构为主体的分类模型;对于无监督学习任务,研究一般关注生成模型,包括量子玻恩机、量子玻尔兹曼机、量子自编码器;对于半监督学习任务,研究关注量子生成对抗网络;部分强化学习模型.一些用于寻找最优拟设结构的量子神经架构搜索算法也被提出.图 4 按时间顺序对各类基于变分量子电路的量子机器学习算法进行了整理汇总.

4.1 有监督学习

在经典机器学习中,有监督学习以分类和回归任务为主.在量子机器学习领域,有监督学习任务中以解决分类任务为主,因此多种不同结构的量子分类器已经被设计出来.

分类问题可以描述为:给定训练数据 $D=\{x_i, y_i\}_{i=1}^{N}=\{X,Y\}$,其中 N 是样本数量,x_i 是训练数据样本,y_i 是与之对应的样本标签.分类任务的目的是训练一个分类器,使其对样本的类别预测尽可能与样本标签接近,一般用如下损失函数表示:

$$\boldsymbol{\theta}^{*}=\arg\min_{\boldsymbol{\theta}}\frac{1}{N}\sum_{i=1}^{N}L(f(x_i,\boldsymbol{\theta}),y_i)+R(\boldsymbol{\theta}), \tag{16}$$

其中 f 是从 X 到 Y 的映射函数,$\boldsymbol{\theta}$ 是可训练的参数,L 是对预测误差的度量函数,R 是正则化函数.在优化过程中不断更新 $\boldsymbol{\theta}$,找到使损失函数最小的参数值,此时达到理论上的最优分类情况.

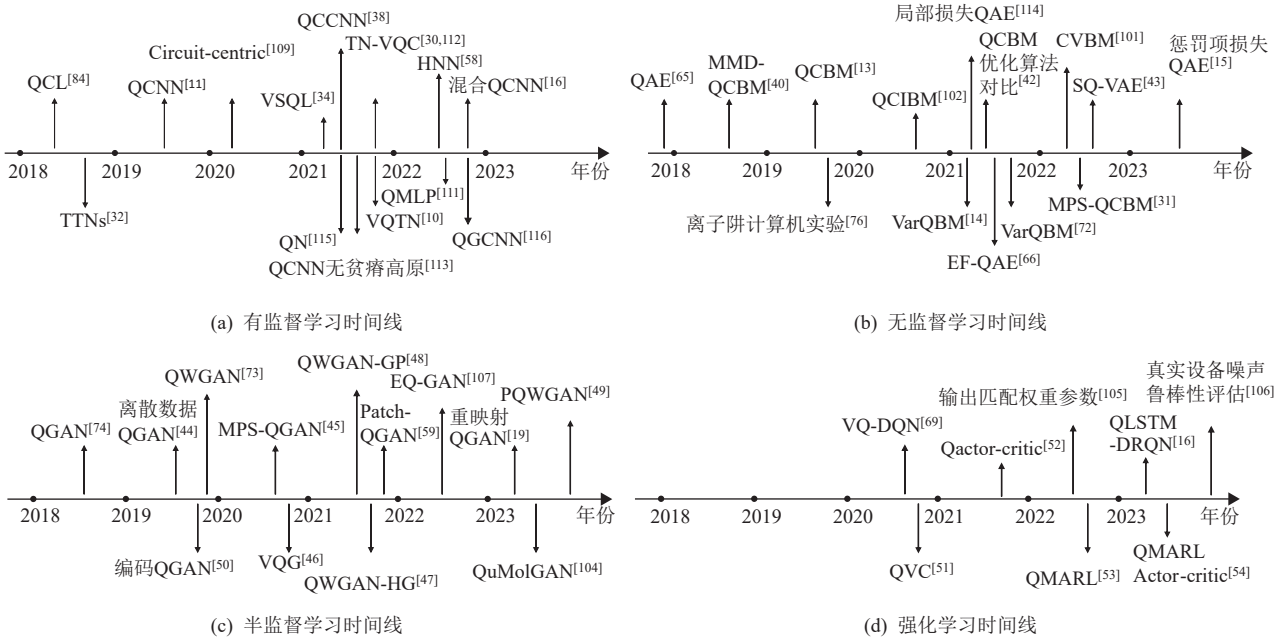


Fig. 4 Timelines of quantum machine learning algorithms based on variational quantum circuit

图 4 基于变分量子电路的量子机器学习算法时间线

表 6 对各分类任务上基于变分量子电路的机器学习算法及其性能进行了总结与对比.

4.1.1 变分量子分类器

使用变分量子电路作为主体直接完成分类任务的分类器一般被称为变分量子模型(VQM). 分类任务的损失函数一般定义为数据真实标签与一个易于测量的可观察量的期望值之间的距离, 观察值一般使用厄米哈密顿量来计算:

$$L=\sum_i(\langle\psi_i|\boldsymbol{U}^{\dagger}\boldsymbol{H}\boldsymbol{U}|\psi_i\rangle-y_i)^2, \tag{17}$$

其中 $\boldsymbol{H}$ 为系统的厄米哈密顿量矩阵, $\langle\psi_i|$和$|\psi_i\rangle$表示经过编码后的量子态数据, $\boldsymbol{U}$ 表示变分电路进行的酉变换. 将编码电路和变分电路的变换代入式(17)得到整体损失函数:

$$C(\boldsymbol{\theta})=\sum_i(\langle\psi_0|\boldsymbol{V}^{\dagger}(x_i)\boldsymbol{U}^{\dagger}(\boldsymbol{\theta})\boldsymbol{H}\boldsymbol{U}(\boldsymbol{\theta})\boldsymbol{V}(x_i)|\psi_0\rangle-y_i)^2. \tag{18}$$

编码电路和量子分类器使用的参数化量子电路的结构都具有多种形式, 对于不同的分类任务和不同类型的数据, 选择合适的编码电路和量子分类器能够有效地提高算法的性能.

最初 Farhi 等人[98]设计了一个能够对二进制串进行二分类的分类器, 只使用单位门和非门就完成了分类任务. 在这个基础上引入了含参量子门, 在 MNIST 数据集上完成了二分类任务. 同时在简单构建的量子态数据集上进行了实验, 准确率达到了 97%. 基于这种框架发展出的分类器被称作量子电路学习(quantum circuit learning, QCL)[84].

在近期对量子分类器的研究中, 由于 NISQ 设备可供使用的量子比特较少, 因此在编码前需要对输入数据进行降维. 一般采用特征提取器和变分量子电路相结合的方式, 特征提取器从输入数据中提取特征并产生低维特征向量, 将其输入到变分量子电路中进行分类. 主成分分析是一种经典降维算法, 将其作为变分量子电路预处理的方法(PCA-VQC)[30]可以作为变分量子分类器的基线. 结合不同的降维算法的变分量子电路分类器将提高量子分类算法性能作为一个主要研究方向, 一般通过经典降维算法进行预处理和经典神经网络作为后处理完成分类任务, 如图 5 所示.

Grant 等人[32]提出了量子 TTN 分类器和 MERA 分类器. TTN 分类器的量子电路结构采用最近邻双量子比特酉门, 每次输出取 1 个量子比特, 故下层量子比特数减半. 反复进行如上酉变换和丢弃操作直到仅剩 1 个量子比特, 测量单量子比特算子 M 的期望值, 经过后处理得到最终分类结果. MERA 结构同 TTN 类似, 在每层中原有的每 2 个酉门之间添加 1 个双量子比特酉门, 这些额外添加的酉门有助于在特定尺度上获取量子关联. 如图 6 所示, 8 量子比特 TTN 结构和 MERA 结构通过适当的策略能够在多个 1 维、2 维数据集上得到较好的效果, 后续以这 2 种张量网络结构的量子电路为主体发展出多种变分量子分类器.

Table 6 Machine Learning Algorithms Based on Variational Quantum Circuits for Classification Tasks

表 6 分类任务上基于变分量子电路的机器学习算法

模型	数据集	任务	环境	量子位	参数量	准确率/%	
						训练集	测试集
VQM[98]	MNIST	二分类	模拟	17	136		90
TTN[32]	Iris	二分类	模拟	4	7		98.92
TTN[32]	MNIST	二分类	模拟	8	7		97.63
MERA[32]	MNIST	二分类	模拟	8	11		98.86
Hybrid[32]（TTN 预训练过的 MERA）	MNIST	二分类	模拟	8	11		98.46
TTN[32]	合成量子数据集	二分类	模拟	8	7		60.45
PCA-VQC[30]	MNIST	二分类	模拟	4	12	87.29	87.34
MPS-VQC[30]	MNIST	二分类	模拟	4	12	99.91	99.44
QTN-VQC[33]	MNIST	二分类	模拟	8	328		91.43
QTN-VQC[33]	MNIST	二分类	模拟	12	4 464		92.36
QTN-VQC[33]	MNIST	二分类	模拟	16	600		92.28
VQTN[10]	Iris	三分类	模拟	2	3		100
VQTN(TTN)[10]	MNIST	二分类	模拟	8	12		97.80
VQTN(TTN)[10]	MNIST	二分类	模拟	16	28		97.45
VQTN(MERA)[10]	MNIST	二分类	模拟	8	18		97.92
VQTN[10]	MNIST-4	四分类	模拟				82.19
QCNN[70]	MNIST	十分类	模拟	4	6		95
Noisy QCNN[71]	MNIST	二分类	模拟	14	46	94.8	96.0
Noisy QCNN[71]	MNIST	十分类	模拟	14	379	74.2	74.0
Noisy-free QCNN[71]	MNIST	二分类	模拟	14	46	95.4	96.3
Noisy-free QCNN[71]	MNIST	十分类	模拟	14	379	75.6	74.3
QCCNN[38]	Tetri	二分类	模拟	4	16		≈100
QCCNN[38]	Tetri	四分类	模拟	4	16		≈100
QMLP[111]	MNIST	十分类	模拟	16	128		75
QMLP[111](比特翻转)	MNIST	十分类	模拟	16	128		63
QMLP[111](相位翻转)	MNIST	十分类	模拟	16	128		67
VSQL[34]	MNIST	二分类	模拟	2	35		99.52
VSQL[34]	MNIST（1 000 个样本）	十分类	模拟	9	928		87.39
VSQL[34]	含噪量子态	二分类	模拟	2			100
VSQL[34]	不含噪量子态	三分类	模拟	2			100
HNN[58]	BAS	二分类	模拟	10	20		100
HNN[58]	BAS	二分类	量子	10	20		33.33
HNN[58]	Iris	三分类	模拟	10	20	89.88	91.5
HNN[58]	Iris	三分类	量子	10	20	28.12	37.5

注：数据取相应论文给出的最优模型数据，使用相同数据集的相同任务之间仍存在差异，例如 MNIST 数据集二分类任务可以为 2 个数字的分类、是否为偶数的分类、是否大于 4 的分类等，并非完全一致. 模拟环境是指使用经典计算机模拟的环境，量子环境是指使用量子计算机上运行相应算法.

Chen 等人[30,112]提出了基于张量网络和变分量子电路的模型(TN-VQC)，使用张量网络进行特征提取，允许端到端的训练. 在 MNIST 数据集上完成的二分类任务主要依靠矩阵乘积态作为特征提取器，实验表明 MPS-VQC 方案远超 PCA-VQC 方案.

Qi 等人[33]设计了一种用于量子嵌入的张量积编码(TPE)，与 TTN 结合作为量子张量网络(QTN)进行量子嵌入，提出了 QTN-VQC 模型，在 MNIST 数据集上性能有所提升.

Huang 等人[10]提出了 VQTN 算法，使用核编码方

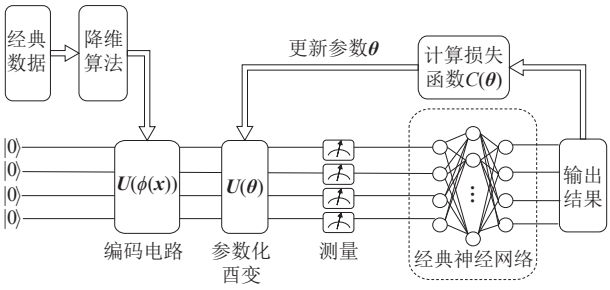


Fig. 5 Quantum classifiers with classical pre-processing and post-processing

图 5 经典前后处理的量子分类器

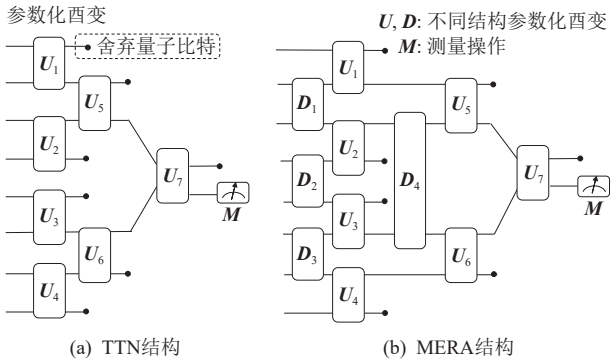


Fig. 6 Structures of TTN and MERA

图 6 TTN 和 MERA 结构

案，TTN 和 MERA 作为电路主体，在对变分量子电路模型输出态测量后加入一个经典神经网络结构用于提取隐藏特征，在 VQTN 中经典优化器更新的参数包括变分量子电路的参数和经典神经网络的参数. VQTN 在 Iris 数据集上和 MNIST 数据集上相对 QTN 算法有所提升. 同时，VQTN 使用的参数量和量子位以及算法空间复杂度均有所减少.

除了量子电路学习框架外，Schuld 等人[109]提出一种具有新结构的中心电路量子分类器(circuit-centric quantum classifier)，其可学习的参数量与输入维度多对数相关，即$O(poly(\mathrm{lb}n))$. 该方法实现了量子随机失活(Dropout)，并且提出了一种混合量子-经典梯度下降的方式训练模型. 在 MNIST, Cancer 等多个数据集上的结果均表现出对噪声鲁棒性良好.

在分类任务中，为了在没有训练过的数据样本上保持较高的分类精度，一般会在训练时引入正则项. Chen 等人[30]提到 VQC 相当于对 MPS 提供了正则化，Schuld 等人[109]提出的 Dropout 也是一种正则化方法，如何在参数化量子电路构成的分类器中引入正则化技术，将是未来研究的方向之一.

4.1.2 量子卷积神经网络

变分量子电路被证明在模拟经典人工神经元上有着良好的前景，相比经典神经元具有更好的可表达性[22]. 因此，很多研究关注于使用变分量子电路模拟经典神经网络，得到具有量子优势的量子神经网络. 同时，将量子神经网络与经典神经网络相结合的方法也正在迅速发展，通过经典神经网络帮助量子神经网络解决一些实际应用中的问题，例如降噪等. 对于经典的神经网络，研究其量子版本是极为重要的方向. 与其他量子神经网络不同的是，Pesah 等人[113]表明由$O(\mathrm{lb}n)$层组成构建的量子卷积神经网络(quantum convolutional neural network, QCNN)不会出现贫瘠高原[114]的现象，这一结果保证了量子卷积神经网络在随机初始化参数时的可训练性，这一特点使得量子卷积神经网络成为量子有监督学习的研究重点.

Cong 等人[11]引入了 QCNN，如图 7 所示. 使用量子电路模仿给定滤波器的卷积层，实现了量子版的卷积层、池化层与全连接层. 该模型仅有$O(\mathrm{lb}N)$个变分参数，其中 N 是量子比特数. 电路的输入是一个未知的量子态 $\boldsymbol{\rho}$，将单准局部酉元 $\boldsymbol{U}$ 应用在有限深度的电路上作为卷积层，通过测量部分量子比特，根据测量结果决定其附近量子比特的酉旋转 $\boldsymbol{V}$ 的方式实现池化，最后使用一个酉元 $\boldsymbol{F}$ 作为全连接层. 该 QCNN 的超参数与经典情况一样都是确定的. Oh 等人[70]在 MNIST 上进行了实验，成功构建了 QCNN 并完成了分类任务.

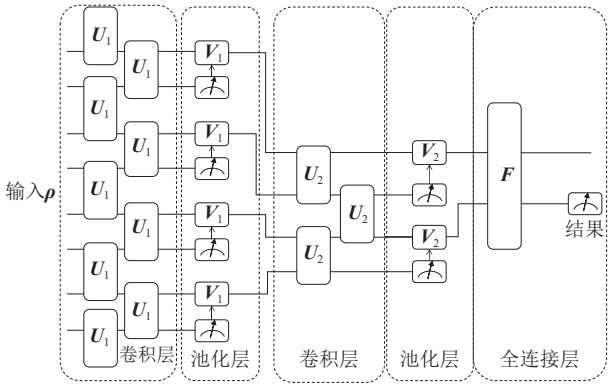


Fig. 7 Quantum convolutional neural network

图 7 量子卷积神经网络

Wei 等人[71]提出一种不同结构的 QCNN，实现了基于酉算子线性组合构建的卷积层 LCU，通过舍弃部分量子比特的方式模拟池化层效果，全连接层通过测量一个参数化哈密顿量的期望值并使用一个非线性激活函数后处理实现，可以提供精确对应于特定经典卷积核的可实现变分量子线路. 该模型仅有$O((\mathrm{lb}M)^6)$的基本门复杂度，相对经典卷积神经网络$O(M^2)$的参数量，计算复杂度有指数级的降低，其中

M 是输入方阵的维度. 然而文献 [11, 70−71] 方法具有显著的缺点, QCNN 属于不含任何经典结构的全量子神经网络, 具有输入需要量子比特数较多的特点, 而且需要量子随机存取器的辅助, 只适用于特定类型的量子数据嵌入, 不适合在 NISQ 计算机上实现.

Henderson 等人[110] 使用非参数随机量子电路进行特征映射, 使用参数化量子电路构建的卷积层和经典池化层相结合共同提取特征图, 这种方法被称为量子化卷积神经网络(quanvolutioanl nertual network). Liu 等人[38] 将其与 QCNN 结合提出了量子-经典卷积神经网络(quantum-classical convolutional neural network, QCCNN). 该方法需要的量子比特数只和特征图大小有关, 只需要 mn 的量子比特数, 其中 m 和 n 表示滤波窗口的形状, 可以基于 QCNN 进行各种判别模型和生成模型的实现, QCNN 的问题在于量子特征映射需要 $O(Nd)$ 次量子门运算, 其中 N, d 分别表示量子比特数和电路层数, 相比 $O(N)$ 次浮点数运算的 CNN 不具有优势.

Chen 等人[16] 提出了基于 QCNN 的盒子计数分形特征的二值图像分类方法, 构建的 QCNN、混合 QCNN、多量子层混合 QCNN 模型在乳腺癌数据上能够达到 83%~88% 的精确度, 比经典卷积神经网络模型高了 1~4 个百分点.

4.1.3 影子量子电路

受到影子学习的启发, Li 等人[34] 提出了变分影子量子学习(variational shadow quantum learning, VSQL), 4 量子比特结构示意图如图 8 所示.

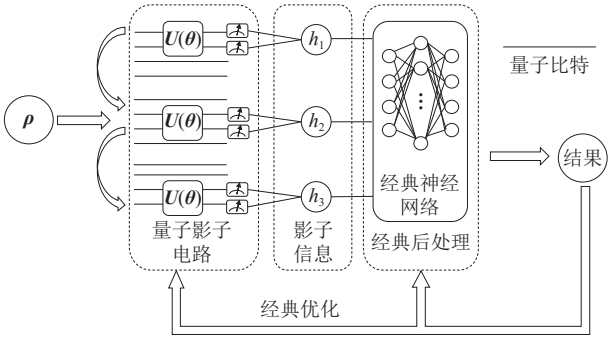


Fig. 8 Illustration of VSQL

图 8 变分影子量子学习示意图

准备一个变分量子电路, 每次作用于部分输入量子比特并对结果进行测量, 即每次从希尔伯特空间的子空间上学习部分特征, 每次测量后将所有门向下平移 1 个量子比特. 每次的测量结果被称为一个影子特征, 将所有影子特征输入至经典神经网络进行后处理. 最后通过经典优化算法对变分量子电路和经典神经网络的参数进行调整.

变分影子电路在含噪量子态分类的实验上能够达到 100% 的准确率, 鲁棒性较强. 由于每次获取的特征是局部特征, 不易受拓扑连接限制. 仅使用 1 个变分量子电路来获取所有影子特征使得模型参数量减少. 变分影子电路还能够避免产生贫瘠高原问题. 以上优势使得变分影子电路适用于 NISQ 设备.

4.1.4 其 他

除此之外, 还有一些值得关注的工作. Monteiro 等人[115] 提出二分类感知器的量子版本. Chu 等人[111] 提出了一种量子多层感知器, 其结构具有容错编码且能够提供丰富的非线性运算. 相对于 QuantumNAS[12] 所寻找到的结构, 精度提高 10%, 同时量子门数量减少 2 倍, 参数量减少了 3 倍.

对于图卷积神经网络, 其量子版 QGCNN(quantum graph convolutional neural network)的实现中使用了变分量子电路提取特征, 用于完成半监督分类任务[116]. Hur 等人[39] 提出了全参数化 QCNN 模型, 只需要最近邻量子比特相互作用, 在 NISQ 设备上实现所需的理想特征, 同时对不同参数下的 QCNN 进行了基准测评.

Arthur[58] 提出一种混合量子-经典神经网络架构 HNN, 使用变分量子电路作为神经元, 如图 9 所示.

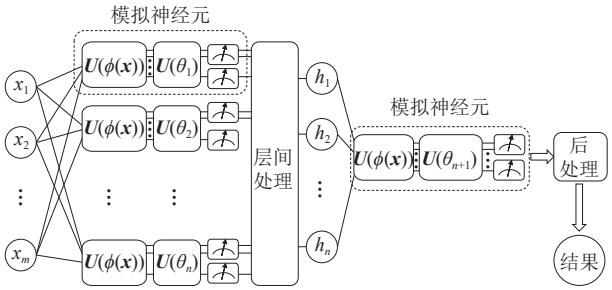


Fig. 9 Illustration of HNN frame with single hidden layer

图 9 单隐藏层 HNN 架构示意图

对于 m 维的输入向量, 采用 n 个变分量子分类器分别对其进行分类, 每个电路输出经过测量得到一个标签, 层间处理将结果堆叠为一个 n 维向量, 再经过一个变换得到该隐藏层的输出, 最后使用一个变分量子电路作为输出层. 共使用 $n\times l+1$ 个变分量子电路, 其中 n 为每个隐藏层神经元数目, l 为隐藏层层数. 这种结构可以看作是一种特殊情况下的影子量子电路, 每个变分量子电路作用于全部输入. 在仿真实验中, 在二进制数据集上的精确度相对单个变分量子电路构建的量子分类器提高了 8~11 个百分点, 在损失函数最优解上表现也更好, 减小了 20~40 个百

分点. 然而在使用量子硬件时, 该神经网络和量子分类器的表现都较差, 推测可能是因为硬件的限制, 量子比特与量子门数量的增加对保真度造成了影响.

4.1.5 总 结

基于变分量子电路的有监督学习的研究将会围绕 5 个方面展开:

1)使用更复杂的特征映射和 Ansatz 构成的变分量子电路;

2)通过变分量子电路的组合, 构造更复杂的网络结构;

3)将经典神经网络中的学习率等经典神经网络中的超参数以调优的训练方式引入;

4)探索固定任务下变分量子电路的最优结构并设计结构搜索算法;

5)将量子分类器与经典神经网络结合构建混合量子-经典结构.

4.2 无监督学习

在无监督学习任务中, 变分量子算法主要围绕生成模型展开. 使用经典机器学习方法完成的生成模型已经被广泛应用到量子态近似、图像生成、药物设计、缺失文本推理等诸多任务之中. 生成模型的目标是学习给定数据中未知的概率分布 p 并进行建模, 根据建模生成新数据样本 $q_{\boldsymbol{\theta}}$. 任务具体表示为

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} D(p, q_{\boldsymbol{\theta}}), \tag{19}$$

其中 D 表示对样本之间相似度的度量方式, 度量结果越小说明相似度越高, 我们希望能够找到使度量结果最小的生成样本.

当前量子机器学习数据集尚不完善, 大多使用经典数据集而非量子数据集, 因此需要将经典数据编码到量子态上, 而精确制备一个 n 量子比特的量子态需要 $O(2^n)$ 量子门, 这会抵消量子算法所带来的量子加速优势. 无监督学习能够有效地将经典数据加载到量子态上, 量子无监督学习模型已经成为一种有效的数据编码方法[15].

4.2.1 量子玻恩机

量子玻恩机(quantum circuit Born machine, QCBM)是基于波函数的概率解释的一种模型, 其目的是生成与给定分布相同的分布, 参数化量子电路被用来从概率分布中抽样, 也称为数据驱动量子电路学习(data-driven quantum circuit learning, DDQCL)[13].

给定数据集 D, 存在数据概率分布 P_D. 参数化量子电路将初态 $|0\rangle$ 演化为 $|\psi_{\boldsymbol{\theta}}\rangle$, 在计算基上测量得到生成数据, 其概率分布为 $P_{\boldsymbol{\theta}}(\boldsymbol{x}) = |\langle \boldsymbol{x} \mid \psi_{\boldsymbol{\theta}}\rangle|^2$. 生成模型的训练目标是使生成数据概率分布 $P_{\boldsymbol{\theta}}(\boldsymbol{x})$ 与目标概率分布 P_D 更接近, 文献 [13] 中使用 KL 散度 $D_{\mathrm{KL}}[P_D|P_{\boldsymbol{\theta}}]$ 作为距离度量:

$$D_{\mathrm{KL}}[P_D|P_{\boldsymbol{\theta}}] = \sum_{\boldsymbol{x}} P_D(\boldsymbol{x}) \ln \frac{P_D(\boldsymbol{x})}{P_{\boldsymbol{\theta}}(\boldsymbol{x})}. \tag{20}$$

使用其负对数似然函数的变体定义损失函数:

$$C(\boldsymbol{\theta}) = -\frac{1}{D}\sum_{d=1}^{D} \ln(\max(\varepsilon, P_{\boldsymbol{\theta}}(\boldsymbol{x}^{(d)}))). \tag{21}$$

文献 [13] 采用 PSO 作为优化算法最小化 KL 散度. Zhu 等人[76] 在离子阱计算机上实现了 QCBM, 在 BAS 数据集上表现出对量子硬件噪声的鲁棒性. 同时对比了 PSO 和 BO 优化算法的效果, 在 KL 散度、qBAS 得分以及纠缠熵的评价指标上 BO 均体现出优势.

Liu 等人[40] 提出 MMD 损失函数, 实现了一种可以使用基于梯度优化算法的 QCBM.

$$\begin{aligned}\mathcal{L} = \left\| \sum_{\boldsymbol{x}} P_{\boldsymbol{\theta}}(\boldsymbol{x})\phi(\boldsymbol{x}) - \sum_{\boldsymbol{x}} P_D\phi(\boldsymbol{x}) \right\|^2 = \\ \underset{\boldsymbol{x}\sim P_{\boldsymbol{\theta}}, \boldsymbol{y}\sim P_{\boldsymbol{\theta}}}{E}[K(\boldsymbol{x},\boldsymbol{y})] - 2\underset{\boldsymbol{x}\sim P_{\boldsymbol{\theta}}, \boldsymbol{y}\sim P_D}{E}[K(\boldsymbol{x},\boldsymbol{y})] + \\ \underset{\boldsymbol{x}\sim P_D, \boldsymbol{y}\sim P_D}{E}[K(\boldsymbol{x},\boldsymbol{y})],\end{aligned} \tag{22}$$

其中 $\phi(\boldsymbol{x})$ 为编码映射函数, 核函数 $K(\boldsymbol{x},\boldsymbol{y}) = \phi(\boldsymbol{x})^{\mathrm{T}}\phi(\boldsymbol{y})$, 该文献使用高斯核 $K(\boldsymbol{x},\boldsymbol{y}) = \mathrm{e}^{-\frac{|\boldsymbol{x}-\boldsymbol{y}|^2}{2\sigma}}$ 来衡量 2 种分布在不同尺度下的差异. MMD 是应用最多的损失函数, 被广泛应用于各种模型[31,41,101]. 其中 Gong 等人[31] 实现了基于 MPS 的 QCBM, 使用 MMD 损失函数进行训练, 通过重置节点的方式在保持 KL 散度指标较低的同时减少了量子比特的使用.

Čepaitė等人[101] 提出了基于连续变分量子计算的连续变分玻恩机(CVBM), 能够学习量子和经典连续数据分布, 同时对噪声鲁棒性较好.

Coyle 等人[102] 提出了量子伊辛玻恩机, 将总变分距离作为损失函数, 当其小于一定阈值时可以实现量子优势. 同时使用 QCIBM 和 QAOA 来学习瞬时量子多项式时间电路的输出分布[117].

QCBM 主要应用在金融领域, 可以用于学习金融数据分布与生成金融数据[41,80,82]. 将 QCBM 应用于真实金融数据的研究表明, 这些量子模型具有不差于经典模型的性能. 作为一个简单的生成模型框架, QCBM 也被用于对比不同优化算法性能[42,76].

4.2.2 量子玻尔兹曼机

量子玻尔兹曼机(quantum Boltzman machine, QBM)通过一组相互作用的自旋量子对表示量子神经网络, 其对应一个可调的 Ising 模型, 可以用于生成

学习和判别学习. 经典玻尔兹曼机最耗时的部分就是计算热平衡平均值的过程, 量子玻尔兹曼机能够用于构造哈密顿量从而提高效率. 过去大多数研究以基于边界训练和基于状态的训练为主[118-119].

Zoufal 等人[14]提出一种基于热场双重技术和变分量子虚时演化的方法 VarQITE 来学习离散概率分布, 使用经典优化器学习哈密顿量的参数, 这种基于变分量子电路实现的量子玻尔兹曼机被称为变分量子玻尔兹曼机(variational quantum Boltzman machine, VarQBM). 通过 VarQITE 制备吉布斯态的近似态 $\boldsymbol{\rho}^{\text{Gibbs}}$, 通过训练其哈密顿参数使得采样概率分布 $\boldsymbol{\rho}^{\text{QBM}}$ 尽可能接近训练数据分布 $\boldsymbol{\rho}^{\text{data}}$. 一般情况下采用交叉熵损失函数:

$$\min_{\boldsymbol{\theta}} L=\min_{\boldsymbol{\theta}}\left(-\sum_{\boldsymbol{x}}\boldsymbol{\rho}_{\boldsymbol{x}}^{\text{data}}\text{lb}\boldsymbol{\rho}_{\boldsymbol{x}}^{\text{QBM}}\right), \tag{23}$$

其中 $\boldsymbol{\rho}_{\boldsymbol{x}}^{\text{data}}$ 和 $\boldsymbol{\rho}_{\boldsymbol{x}}^{\text{QBM}}$ 分别表示结果 $\boldsymbol{x}$ 在训练数据和生成数据中出现的概率. 使用自动微分对参数进行更新, 在模拟实验上完成的生成式任务和判别式任务都表明变分量子玻尔兹曼机具有良好的抗噪性. Shingu 等人[72]也提出一种变分量子玻尔兹曼机, 同样使用了变分量子虚时演化算法. 该方法最小化精确演化和参数化实验状态的距离:

$$\delta\left\|\left(\frac{\partial}{\partial r}+\hat{\boldsymbol{H}}-\langle\psi(\boldsymbol{\theta}(\tau))|\hat{\boldsymbol{H}}|\psi(\boldsymbol{\theta}(\tau))\rangle\right)|\psi(\boldsymbol{\theta}(\tau))\rangle\right\|=0, \tag{24}$$

其中实验状态 $|\psi(\boldsymbol{\theta}(\tau))\rangle$ 通过变分量子电路构造:

$$|\psi(\boldsymbol{\theta}(\tau))\rangle=\boldsymbol{U}(\boldsymbol{\theta}(\tau))|\psi(0)\rangle. \tag{25}$$

2024 年, Huijgen 等人[120]提出使用 β–变分量子本征求解器训练量子玻尔兹曼机.

4.2.3 量子自编码器

自编码器是一种十分有效的维度压缩算法, 目的是对维度压缩后的样本进行重构, 经典自编码器一般通过经典神经网络实现. 量子自编码器(quantum autoencoder, QAE)完成的任务大多数是量子的, 主要的目的是实现对量子态的制备和压缩, 也能够被用于实现经典数据的压缩与重建.

在经典的自编码器中, 一般通过神经网络将 $n+k$ 维的数据压缩至 n 维, 这个过程被称作编码, 压缩后的 n 维数据被称为压缩数据. 如果压缩数据能够经过另一个神经网络生成与初始 $n+k$ 维数据分布相似的数据, 那么我们可以认为这个经典自编码器是良好的, 压缩过程是有效的. Romero 等人[65]首次将变分量子算法引入自编码器中, 将其应用于压缩哈伯德模型和分子哈密顿量的基态. 后续对于量子自编码器的研究几乎都采用了 Romero 等人的框架. 在该框架中, 量子自编码器的输入是 $n+k$ 个量子比特上纯态的集合 $\{p_i,|\psi_i\rangle_{AB}\}$, 子系统 A, B 分别由 n 和 k 个量子比特构成. 在编码的过程中, 量子自编码器将 A, B 系统上的信息编码到 n 个量子比特上.

一般使用保真度的期望 $F(|\psi_i\rangle,\boldsymbol{\rho}_i^{\text{out}})=\langle\psi_i|\boldsymbol{\rho}_i^{\text{out}}|\psi_i\rangle$ 作为初始态和输出量子态偏差的度量, 其中 $|\psi_i\rangle$ 表示初始的量子态, $\boldsymbol{\rho}_i^{\text{out}}$ 表示最终输出的密度矩阵. 当 $F(|\psi_i\rangle,\boldsymbol{\rho}_i^{\text{out}})\approx 1$ 时可以认为编码前的初始量子态和解码后的输出量子态是相似的, 完成了一次成功的编码-解码过程. 通过经典训练方法寻找使保真度最大的酉算子族, 因此损失函数可以通过保真度来定义:

$$C_1(\boldsymbol{\theta})=\sum_i p_i F(|\psi_i\rangle,\boldsymbol{\rho}_{i,\boldsymbol{\theta}}^{\text{out}}). \tag{26}$$

设 $\boldsymbol{U}^{\boldsymbol{\theta}}$ 是 $n+k$ 量子比特上的一个酉算子族, $\boldsymbol{\theta}=(\theta_1,\theta_2,\ldots)$ 是定义酉量子电路的一组实参数, 密度矩阵表示为

$$\boldsymbol{\rho}_{i,\boldsymbol{\theta}}^{\text{out}}=(\boldsymbol{U}^{\boldsymbol{\theta}})_{AB'}^{\dagger}\text{tr}_B\left[(\boldsymbol{U}^{\boldsymbol{\theta}})_{AB}\left[\psi_{i_{AB}}\otimes a_{B'}\right](\boldsymbol{U}^{\boldsymbol{\theta}})_{AB}^{\dagger}\right](\boldsymbol{U}^{\boldsymbol{\theta}})_{AB'}, \tag{27}$$

其中 $|\psi_i\rangle\langle\psi_i|_{AB}=\psi_{i_{AB}}$, $|a\rangle\langle a|_{B'}=a_{B'}$.

由于密度矩阵计算的困难性, 引入了一个 k 量子比特的固定纯态系统 B', 使用 SWAP 电路将对 B' 的追踪代替对系统 B 的追踪, 如图 10 所示. B 系统的压缩状态和 B' 系统的初态的保真度与式(26)定义的保真度等价, 故损失函数也可以定义为

$$C_2(\boldsymbol{\theta})=\sum_i p_i F(\text{tr}_A\left[\boldsymbol{U}^{\boldsymbol{\theta}}|\psi_i\rangle\langle\psi_i|_{AB}(\boldsymbol{U}^{\boldsymbol{\theta}})^{\dagger}\right],|a\rangle_B). \tag{28}$$

Khoshaman 等人[121]最先提出量子变分自编码器, 使用 QBM 代替了 DVAE 中的 RBM. 同时为了克服传统方法对负相位精确采样难以实现这一限制, 提出了 Q-ELBO 下界. Romero 等人[65]和 Khoshaman 等人[121]提出的 QVAE 的实现采用的都是量子退火方法, 下面将着重讨论基于门电路的量子自编码器.

基于变分量子电路的量子自编码器研究主要围绕以下进行改进: 设计具有更复杂结构的 Ansatz、改进损失函数, 以及经典优化器的使用.

在经典任务上, Li 等人[43]提出了基线量子自编

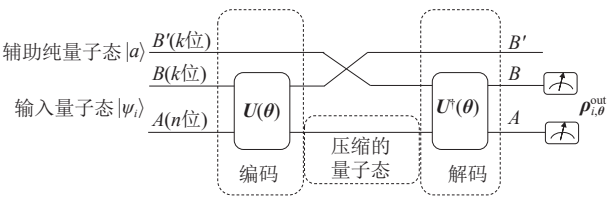


Fig. 10 Illustration of quantum autoencoder

图 10 量子自编码器示意图

码器(BQ-VAE)和可扩展的量子生成自编码器(SQ-VAE). 使用可调量子层深度、异构学习率、补丁量子电路(patch-VQC)等混合量子经典神经网络架构对高维数据集进行学习，并将其用于药物搜索. 使用一个补丁变分量子电路作为主体，其每个补丁具有 3 个量子比特，具有量子输入层、隐藏层、输出层. 输入层作为编码器，使用量子编码的方式将输入数据编码到补丁电路的输入上，隐藏层由重复的 PQC 结构构成，输出层对所有量子比特进行测量，最后使用一个经典全连接层将测量结果连接，图 11 所示为使用 2 个补丁电路的情况. BQ-VAE 在低维数据上的学习速度相对经典 VAE 要高一些，SQ-VAE 用于模拟高维分子，进行经典 VAE 难以实现的高维药物分子的探索.

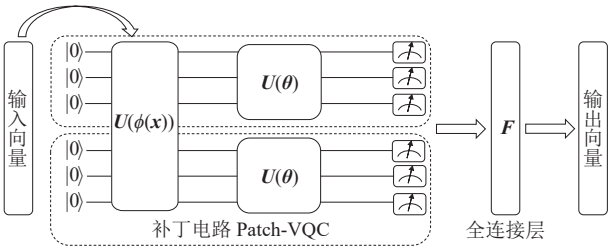


Fig. 11 Illustration of SQ-VAE

图 11 SQ-VAE 示意图

量子自编码器应用之一是进行量子态制备. 一个量子自编码器可以经过训练用于压缩一组特定的量子态，使用解压酉矩阵就能够用于生成和压缩前相似的量子态. 在变分量子算法中，使用量子自编码器作为初态制备电路具有参数量较少的优势. 如果使用经典数据作为输入、经典神经网络作为编码器、变分量子电路作为解码器，可以实现将经典数据编码到量子态上的量子自编码器. 近期，Wu 等人[15]提出了一种其深度与量子比特数线性相关的量子自编码器，能够高精度地构造逼近未知经典分布的量子态. 该量子自编码器引入了新的编码方案和纯度惩罚项 $\mathcal{L}_{\text{penalty}} = -\sum_{i=1}^{n} \text{tr}(\bar{\rho}_i^2)$，损失函数 $\mathcal{L} = \mathcal{L}_{\text{recon}} + \beta \mathcal{L}_{\text{penalty}}$. 使用经典神经网络作为编码器与变分量子电路实现的解码器相结合，在 10 量子比特数据集上的仿真实现表明该方法相对于现有方法具有更快的收敛速度和稳定性，在学习高维数据分布的任务上误差更小、鲁棒性更强. 该模型的训练复杂度为 $O(m^2)$，测试复杂度为 $O(mn)$，m 为数据集样本数量，n 为量子比特数. 在保证量子自编码器的效率和收敛速度的同时，尽可能地减少变分量子电路的深度是 NISQ 时代主要的研究方向.

量子态的压缩能够将量子信息压缩到低维空间的量子自编码器，对于实现量子信息领域的数据自动压缩具有重要意义. Huang 等人[122]分析了实现完美量子自编码器的条件，并从理论上证明了如果输入态的最大线性无关向量的个数不超过隐空间的维数，量子自编码器可以将高维量子信息无损压缩到隐空间. 由此设计了将 2 个双量子比特量子态压缩至 2 个单量子比特量子态的量子自编码器. 其除应用于量子态压缩外，还能够被用于区分 2 组非正交态.

对于变分量子算法，全局形式的损失函数具有贫瘠高原现象[114]，这对变分量子算法的优化过程造成了困难. 由此 Cerezo 等人[123]提出了使用局部损失函数的量子自编码器，使得量子自编码器的可训练性大大提高.

Bravo-Prieto[66]提出了一种增强特征量子自编码器 EF-QAE，将一个特征向量介入变分量子电路中，能够以更高的保真度实现量子态的压缩，在使用相同量子资源的前提下，EF-QAE 使用了额外的经典优化，获得了比标准量子自编码器更好的性能. 使用更多经典优化手段构建的性能更佳、使用的量子资源更少的变分量子算法更适合在近期的量子设备上实现.

4.3 半监督学习

量子半监督学习模型围绕量子生成对抗网络(QGAN)展开. 在经典生成对抗网络(GAN)中，模型主要由生成器和判别器构成. 生成器将噪声映射到真实样本空间中，目标是使噪声样本分布最大程度接近真实样本分布，判别器需要最大程度对混合样本进行区分. 模型通过生成器和判别器的对抗使生成样本逐渐接近于真实样本，进而学习到概率分布. 训练目标可以表示为

$$\min_{G} \max_{D} loss(\boldsymbol{\theta}_G, \boldsymbol{\theta}_D), \tag{29}$$

其中 G 是生成器，D 是判别器.

最早的 QGAN 概念由 Lloyd 等人[74]提出，他们认为 QGAN 的任务可以是经典的或量子的，使用的生成器和判别器也可以是经典的或量子的，由此将 QGAN 按任务和实现方式是否为量子的分为 4 类，并对其可行性和性能进行了理论分析. 配套文献 [124] 进行了一些简单的数值实验，对 QGAN 可行性进行了初步的探索，QGAN 的整体框架如图 12 所示.

QGAN 的训练目标与 GAN 相同，其量子形式为：

$$\min_{\boldsymbol{\theta}_G} \max_{\boldsymbol{\theta}_D} \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} Pr((D(\boldsymbol{\theta}_D, |\lambda\rangle, R(|\lambda\rangle)) = |real\rangle) \cap (D(\boldsymbol{\theta}_D, |\lambda\rangle, G(\boldsymbol{\theta}_G, |\lambda\rangle, z)) = |fake\rangle)), \tag{30}$$

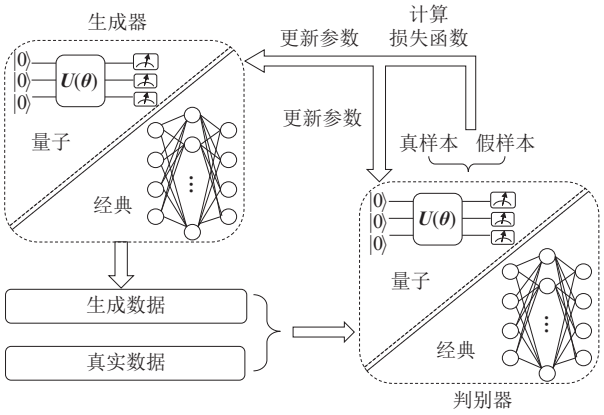


Fig. 12 Quantum generative adversarial networks

图 12 量子生成式对抗网络

其中$|\lambda\rangle$为数据的标签, R 为真实数据源, 其根据$|\lambda\rangle$获得真实数据. G 为生成器, 其根据$|\lambda\rangle$和量子噪声$|z\rangle$ 生成假数据, $\boldsymbol{\theta}_G$为生成器参数. D 为判别器, 对真实或生成数据进行判断, 若判断为真实数据则输出量子态$|real\rangle$, 否则输出量子态$|fake\rangle$, $\boldsymbol{\theta}_D$为判别器参数. Λ 为基数, 表示标签数. Dallaire-Demers 等人[124] 给出的 QGAN 损失函数如下:

$$C(\boldsymbol{\theta}_G,\boldsymbol{\theta}_D)=\frac{1}{2}+\frac{1}{2\Lambda}\sum_{\lambda=1}^{\Lambda}(\cos^2(\phi)\mathrm{tr}(Z\boldsymbol{\rho}_\lambda^{DR}(\boldsymbol{\theta}_D))-$$
$$\sin^2(\phi)\mathrm{tr}(Z\boldsymbol{\rho}_\lambda^{DG}(\boldsymbol{\theta}_D,\boldsymbol{\theta}_G,z))),\tag{31}$$

其中 Z 表示测量结果期望值, 其中$\boldsymbol{\rho}_\lambda^{DR}(\boldsymbol{\theta}_D)$表示初始量子态经过真实数据酉变换和判别器酉变换后的量子态, $\boldsymbol{\rho}_\lambda^{DG}(\boldsymbol{\theta}_D,\boldsymbol{\theta}_G,z)$表示初始量子态$\boldsymbol{\rho}_\lambda^0$经过生成器酉变换和判别器酉变换后的量子态:

$$\begin{aligned}&\boldsymbol{\rho}_\lambda^{DR}=\boldsymbol{U}_D(\boldsymbol{\theta}_D)\boldsymbol{U}_R\boldsymbol{\rho}_\lambda^0\boldsymbol{U}_R^\dagger\boldsymbol{U}_D^\dagger(\boldsymbol{\theta}_D),\\&\boldsymbol{\rho}_\lambda^{DG}=\boldsymbol{U}_D(\boldsymbol{\theta}_D)\boldsymbol{U}_G(\boldsymbol{\theta}_G)\boldsymbol{\rho}_\lambda^0(z)\boldsymbol{U}_G^\dagger(\boldsymbol{\theta}_G)\boldsymbol{U}_D^\dagger(\boldsymbol{\theta}_D).\end{aligned}\tag{32}$$

式(31)中ϕ将真实样本和假样本在总样本中所占比例参数化, 一般取$\phi=\frac{\pi}{4}$, 此时损失函数可以抽象表示为

$$\frac{1}{2}+\frac{1}{4}\mathrm{tr}[|\sigma-\delta|],\tag{33}$$

其中σ和δ分别为用于刻画训练样本和生成样本的密度算符. 大部分 QGAN 采用交替训练的方式, 直到判别器和生成器达到纳什均衡为止.

根据生成器和判别器也可以是经典的或量子的, 实际上可以更细地将 QGAN 分为 8 种情况, Tian 等人[125] 给出了命名方式, 如表 7 所示. 下面对 8 种 QGAN 的可行性和研究进展进行讨论.

使用经典分类器和经典判别器可以对经典数据和量子数据进行处理, 本节主要讨论基于变分量子电路实现的模型. 近期研究的 QGAN 旨在量子计算机上实现生成器或判别器, 从而加速生成网络或判别网络的速度.

Table 7 Classification of QGANs and Related Researches

表 7 QGAN 分类及相关研究

任务	生成器	判别器	名称	相关研究
经典	经典	经典	CT-CGCD	文献 [126]
经典	经典	量子	CT-CGQD	文献 [46, 104, 107]
经典	量子	经典	CT-QGCD	文献 [19, 44−49, 59]
经典	量子	量子	CT-QGQD	文献 [46, 59]
量子	经典	经典	QT-CGCD	文献 [127]
量子	经典	量子	QT-CGQD	
量子	量子	经典	QT-QGCD	文献 [50, 62]
量子	量子	量子	QT-QGQD	文献 [73−74, 124, 128−129]

注: 采用文献 [125] 给出的命名方式, 名称中的字母 T, G, D 分别表示任务、生成器、判别器. C, Q 分别表示是通过经典还是量子方法完成的. 经典生成器与量子判别器构成的 QGAN 对于量子数据无法收敛到纳什均衡, 无法完成量子任务.

1)CT-CGCD. CT-CGCD 是经典 GAN 模型[126], 目前已经有很多扩展研究, 但不属于量子机器学习模型.

2)QT-CGCD. 使用经典 GAN 模型完成在量子态数据上的任务. 尽管经典生成器无法精确拟合量子数据, 而由于经典判别器也无法对经典生成器生成的经典数据与量子数据进行合适的比较, 因此在这种情况下纳什均衡是可以达到的, 但模型效果较差. 由于上述问题, 在量子机器学习任务中极少使用 CGCD 模型, CGCD 的量子任务一般在于辅助进行物理问题研究, 例如量子态断层扫描[127].

3)QT-QGQD. 最初提出的 GAN 的量子版 Qu-GAN[124] 是 QT-QGQD. Lloyd 等人[74] 证明对于经典任务, CGCD 计算梯度需要$O(N^2)$个经典比特和$O(N^2)$的时间复杂度. QGQD 在计算协方差矩阵时, 对于低秩稀疏矩阵只需要$O(\mathrm{lb}n)$量子比特和$O(poly(\mathrm{lb}n))$个量子门.

对于量子态数据样本, 最初 Hu 等人[128] 提出了一种由量子生成器和量子判别器构成的 QGAN, 并在超导量子电路上进行了第 1 个原理证明演示实验, 生成保真度达到了 98.8%, 为通过实验探索使用 NISQ 设备执行机器学习任务奠定了基础. QGAN 在未来将被更广泛地应用到对量子数据的学习和处理中. 经典 GAN 在被提出后经过多次改进, WGAN 的提出解决了 GAN 的部分缺陷. Chakrabarti 等人[73] 根据 WGAN 的改进方式对 QGAN 进行了改进, 提出了对量子数据的 Wasserstein 半度量, 设计出了 QWGAN, 通过仿真分析得出其有着相较于 QGAN 更为优异的性能.

但是 QWGAN 能否应用于实际的量子计算机和应用于更大的量子系统都是未知的. QWGAN 在被提出时并非使用变分量子电路实现, 由于其使用简单的参数化量子电路实现, 具有与变分量子算法相似的特点, QWGAN 也可以通过变分量子电路实现. 使用变分量子算法实现对经典 GAN 的改进版的量子化将是未来改进 QGAN 模型的一个研究方向.

4) CT-QGQD. 对于经典任务, Huang 等人[59] 利用 Patch 电路和 Batch 电路构成的量子生成器设计了 PATCH-GAN 和 Batch-GAN, 首次在超导量子处理器上实现了手写数字图像生成, 在减少训练参数的同时能保证与经典 GAN 相同的性能.

5) CT-QGCD. 在经典数据样本上, 对于离散型随机变量, Zeng 等人[44] 和 Situ 等人[45] 提出了不同的量子生成器, 结合经典判别器实现了 QGAN. Zhou 等人[19] 对 QGAN 的量子生成器部分结构进行了优化改进, 减少了模型所需参数量, 结合重映射方法实现了一种图像生成方法, 在 Bars 和 Stripes 数据集上的仿真实验表明减少参数量后的模型没有明显的性能损失. 这种方法能够更大程度上利用量子器件, 更适用于 NISQ 设备.

对于连续型随机变量, Romero 等人[46] 提出了一种量子变分生成器 VQG, 其可以用于模拟经典连续概率分布. 特别地, 该模型在使用量子生成器时, 判别器可以为经典的或量子的; 在使用经典生成器时, 可以使用量子判别器. 故该模型同时属于 CT-QGCD, CT-QGQD, CT-CGQD. Li 等人[47] 使用量子生成器减少了模型的参数量, 提出具有新结构的 QGAN-HG 并在药物生成上进行了实验.

Herr 等人[48] 将变分量子算法引入 QWGAN, 设计了使用量子-经典混合生成器和经典判别器构建的 QWGAN-GP, 并将其与 AnoGAN 方法结合, 实现了针对信用卡异常检测的新框架. Tsang 等人[49] 在 QWGAN 的基础上进行了改进, 提出了 PQWGAN, 能够在不降维的前提下生成高分辨率图像, 这使得模型能够学习到更多信息. 由于 PQWGAN 的设计目的是能够在 NISQ 设备上运行, 因此判别器并没有和 QWGAN 一样使用量子判别器, 而是沿用了 WGAN 中的经典神经网络作为判别器. Huang 等人[59] 提出的模型也支持使用经典判别器, 目的是在资源有限的情况下完成大规模学习任务.

6) QT-QGCD. 在经典数据的任务上, 量子算法对经典算法有着极大的潜在优势, 量子优势的发挥极大程度上依赖于经典数据到量子态的加载过程, 然而目前的方法中精确地加载需要 $O(2^n)$ 的门, 有可能抵消量子算法的加速优势. Zoufal 等人[50] 提出使用 QGAN 来学习经典数据的概率分布并将其加载到量子态上, 并且只使用 $O(poly(n))$ 的门, 这促进了量子算法在经典数据上的应用. Agliardi 等人[62] 对多元数据加载中的优化调优问题进行了研究, 进行了效率优化.

7) QT-CGQD. 经典生成器不能精确地拟合量子态数据分布, 无法达到纳什均衡. 因此对于量子任务, 不存在经典生成器-量子判别器构成的模型.

8) CT-CGQD. Niu 等人[107] 提出了 EQ-GAN 架构来克服非凸性和模式崩溃问题, 生成器与判别器都可以是经典的或量子的, 同时引入辅助量子神经网络来加速训练. Kao 等人[104] 设计了 Wasserstein 度量下的 MolGAN 的量子实现 QuMolGAN, 使用变分量子电路实现量子噪声生成器和量子判别器, 药物分子探索的实验表明其生成分子性质和 KL 散度等指标均好于经典 GAN.

总的来说, QGAN 相对于 GAN 能够应用到更多的任务上, 尤其是量子任务. 在经典任务上也相对于经典 GAN 系列展现出潜在的优势. 在量子任务中, QGAN 已经被用于量子态的学习和制备[73]、数据编码[50,62]、纠缠探测[130] 等. 在经典任务上, QGAN 被用于图像生成[59]、药物设计[47,104]、金融领域[50]、异常检测[48,131] 等. 对于模式崩溃[107] 等经典 GAN 难以解决的问题, QGAN 也有良好的前景.

4.4 强化学习

强化学习用于解决智能体通过与环境交互不断学习最优策略以获得最大化奖励的问题. 在智能体与环境交互过程中, 根据获得奖励的大小来调整策略, 以获得最大奖励值为目标, 反复迭代寻找相应最优策略. 在经典机器学习中, 一般使用神经网络作为决策模型. 在量子强化学习 (quantum reinforcement learning, QRL) 中, 可以通过使用变分量子电路代替神经网络作为决策模型的方法实现基于变分量子算法的量子强化学习模型.

一般情况下, 量子强化学习系统流程如下: 首先由量子-经典混合智能体接收来自环境的状态信息, 然后通过变分量子算法决定智能体策略, 最后使用经典的算法对策略进行评估和更新, 如图 13 所示.

表 8 对量子强化学习算法进行了总结.

在强化学习中, 基于值函数的方法是非常常见的方法, 在量子强化学习中, 基于变分量子电路可以实现其量子版本的同时获得性能上的提升. Q 学习 (Q-learning)[132] 是一种经典的无模型、off-policy 强化

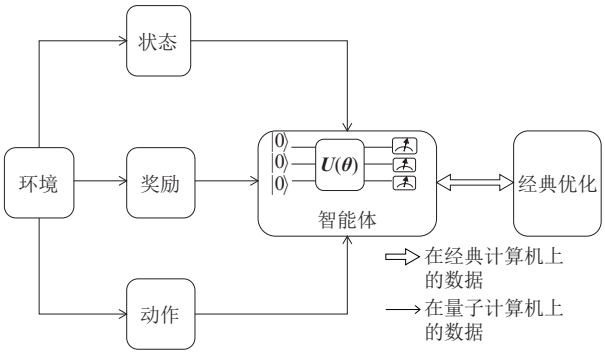


Fig. 13 Illustration of quantum reinforcement learning

图 13 量子强化学习示意图

学习算法, 通过寻找在给定状态下不同行为得到的最大奖赏的方式更新 Q 值, 再通过 Q 值决定智能体下一步进行的策略. 在深度强化学习中, 深度 Q 学习(deep Q-learning, DQN)[133] 是一种主要的算法, 它使用深度神经网络进行 Q 值的函数拟合.

Chen 等人[69] 最先提出并实现了基于变分量子电路的深度 Q 学习(variational quantum deep Q-learning, VQ-DQN)网络, 能够在离散状态空间中完成强化学习任务. 相比于 DQN 算法, VQ-DQN 使用变分量子电路代替经验回放和目标网络来近似 Q 函数. 在 frozen-lake 和 cognitive-radio 环境下的实验表明这种方法具有训练和推理速度快以及内存消耗和模型参数量少等优点, VQ-DQN 相对 DQN 减少了 56.25% 的参数量, 使用基编码的计算复杂度为 $O(n)$, 使用概率幅编码的计算复杂度为 $O(\mathrm{lb}n)$, 而经典 DQN 方法复杂度为 $O(n^2)$, 其中 n 为输入向量维度, 这也意味着基于变分量子电路可以实现更复杂的量子强化学习模型. 同时 VQ-DQN 在量子硬件的实验中也表现良好. Lockwood 等人[51] 提出一种能够在离散和连续状态空间下进行 Q 学习的变分量子电路, 将连续空间的输入转化为对初始层旋转门角度的限制编码, 使用 Z 基测量值来表示 Q 函数值:

$$Q(s,a)=\langle\psi_0|\boldsymbol{U}_{\boldsymbol{\theta}}^{\dagger}(s)\hat{\boldsymbol{Z}}_a\boldsymbol{U}_{\boldsymbol{\theta}}(s)|\psi_0\rangle. \tag{34}$$

各个实验存在其特殊的训练指标, 例如处理时间、加载精度等, 本文仅列出了相同训练指标以对比.

在基于变分量子电路的 Q 学习中, 量子电路的输出范围最好能够与最优 Q 值相匹配. Skolik 等人[105] 将该方法进行改进, 在计算 Q 函数时引入了可训练的权重参数, 使其能够更加灵活地进行匹配:

$$Q(s,a)=\langle\psi_0|\boldsymbol{U}_{\boldsymbol{\theta}}^{\dagger}(s)\hat{\boldsymbol{Z}}_a\boldsymbol{U}_{\boldsymbol{\theta}}(s)|\psi_0\rangle\boldsymbol{w}_{\boldsymbol{Z}_a}. \tag{35}$$

由于动作数量通常比较少, 每个动作对应一个权重参数并不会造成参数量过多和计算复杂度增大的情况.

Lockwood 等人[51] 讨论了 QDDQN 框架, 将基于变分量子电路的 DQN 与经典双 DQN(double DQN, DDQN)[134] 结合. Heimann 等人[60] 将其应用到机器人

Table 8 Quantum Reinforcement Learning Algorithms

表 8 量子强化学习算法

模型	测试环境	环境	量子位	参数量	回合数	回报
VQ-DQN[69]	frozen-lake	模拟	4	28	198	0.9
VQ-DQN(pretrianed)[69]	frozen-lake	量子	4	28	1	0.95
VQ-DQN[69]	cognitive-radio	模拟	4	28	10*	100
VQ-DQN(pretrianed)[69]	cognitive-radio	量子	4	28	1	100
Quantum-DQN[105]	frozen-lake v0	模拟	4	5 层	3 100	1.0
Quantum-DQN[105]	frozen-lake v0	模拟	4	10 层	2 200	1.0
Quantum-DQN[105]	frozen-lake v0	模拟	4	15 层	1 700	1.0
Quantum-DQN[105]	Cart Pole v0(optimal)	模拟	4	62	186	195
Quantum-DQN[105]	Cart Pole v0(sub-optimal)	模拟	4	62	3 000	176
Quantum Actor-critic[52]	Cart Pole	模拟	4	36	6 000	105
QLSTM-DRQN-1[16]	Cart Pole(Full Observable)	模拟	8	150	350*	100*
QLSTM-DRQN-1[16]	Cart Pole(Partially Observable)	模拟	8	146	675*	150*
QLSTM-DRQN-2[16]	Cart Pole(Full Observable)	模拟	8	270	420*	125*
QLSTM-DRQN-2[16]	Cart Pole(Partially Observable)	模拟	8	266	750*	100*
QMARL[53]	Single-Hop Offloading	模拟	4	50	500	−3.0
改进 CTDE QMARL[54]	Smart Factory	模拟	16	54	980	−37.0

注: 带*的数值表示原论文中未给出精确数值, 本文进行估算后得到的数值. 结果为各论文中给出的最优参数的模型. Quantum-DQN 模型中未给出具体参数量, 层数与参数量正相关.

导航任务中. 实验表明该模型能够使用远远小于经典 DDQN 使用的参数量获得相同的奖励，代价是需要更多的步数进行训练.

除了上述方法之外，Chen 等人[16]提出一种基于 QLSTM 的量子深度循环 Q 学习（QLSTM-DRQN）在求解 cart-pole 等基准测试时，尤其是在只能观察到部分环境的测试中，相对于经典的 DRQN 具有更好的稳定性和高的分数.

在强化学习中，基于值函数的方法存在策略退化的问题，还存在不会出现策略退化的基于策略梯度的算法，这些算法的量子版本也可以基于变分量子电路实现并进行改进. 一般通过将变分量子电路代替 DNN 的方式实现，使用策略变分量子电路（policy-VQC）决定智能体的动作. Jerbi 等人[135]最先提出了基于策略梯度的变分量子算法 policy-PQC，其提出的 softmax-PQC 模型中 softmax 策略由变分量子电路计算获得，同时添加了输入缩放参数和可观测值权重参数以提高电路的表达能力和灵活性，其性能与使用深度神经网络相当. 电路使用交替编码酉算子和变分酉算子，编码酉算子的输入主要包括缩放参数 λ 和状态 s，如图 14 所示为使用双量子比特的单层结构. 一般情况下，模型与环境的交互是经典的，即使用经典策略梯度方法.

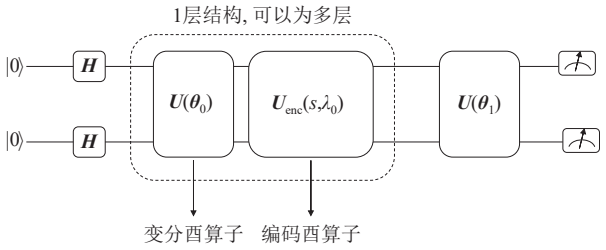


Fig. 14 Illustration of policy-PQC

图 14 policy-PQC 示意图

Jerbi 等人[136]提出量子策略梯度算法，可以加速数值方法的数值和解析梯度估计，该方法是能够在大状态动作空间中完成强化学习任务的全量子方法，在允许适当的量子访问的情况下，这种算法可以在策略参数更新上提供平方级加速. 将量子策略梯度算法与经典策略梯度算法相结合，构建混合策略梯度算法将成为变分量子强化学习的重要研究方向之一.

对于基于策略-价值的方法，演员-评论家（actor-critic）算法的量子版本也被相继提出. 最初 Wu 等人[137]提出分别使用一个变分量子电路作为 actor 和 critic 部分，构造了一个适用于离散状态空间和连续状态空间的框架. Kwak 等人[52]实现了 actor-critic 的量子版本，使用量子 actor 和经典 critic 的组合，如图 15 所示. 在模拟实验学习过程中获得奖励的偏差较高，推测是量子系统内部的不确定性造成的，这使得该模型在参数较多时难以维持良好效果. 如何利用量子系统不确定性或提高相应鲁棒性是量子强化学习的重要挑战之一.

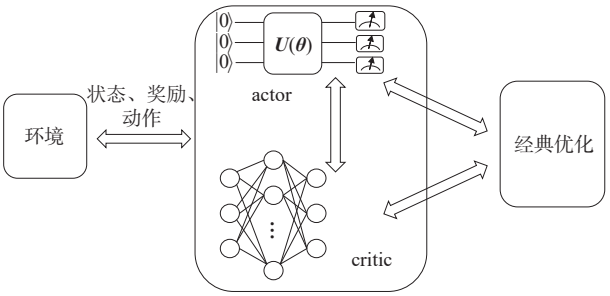


Fig. 15 Quantum actor combines with classical critic

图 15 量子 actor 结合经典 critic

除此之外，Skolik 等人[106]将变分强化学习算法对真实量子硬件上存在的各类型噪声的鲁棒性进行了评估，包括散粒噪声、相干误差以及非相干误差. 同时提出了一种在几乎不影响最终表现的情况下减少变分 Q 学习模型测量次数的方法.

同样地，变分量子电路也可以集成应用到经典多智能体强化学习中. Yun 等人[53]将量子强化学习扩展到量子多智能体强化学习中（quantum multi-agent reinforcement learning, QMARL），提出一种基于变分量子电路的多智能体强化学习方法. 对于经典的多智能体强化学习，智能体之间的相互作用会影响 MARL 训练的收敛，一般采用 CTDE 作为训练范式以解决问题. 尽管可以使用变分量子电路实现 CTDE，但是需要用量子比特表示状态-动作对，所需量子比特的数量随智能体的数量的增加而增加，误差也会随之增大，这并不适用于 NISQ 设备. 因此设计了一种量子态编码的方法将状态-动作对的维度降低，从而减少使用的量子比特数. 经过模拟，这种 QMARL 比经典 MARL 获得的总奖励提高了 57.7%. 一些量子或经典降维算法可以与 QMARL 进行结合以获得更好的性能. Yun 等人[54]对 CTDE QMARL 框架进行了改进，使用参数共享策略减少了参数量，并对 VQC 和编码方式进行了改进.

由于量子硬件的限制，目前提出的绝大多数量子强化学习算法都停留在理论层次上，仅仅通过经典计算机的模拟对可行性进行了说明，物理实验方案的实现十分匮乏.

4.5 量子电路结构搜索

在NISQ时代,深度浅、参数量小的变分量子算法更有可能应用于NISQ计算机上并体现出其量子优势,因此很多研究关注于寻找给定任务下最优变分量子电路结构.在量子机器学习中,当前量子电路结构搜索集中于完成分类任务.

量子结构搜索(quantum architecture search, QAS)是由神经结构搜索(neural architecture search, NAS)改进的一种用于寻找最优拟设结构的方法.对于一个机器学习任务,QAS自动寻找一个近似最优的近似解,即最优电路结构,以平衡增加更多噪声量子门所带来的优劣势.Zhang等人[55]提出使用神经网络预测器引导的QAS,能够端到端地自动设计变分量子电路,性能远远好于随机搜索设计的量子电路,有助于使用更少的电路评估找到更优的解决方案.Du等人[108]提出一种基于噪声感知方法来进行拟设结构搜索的QAS算法,将噪声抑制和增强可训练性统一为一个学习问题,具有对噪声鲁棒性良好、不需要辅助量子比特资源、兼容于所有量子平台、可以与其他噪声抑制和解决贫瘠高原问题的算法无缝集成等诸多优点.对于相同的迭代次数T,该QAS算法与传统变分量子算法具有相同的时间复杂度$O(dT+K)$,空间复杂度$O(Td+TNL+Kd)$也仅为后者的T倍,K为采样步骤中平衡效率和性能的超参数,一般取$O(poly(QNL))$,其中d, N, L, Q分别表示可训练参数量、量子比特数、电路深度和量子门类数.

Wang等人[12]提出一种将变分量子电路与噪声自适应的量子位映射相结合的框架QuantumNAS,考虑了硬件拓扑,其性能显著优于无噪声搜索、人工设计、随机搜索和现有的噪声自适应量子比特映射基准,通过剪枝的方法减少了参数量,提高了分类准确率.对于分类任务,经过在MNIST, Fashion, Vowel等数据集上,使用多台量子计算机进行的实验表明,QuantumNAS是第1个在真实量子计算机上超过95%的二分类、85%的四分类和32%的十分类准确率的模型.QuantumNAS模型是分类任务在真实量子环境下的SOTA模型.

Lu等人[56]提出马尔可夫量子神经进化算法MQNE,将量子电路和有向图之间进行了一一对应映射,能够将寻找最合适门序列的问题归结为在图中寻找路径的问题,并将其转化为马尔可夫过程,能够为各机器学习任务找到近似最优的拟设结构.在MNIST等数据集和SPT量子态分类上的实验表明,MQNE在保持分类精度的前提下,找到了深度较浅的Ansatz参数化量子电路.

强化学习的方法可以被用于优化量子电路的结构.参数化量子电路结构可以构建成一系列由可训练的策略所生成的动作,根据变分量子电路的损失函数设计最终奖励函数,奖励函数引导策略更新,使回报期望最大化,从而选择出一系列动作并构建电路结构.Ostaszewski等人[138]将负反馈课程学习引入变分量子电路,提出一种基于深度强化学习的变分量子本征求解器通用优化方法,将架构应用于寻找氢化锂分子基态能量基准问题,能够在保持低深度电路的情况下获得栅极效率上最先进的结果.这种方法适用于所有基于变分量子电路的算法,故同样可以被应用于解决机器学习任务的变分量子电路中.

噪声会显著影响优化器的性能和最终结果,同样地,对量子电路结构搜索算法的影响也是多方面的[139],故对噪声鲁棒性强的电路结构搜索算法也是一个关键的研究方向.Patel等人[78]设计了基于课程强化学习的量子电路结构搜索算法(curriculum reinforcement learning for quantum architecture search, CRLQAS).该算法引入一种基于深度感知的3维编码,状态由3维网格描述,通过量子位索引、深度和门类型将量子电路编码到网格上.状态和奖励通过多层感知器进行处理,智能体根据策略选择出相应的动作,即对量子门的操作.经典优化器对修改后的电路完成训练后提供奖励,继续选择下一个动作.该方法还设计了一种优化机制,通过设计非法行为来避免连续应用类似的量子门,减小了搜索空间,同时加入了随机停止机制以引导智能体寻找更浅的电路.此外,还设计了SPAS和Adam算法的变体,在应用该搜索算法时能够获得更快的收敛速度和更好的鲁棒性.该算法在量子化学任务数值实验上表现良好,具有应用于量子机器学习任务的极大潜力.

在以上的讨论中,绝大部份工作都集中于在门级使用变分量子电路完成机器学习任务.然而,较少的参数数量制约了其灵活性和表达能力,Liang等人[140]提出变分量子脉冲学习(VQP),使用更底层的具有额外的控制参数的脉冲作为基本成分来构建变分量子电路的拟设结构,直接训练量子脉冲来完成图像二分类任务,并在噪声环境下获得了稳定的结果,相较于基于门的变分量子电路具有延迟优势.

Meitei等人[141]提出了Ctrl-VQE,通过调整脉冲状态来进行状态准备,大大缩短了状态制备时间.这种方法依赖于小量子系统的经典模拟,鲁棒性差,只使

用单量子比特脉冲，并且只能应用于 VQE 任务而不能完成量子机器学习任务. 进一步地，Liang 等人[142]提出一种中级变分原生脉冲拟设 NAPA，在该算法中提出渐进式脉冲学习，可以自适应地生成脉冲拟设. NAPA 算法经历多次脉冲添加步骤，每个步骤增添一组新的原生脉冲，前面的脉冲参数不变，新脉冲参数通过优化器多次迭代更新得到. 递进添加的停止条件是不能通过添加新步骤来提高精度. NAPA 中所用到的参数化脉冲具有更高的灵活性，能够在相同的电路延迟内集成更多的参数. NAPA 相较 Ctrl-VQE 增加了双量子比特脉冲，与门级电路 SOTA 模型 QuantumNAS 在不同规模量子比特下的对比实验表明，NAPA 在小分子 VQE 任务上的平均延迟减少了 86%，准确率高达 99.895%，在大分子 VQE 任务上平均准确率达到了 97.27%. 在组合优化问题最大割任务上提供了可行性，在机器学习任务上也具有极大的潜力. 探索 NAPA 在不同类型的变分量子机器学习算法中的应用是重要的研究方向之一.

一些相关的研究在其他领域的任务上表现良好，其中以变分量子特征求解器居多[143]，未来有望将这些算法用于其他量子机器学习任务的求解上.

表 9 对电路结构搜索算法进行了总结.

Table 9 Quantum Architecture Searching Algorithms

表 9 量子架构搜索算法

模型	数据集	任务	环境	量子位	最优结构参数量	准确率/%
QuantumNAS[12]	MNIST	二分类	量子	5	22	95
QuantumNAS[12]	MNIST	四分类	量子	5	22	75
QuantumNAS[12]	MNIST	十分类	量子	15		32.5
QuantumNAS[12]	Fashion-2	二分类	量子	5	22	92
QuantumNAS[12]	Fashion-4	四分类	量子	5	36	85
MQNE[56]	MNIST	二分类	模拟	9	106	97
MQNE[56]	Cancer	二分类	模拟	7	68	94.6
MQNE[56]	SPT	量子态分类	模拟	8	46	100
QAS[55]	Fashion-MNIST	二分类	模拟	10		92.4
QAS[108]	合成数据集（无噪声）[144]	二分类	模拟	3		>90
QAS[108]	合成数据集（有噪声）[144]	二分类	模拟	3		100
CRLQAS[78]		VQE	模拟			
NAPA[142]		VQE 最大割	量子			

5 挑战与展望

通过理论的分析和经典计算机的模拟，很多基于变分量子算法的量子机器学习模型在解决部分特定任务时有显著的性能提升. 然而在效率提高的同时，变分量子算法在机器学习领域的应用还面临着许多挑战.

5.1 噪 声

模型的脆弱性指模型对抗噪声干扰的能力，噪声分为自然噪声和对抗性噪声 2 种，NISQ 设备中产生的扰动属于自然噪声，在 QGAN 中生成器部分产生的噪声属于对抗性噪声，我们在设计变分量子电路时希望能够降低脆弱性和提高抗干扰能力. 变分量子算法设计的目的是在 NISQ 设备上进行应用，NISQ 设备的特点是存在随电路深度增大而增大的噪声，Stilck 等人[145]的研究表明当噪声水平超过一定阈值时，量子计算机的性能将低于经典计算机. 故而在量子机器学习模拟实验中也需要对噪声进行模拟，常见的包括结合纯态分解和低维基投影的含噪声量子电路模拟方法[146]等.

CRLQAS[78]和 NAPA[142]模型未进行量子机器学习任务相关实验，并非是为特定任务设计，其具有完成机器学习任务的潜力，需要进一步的研究.

变分量子算法最适合应用于 NISQ 设备的原因之一是具有良好的抗噪声干扰能力，Sharma 等人[147]认为变分量子算法的损失函数的全局极小值可能不受噪声的影响.

噪声对变分量子算法的限制在于可训练性上，是导致贫瘠高原[114]的直接原因之一. Wang 等人[139]

证明了对于局部泡利噪声, 如果拟设深度随量子比特数线性增长, 那么梯度会随量子比特数指数级消失, 这种噪声诱导的贫瘠高原(noise-induced barren plateau, NIBP)的成因被证明是噪声导致的量子态向最大混合态偏移, 同时噪声损失函数逐渐向最大混合态对应值靠近. NIBP 问题不同于无噪声场景下的贫瘠高原现象, 随着问题规模的增大, 前者在损失函数的每个点都会产生梯度消失, 而非后者的概率性梯度消失, 无噪声场景下避免贫瘠高原的方法对于 NIBP 问题几乎没有作用, 例如分层训练[148]、关联参数[149]和其他一些策略[123]. 目前缓解 NIBP 现象的方法只有 2 种: 降低量子硬件噪声水平和降低电路深度, 而前者依赖于物理理论和工业水平. 使用错误缓解策略是否能够缓解 NIBP 现象是一个尚未解决的问题, Endo 等人[150]通过错误缓解从含噪声的结果中提取无噪声的结果, 而 Wang 等人[139]认为只包括后处理噪声电路的错误缓解技术几乎不可能缓解 NIBP 现象. 因此最关键的研究方向依旧是降低变分量子电路深度的方法, 但是这些方法都会限制我们对电路的设计从而影响模型的性能.

针对特定的问题和硬件来设计拟设结构也能够抑制噪声[151], 不过我们更希望能够找到更通用的噪声缓解策略. 故对于量子噪声的理论研究还需要进一步探索, Weber 等人[152]为量子分类器可以容忍的噪声提供了严格的鲁棒性条件. Stilck 等人[145]证明了噪声过大的场景会导致变分量子算法相对经典算法不具有量子优势.

此外, 噪声也具有一些潜在的好处: Du 等人[153]证明了能够利用去极化噪声来保护分类算法免受类似的对抗样本的攻击, 对于量子数据和经典数据均适用. Liu 等人[154]发现测量引起的随机性噪声可以帮助优化器跳出鞍点. 目前尚不清楚可容忍的噪声量和有害的噪声量之间的界限[155], 对不同类型的学习算法和噪声源很难给出通用的解释.

5.2 可训练性

在变分量子算法中, 参数化量子电路的可训练性是十分重要的. Mcclean 等人[114]发现在一些训练过程中损失函数的梯度的方差随着系统平均规模的增大而指数级递减, 故解决噪声以及确定最小化的方向都需要指数级的精度. 这在变分量子算法中被称为贫瘠高原现象, 对经典优化部分造成了一定的限制, 可能导致变分量子算法带来的量子优势消失, 可训练性成为了发展瓶颈之一.

在对贫瘠高原现象的探究中发现了一些与其成因相关的因素, 包括文献 [139] 提到的噪声等, 针对这些原因也提出了一些缓解贫瘠高原的方法.

贫瘠高原和 Ansatz 电路参数随机初始化有关, 解决贫瘠高原现象的一种方法是设计初始化参数的策略. Zhang 等人[156]证明对于具有一定方差的满足高斯分布的初始化参数, 随着量子比特数和电路深度的增加, 梯度消失速度至多是多项式的, 这种方法在量子模拟和量子化学的模拟实验中均优于均匀初始化参数, 有望用于量子机器学习模型参数初始化策略. 编码过程可以看作是对部分参数的初始化, 寻找一种合适的编码策略有助于缓解贫瘠高原[157].

当非结构化的 Ansatz 电路的深度随量子比特数多项式地增长时, 该电路满足西 2-设计, 这种参数化量子电路会最大化其可表达性[158]. Holmes 等人[159]证明了高表达能力的电路会导致更平坦的损失函数, 通过适当降低电路可表达性的方式能够缓解贫瘠高原现象, 提出了几种方法, 例如限制参数化量子门(即旋转门)的旋转角度或者关联参数等. Bilkis 等人[151]提出一种变结构的方法 VAns(variable Ansatz), 在优化过程中不断增加和删除量子门, 以此保持 Ansatz 维持在一个较浅的深度附近, 从而缓解可训练性和噪声的问题. 使用这种方法构建了量子自编码器, 成功缓解了贫瘠高原现象, 同时表现出对噪声良好的鲁棒性.

过多的量子纠缠态的引入也有可能导致贫瘠高原现象[160], 这是可观测量子比特和隐藏层中的量子比特纠缠导致的, 在变分量子电路中 CNOT 门会引入纠缠, 类似的门的大量使用可能是贫瘠高原出现的原因. 这种贫瘠高原通过控制量子比特之间的纠缠能够进行缓解.

Cerezo 等人[123]推测使用全局损失函数是贫瘠高原现象出现的原因之一, 因此提出了一种使用多个局部损失函数C_{L}代替全局损失函数C_{G}的方法来延缓贫瘠高原的出现, 并以此设计了使用局部损失函数的量子自编码器.

$$\begin{aligned} &C_{\mathrm{G}}=\mathrm{tr}\left[\boldsymbol{O}_{\mathrm{G}}\boldsymbol{V}(\boldsymbol{\theta})|\psi_0\rangle\langle\psi_0|\boldsymbol{V}(\boldsymbol{\theta})^{\dagger}\right],\\ &\boldsymbol{O}_{\mathrm{G}}=\boldsymbol{1}-|0\rangle\langle 0|. \end{aligned} \tag{36}$$

$$\begin{aligned} &C_{\mathrm{L}}=\mathrm{tr}\left[\boldsymbol{O}_{\mathrm{L}}\boldsymbol{V}(\boldsymbol{\theta})|0\rangle\langle 0|\boldsymbol{V}(\boldsymbol{\theta})^{\dagger}\right],\\ &\boldsymbol{O}_{\mathrm{L}}=1-\frac{1}{n}\sum_{i=1}^{n}|0\rangle\langle 0|_i\otimes\boldsymbol{1}_i. \end{aligned} \tag{37}$$

这些局部损失函数被证明在训练中梯度消失是多项式级别的而非指数级, 因此在大规模数据的训练上有着极大的潜力. 然而在近期的研究中, Kashif

等人[161]发现在分类问题上局部损失函数并没有体现出优势，在多分类问题上甚至远不如全局损失函数. 未来的研究将会对比全局损失函数和局部损失函数在其他机器学习任务上的性能，例如回归任务或无监督学习任务等.

在机器学习领域中，除了量子卷积神经网络由于其特殊的 Ansatz 结构没有表现出贫瘠高原的现象以外[113]，其他基于变分量子电路的量子机器学习模型仍然因可训练性问题被制约量子优势. 对于变分算法整体的研究，可训练性的问题是其在所有应用领域发展的重大挑战.

5.3 可表达性

变分量子算法另一个重要的性质是可表达性，这决定了参数化量子电路能够产生的量子态种类的丰富程度. 可表达性利用 Haar 度量定义：

$$\|\boldsymbol{A}^{(t)}\|_{\mathrm{HS}}^2=\left\|\int_{\mathrm{Haar}}(|\psi\rangle\langle\psi|)^{\otimes t}\mathrm{d}\psi-\int_{\theta}(|\psi_{\theta}\rangle\langle\psi_{\theta}|)^{\otimes t}\mathrm{d}\psi_{\theta}\right\|_{\mathrm{HS}}^2, \tag{38}$$

其中 $\|\boldsymbol{A}\|_{\mathrm{HS}}^2=\mathrm{tr}(\boldsymbol{A}^{\dagger}\boldsymbol{A})$，表示希尔伯特-施密特（Hilbert-Schmidt）范数. 对于一个 Ansatz 电路 $\boldsymbol{U}$，$A_U^{(t)}$越小表示其可表达性越强，当$A_U^{(t)}=0$时可表达性达到最高，能够近似希尔伯特空间中任意量子态.

由于量子机器学习中的变分量子算法经常重复多次使用多量子门，例如 CNOT 等，根据电路的可扩展性和易计算性，Sim 等人[162]提出了使用 MW 纠缠测度来量化和评估这种强纠缠电路的能力，还有一些其他纠缠能力度量方法[163].

5.4 可达性

可达性指的是变分量子算法应当具有能够收敛到损失函数最小值的性质. Akshay 等人[17]给出了可达性的目标函数形式：

$$f_R=\min_{\psi\in\mathcal{H}}\langle\psi|\boldsymbol{O}|\psi\rangle-\min_{\boldsymbol{\theta}}\langle\Psi(\boldsymbol{\theta})|\boldsymbol{O}|\Psi(\boldsymbol{\theta})\rangle. \tag{39}$$

即用在希尔伯特空间中所有量子态计算的损失函数中的最小值与参数化量子电路能够产生的量子态计算损失函数中的最小值做差. 当$f_R=0$时，该变分量子算法能够收敛到函数最小值.

对变分量子算法可达性的研究基本都是围绕 QAOA 展开的[17]，对于量子机器学习模型的讨论较少，将 QAOA 的研究结果转化到量子机器学习中是非常有必要的. 除此之外，近期 Anand 等人[18]提出了结合量子最优控制理论的可达性概念，在 VQE 框架下进行了讨论.

5.5 量子存储

绝大多数量子算法都需要量子随机存储器（quantum random access machine, QRAM）才能体现出其量子优势[164]，基于门电路的 QRAM 的实现更是变分量子电路应用于量子机器学习任务上所必须解决的问题[165-166]，然而目前距离实现满足量子算法所需的 QRAM 还比较遥远. 参数化量子电路被用于实现 QRAM，量子机器学习算法也能够用于研究更优良的 QRAM 结构. 精度更高、规模更大的 QRAM 亟待实现. 此外，也存在一些不需要 QRAM 即可实现的算法，Henderson 等人[110]提出的量子卷积神经网络只需要在单个数据点上运行量子电路，不涉及大规模数据的存储.

5.6 I/O 瓶颈

在使用量子机器学习模型完成经典任务时，对于将经典数据有效地编码到量子态上是比较困难的，计算复杂度往往与维度有关，在高维数据前算法产生的量子优势会被抵消，故而编码问题也被称作读入瓶颈. 目前的缓解方法主要包括采取量子-经典混合模型，先通过经典算法对数据进行降维再应用量子模型[30,38]. 然而这种情况也存在问题，首先很难对混合模型进行批量训练，大规模数据训练难以完成. 其次就是难以确定在任务中起主要作用的是量子模型还是经典神经网络. 除此之外，在经典机器学习任务上，将量子态通过测量提取经典数据也是十分困难的.

5.7 量子机器学习挑战

可训练性和可达性的问题在所有的变分量子算法中是共通的. 接下来对比经典机器学习概念，列举一些基于变分量子电路的量子机器学习算法所面临的问题.

5.7.1 数据集

在目前对量子机器学习的研究中，设计的模型大多是在经典数据集上进行基准测试，从而体现出其对经典模型所具有的量子优势. 虽然目前已经提出了一些编码方案，但是如何最有效地进行编码仍然是未知的. 量子机器学习算法可能在量子数据集上更能体现出量子优势，然而目前提出的量子数据集很少，需要提出更多的、更易于制备的量子态数据集来为量子机器学习模型提供基准测试.

5.7.2 反向传播

在经典机器学习中，反向传播算法是更新参数的重要手段，大大提高了训练复杂模型的效率. 通过重复应用链式法则计算损失函数对各个参数的梯度，该过程需要用到向前传播过程中各子表达式的结果. 由于计算过程中不能对量子态进行访问，因此量子

反向传播算法难以实现. 多数情况下, 我们使用参数平移规则来获得梯度[84], 因此可以避开对中间量子态的访问. 然而当参数量增加、电路深度加深时, 使用参数平移规则计算梯度所花费的时间逐渐增加, 而反向传播算法的计算时间远小于参数平移规则的计算时间, 计算时间的减少能够提高更复杂模型的可训练性. 尽管在经典定义上的反向传播算法难以实现, 一些研究设计了类似反向传播算法[167], 能够有效提高训练效率.

5.7.3 非线性运算

经典机器学习中神经网络的关键是引入了非线性运算, 如果参数化量子电路也能够运用非线性, 那么在参数化量子电路中可以实现经典神经元, 用于代替一些经典神经网络完成任务. 量子门是酉变换运算, 酉门组合而成的电路整体运算是线性的, 在其中实现 Sigmoid 函数等非线性激活函数较为困难, 甚至能否量子化都是不确定的. 可以采用将变分量子电路与经典神经网络结合的方法, 但是如何衡量在混合模型中量子神经网络和经典神经网络的作用比例仍然是一个问题.

6 结 论

本综述介绍了变分量子算法的基本概念、基于变分量子电路实现的量子机器学习模型, 以及变分量子算法目前遇到的诸多挑战. 变分量子算法在分类、对抗学习、自编码器、强化学习等多种机器学习任务上已经体现出其优势或是被证明具有发展潜力, 将量子模型与经典优化相结合的方式更适用于近期 NISQ 量子设备. 然而, 变分量子算法在量子机器学习领域的发展也面临诸多挑战, 算法缺少严格的理论分析, 模型缺乏在真实量子环境下的实验验证、算法规模的可扩展性不足等, 这些瓶颈制约着变分量子算法在机器学习领域的实际应用. 随着量子硬件水平的发展, 变分量子算法在量子机器学习领域的优势将逐渐展现出来, 但是在近期相关研究还是集中于如何充分利用 NISQ 量子设备的潜力, 将经典机器学习方法, 尤其是经典神经网络与变分量子算法相结合的方式将成为未来重要的发展方向. 强化学习领域中的变分量子算法尽管起步较晚, 但是发展潜力巨大, 将是近期主要研究方向. 在未来, 随着量子硬件水平的不断发展和变分量子算法理论的进步, 变分量子算法将从机器学习领域实用化, 扩展到工程中成为智能系统的一个重要组件也有望实现.

作者贡献声明: 于瑞祺提出论文整体思路和框架并撰写论文; 张鑫云提出论文思路; 任爽提出指导意见并修改论文.

参 考 文 献

[1] Hilbert M, López P. The world's technological capacity to store, communicate, and compute information[J]. Science, 2011, 332(6025): 60−65

[2] Arute F, Arya K, Babbush R, et al. Quantum supremacy using a programmable superconducting processor[J]. Nature, 2019, 574(7779): 505−510

[3] Huang Cupjin, Zhang Fang, Newman M, et al. Classical simulation of quantum supremacy circuits[J]. arXiv preprint, arXiv: 2005.06787, 2020

[4] Pan Feng, Zhang Pan. Simulating the Sycamore quantum supremacy circuits[J]. arXiv preprint, arXiv: 2103.03074, 2021

[5] Zhu Qingling, Cao Sirui, Chen Fusheng, et al. Quantum computational advantage via 60-qubit 24-cycle random circuit sampling[J]. Science Bulletin, 2022, 67(3): 240−245

[6] Feng Congcong, Zhao Bo, Zhou Xin, et al. An enhanced quantum *k*-nearest neighbor classification algorithm based on polar distance[J]. Entropy, 2023, 25(1): 127

[7] Li Jing, Gao Fei, Lin Song, et al. Quantum *k*-fold cross-validation for nearest neighbor classification algorithm[J]. Physica A: Statistical Mechanics and Its Applications, 2023, 611: 128435

[8] Cerezo M, Sharma K, Arrasmith A, et al. Variational quantum state eigensolver[J]. NPJ Quantum Information, 2022, 8(1): 113

[9] Zhou Zeqiao, Du Yuxuan, Tian Xinmei, et al. Qaoa-in-qaoa: Solving large-scale maxcut problems on small quantum machines[J]. Physical Review Applied, 2023, 19(2): 024027

[10] Huang Rui, Tan Xiaoqing, Xu Qingshan. Variational quantum tensor networks classifiers[J]. Neurocomputing, 2021, 452: 89−98

[11] Cong I, Choi S, Lukin M D. Quantum convolutional neural networks[J]. Nature Physics, 2019, 15(12): 1273−1278

[12] Wang Hanrui, Ding Yongshan, Gu Jiaqi, et al. QuantumNAS: Noise-adaptive search for robust quantum circuits[C]//Proc of the 28th Annual Int Symp on High-Performance Computer Architecture. Piscataway, NJ: IEEE, 2022: 692−708

[13] Benedetti M, Garcia-Pintos D, Perdomo O, et al. A generative modeling approach for benchmarking and training shallow quantum circuits[J]. NPJ Quantum Information, 2019, 5(1): 45

[14] Zoufal C, Lucchi A, Woerner S. Variational quantum Boltzmann machines[J]. Quantum Machine Intelligence, 2021, 3(1): 7

[15] Wu S R, Li C T, Cheng H C. Efficient data loading with quantum autoencoder[C/OL]//Proc of the 48th Int Conf on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE, 2023[2023-09-14]. https://ieeexplore.ieee.org /abstract/document/10096496

[16] Chen Guoming, Chen Qiang, Long Shun, et al. Quantum convolutional neural network for image classification[J]. Pattern

Analysis and Applications, 2023, 26(2): 655−667

[17] Akshay V, Philathong H, Morales M E, et al. Reachability deficits in quantum approximate optimization[J]. Physical Review Letters, 2020, 124(9): 090504

[18] Anand A, Alperin-Lea S, Choquette A, et al. Exploring the role of parameters in variational quantum algorithms[J]. arXiv preprint, arXiv: 2209.14405, 2022

[19] Zhou Nanrun, Zhang Tianfeng, Xie Xinwen, et al. Hybrid quantum classical generative adversarial networks for image generation via learning discrete distribution[J]. Signal Processing: Image Communication, 2023, 110: 116891

[20] Romero J, Babbush R, Mcclean J R, et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz[J]. Quantum Science and Technology, 2018, 4(1): 014008

[21] Bang J, Lim J, Kim M S, et al. Quantum learning machine[J]. arXiv preprint, arXiv: 0803.2976, 2008

[22] Cerezo M, Arrasmith A, Babbush R, et al. Variational quantum algorithms[J]. Nature Reviews Physics, 2021, 3(9): 625−644

[23] Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces[J]. Physical Review Letters, 2019, 122(4): 040504

[24] Lloyd S, Schuld M, Ijaz A, et al. Quantum embeddings for machine learning[J]. arXiv preprint, arXiv: 2001.03622, 2020

[25] Schuld M. Supervised quantum machine learning models are kernel methods[J]. arXiv preprint, arXiv: 2101.11020, 2021

[26] Grover L, Rudolph T. Creating superpositions that correspond to efficiently integrable probability distributions[J]. arXiv preprint, quant-ph/0208112, 2002

[27] Kitaev A, Webb W A. Wavefunction preparation and resampling using a quantum computer[J]. arXiv preprint, arXiv: 0801.0342, 2008

[28] Lloyd S. Universal quantum simulators[J]. Science, 1996, 273 (5278): 1073−1078

[29] Kandala A, Mezzacapo A, Temme K, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets[J]. Nature, 2017, 549(7671): 242−246

[30] Chen S Y C, Huang C M, Hsing C W, et al. Hybrid quantum-classical classifier based on tensor network and variational quantum circuit[J]. arXiv preprint, arXiv: 2011.14651, 2020

[31] Gong Lihua, Xing Lingzhi, Liu Sihang, et al. Born machine model based on matrix product state quantum circuit[J]. Physica A: Statistical Mechanics and its Applications, 2022, 593: 126907

[32] Grant E, Benedetti M, Cao Shuxiang, et al. Hierarchical quantum classifiers[J]. NPJ Quantum Information, 2018, 4(1): 65

[33] Qi Jun, Yang Chaohan, Chen Pinyu. QTN-VQC: An end-to-end learning framework for quantum neural networks[J]. Physica Scripta, 2023, 99(1): 015111

[34] Li Guangxi, Song Zhixin, Wang Xin. VSQL: Variational shadow quantum learning for classification[C]//Proc of the 35th AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2021: 8357−8365

[35] Grimsley H R, Economou S E, Barnes E, et al. An adaptive variational algorithm for exact molecular simulations on a quantum computer[J]. Nature Communications, 2019, 10(1): 3007

[36] Bittel L, Kliesch M. Training variational quantum algorithms is NP-hard[J]. Physical Review Letters, 2021, 127(12): 120502

[37] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint, arXiv: 1412.6980, 2014

[38] Liu Junhua, Lim K H, Wood K L, et al. Hybrid quantum-classical convolutional neural networks[J]. Science China Physics, Mechanics & Astronomy, 2021, 64(9): 290311

[39] Hur T, Kim L, Park D K. Quantum convolutional neural network for classical data classification[J]. Quantum Machine Intelligence, 2022, 4(1): 3

[40] Liu Jinguo, Wang Lei. Differentiable learning of quantum circuit born machines[J]. Physical Review A, 2018, 98(6): 062324

[41] Coyle B, Henderson M, Le J C J, et al. Quantum versus classical generative modelling in finance[J]. Quantum Science and Technology, 2021, 6(2): 024013

[42] Leyton-Ortega V, Perdomo-Ortiz A, Perdomo O. Robust implementation of generative modeling with parametrized quantum circuits[J]. Quantum Machine Intelligence, 2021, 3(1): 17

[43] Li Junde, Ghosh S. Scalable variational quantum circuits for autoencoder-based drug discovery[C]//Proc of the 25th Design, Automation and Test in Europe Conf and Exhibition. Piscataway, NJ: IEEE, 2022: 340−345

[44] Zeng Jinfeng, Wu Yufeng, Liu Jinguo, et al. Learning and inference on generative adversarial quantum circuits[J]. Physical Review A, 2019, 99(5): 052306

[45] Situ Haozhen, He Zhimin, Wang Yuyi et al. Quantum generative adversarial network for generating discrete distribution[J]. Information Sciences, 2020, 538: 193−208

[46] Romero J, Aspuru-Guzik A. Variational quantum generators: Generative adversarial quantum machine learning for continuous distributions[J]. Advanced Quantum Technologies, 2021, 4(1): 2000003

[47] Li Junde, Topaloglu R O, Ghosh S. Quantum generative models for small molecule drug discovery[J]. IEEE Transactions on Quantum Engineering, 2021, 2: 3103308

[48] Herr D, Obert B, Rosenkranz M. Anomaly detection with variational quantum generative adversarial networks[J]. Quantum Science and Technology, 2021, 6(4): 045004

[49] Tsang S L, West M T, Erfani S M, et al. Hybrid quantum-classical generative adversarial network for high resolution image generation[J]. IEEE Transactions on Quantum Engineering, 2023, 4: 3102419

[50] Zoufal C, Lucchi A, Woerner S. Quantum generative adversarial networks for learning and loading random distributions[J]. NPJ Quantum Information, 2019, 5(1): 103

[51] Lockwood O, Si Mei. Reinforcement learning with quantum variational circuit[C]//Proc of the 16th AAAI Conf on Artificial Intelligence and Interactive Digital Entertainment. Menlo Park, CA: AAAI, 2020: 245−251

[52] Kwak Y, Yun W J, Jung S, et al. Introduction to quantum

reinforcement learning: Theory and pennylane-based implementation[C]//Proc of the 12th Int Conf on Information and Communication Technology Convergence. Piscataway, NJ: IEEE, 2021: 416−420

[53] Yun W J, Kwak Y, Kim J P, et al. Quantum multiagent reinforcement learning via variational quantum circuit design[C]//Proc of the 42nd Int Conf on Distributed Computing Systems. Piscataway, NJ: IEEE, 2022: 1332−1335

[54] Yun W J, Kim J P, Jung S, et al. Quantum multi-agent actor-critic neural networks for internet-connected multirobot coordination in smart factory management[J]. IEEE Internet of Things Journal, 2023, 10(11): 9942−9952

[55] Zhang Shixin, Hsieh Changyu, Zhang Shengyu, et al. Neural predictor based quantum architecture search[J]. Machine Learning: Science and Technology, 2021, 2(4): 045027

[56] Lu Zhide, Shen Peixin, Deng Dongling. Markovian quantum neuroevolution for machine learning[J]. Physical Review Applied, 2021, 16(4): 044039

[57] Robbins H, Monro S. A stochastic approximation method[J]. The Annals of Mathematical Statistics, 1951, 22(3): 400−407

[58] Arthur D. A hybrid quantum-classical neural network architecture for binary classification[J]. arXiv preprint, arXiv: 2201.01820, 2022

[59] Huang Heliang, Du Yuxuan, Gong Ming, et al. Experimental quantum generative adversarial networks for image generation[J]. Physical Review Applied, 2021, 16(2): 024051

[60] Heimann D, Hohenfeld H, Wiebe F, et al. Quantum deep reinforcement learning for robot navigation tasks[J]. arXiv preprint, arXiv: 2202.12180, 2022

[61] Reddi S J, Kale S, Kumar S. On the convergence of adam and beyond[J]. arXiv preprint, arXiv: 1904.09237, 2019

[62] Agliardi G, Prati E. Optimal tuning of quantum generative adversarial networks for multivariate distribution loading[J]. Quantum Reports, 2022, 4(1): 75−105

[63] Nocedal J, Wright S J. Numerical Optimization[M]. New York: Springer, 2006

[64] Zhu Ciyou, Byrd R H, Lu Peihuang, et al. Algorithm 778: LBFGS-B: Fortran subroutines for large-scale bound-constrained optimization[J]. ACM Transactions on mathematical software, 1997, 23(4): 550−560

[65] Romero J, Olson J P, Aspuru-Guzik A. Quantum autoencoders for efficient compression of quantum data[J]. Quantum Science and Technology, 2017, 2(4): 045001

[66] Bravo-Prieto C. Quantum autoencoders with enhanced data encoding[J]. Machine Learning: Science and Technology, 2021, 2(3): 035028

[67] Sutskever I, Martens J, Dahl G, et al. On the importance of initialization and momentum in deep learning[C]//Proc of the 30th Int Conf on Machine Learning. New York: PMLR, 2013: 1139−1147

[68] Tieleman T. Lecture 6.5-RMSProp: Divide the gradient by a running average of its recent magnitude[J]. COURSERA: Neural Networks for Machine Learning, 2012, 4(2): 26−31

[69] Chen S Y C, Yang C H H, Qi Jun, et al. Variational quantum circuits for deep reinforcement learning[J]. IEEE Access, 2020, 8: 141007−141024

[70] Oh S, Choi J, Kim J. A tutorial on quantum convolutional neural networks (QCNN)[C]//Proc of the 11th Int Conf on Information and Communication Technology Convergence (ICTC). Piscataway, NJ: IEEE, 2020: 236−239

[71] Wei Shijie, Chen Yanhu, Zhou Zengrong, et al. A quantum convolutional neural network on NISQ devices[J]. AAPPS Bulletin, 2022, 32(1): 2

[72] Shingu Y, Seki Y, Watabe S, et al. Boltzmann machine learning with a variational quantum algorithm[J]. Physical Review A, 2021, 104(3): 032413

[73] Chakrabarti S, Huang Yiming, Li Tongyang, et al. Quantum wasserstein generative adversarial networks[C]//Proc of the 33rd Conf on Neural Information Processing Systems (NeurIPS). La Jolla, CA: NIPS, 2019: 6781−6792

[74] Lloyd S, Weedbrook C. Quantum generative adversarial learning[J]. Physical Review Letters, 2018, 121(4): 040502

[75] Kennedy J, Eberhart R. Particle swarm optimization[C]//Proc of ICNN'95. Piscataway, NJ: IEEE, 1995: 1942−1948

[76] Zhu Daiwei, Linke N M, Benedetti M, et al. Training of quantum circuits on a hybrid quantum computer[J]. Science Advances, 2019, 5(10): eaaw9918

[77] Spall J C. A one-measurement form of simultaneous perturbation stochastic approximation[J]. Automatica, 1997, 33(1): 109−112

[78] Patel Y J, Kundu A, Ostaszewski M, et al. Curriculum reinforcement learning for quantum architecture search under hardware errors[J]. arXiv preprint, arXiv: 2402.03500, 2024.

[79] Hansen N, Müller S D, Koumoutsakos P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)[J]. Evolutionary Computation, 2003, 11(1): 1−18

[80] Alcazar J, Leyton-Ortega V, Perdomo-Ortiz A. Classical versus quantum models in machine learning: Insights from a finance application[J]. Machine Learning: Science and Technology, 2020, 1(3): 035003

[81] Las Heras U, Alvarez-Rodriguez U, Solano E, et al. Genetic algorithms for digital quantum simulations[J]. Physical Review Letters, 2016, 116(23): 230504

[82] Kondratyev A. Non-differentiable leaning of quantum circuit Born machine with genetic algorithm[J]. Wilmott, 2021, 2021(114): 50−61

[83] Ding Yongcheng, Lamata L, Sanz M, et al. Experimental implementation of a quantum autoencoder via quantum adders[J]. Advanced Quantum Technologies, 2019, 2(7/8): 1800065

[84] Mitarai K, Negoro M, Kitagawa M, et al. Quantum circuit learning[J]. Physical Review A, 2018, 98(3): 032309

[85] He Guangping. Computing the gradients with respect to all parameters of a quantum neural network using a single circuit[J]. arXiv preprint, arXiv: 2307.08167, 2023

[86] Li Jun, Yang Xiaodong, Peng Xinhua, et al. Hybrid quantum-

classical approach to quantum optimal control[J]. Physical Review Letters, 2017, 118(15): 150503

[87] Nakanishi K M, Fujii K, Todo S. Sequential minimal optimization for quantum-classical hybrid algorithms[J]. Physical Review Research, 2020, 2(4): 043158

[88] Parrish R M, Iosue J T, Ozaeta A, et al. A Jacobi diagonalization and Anderson acceleration algorithm for variational quantum algorithm parameter optimization[J]. arXiv preprint, arXiv: 1904.03206, 2019

[89] Ostaszewski M, Grant E, Benedetti M. Structure optimization for parameterized quantum circuits[J]. Quantum, 2021, 5: 391

[90] Shor P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer[J]. SIAM Review, 1999, 41(2): 303−332

[91] Huang H Y, Broughton M, Mohseni M, et al. Power of data in quantum machine learning[J]. Nature Communications, 2021, 12(1): 2631

[92] Caro M C, Huang H Y, Cerezo M, et al. Generalization in quantum machine learning from few training data[J]. Nature Communications, 2022, 13(1): 4919

[93] Chia N H, Gilyén A P, Li T, et al. Samplingbased sublinear low-rank matrix arithmetic framework for dequantizing quantum machine learning[J]. Journal of the ACM, 2022, 69(5): 33

[94] Huang H Y, Kueng R, Torlai G, et al. Provably efficient machine learning for quantum many-body problems[J]. Science, 2022, 377(6613): eabk3333

[95] Huang H Y, Broughton M, Cotler J, et al. Quantum advantage in learning from experiments[J]. Science, 2022, 376(6598): 1182−1186

[96] Aharonov D, Cotler J, Qi X L. Quantum algorithmic measurement[J]. Nature Communications, 2022, 13(1): 887

[97] Bravyi S, Gosset D, König R. Quantum advantage with shallow circuits[J]. Science, 2018, 362(6412): 308−311

[98] Farhi E, Neven H. Classification with quantum neural networks on near term processors[J]. arXiv preprint, arXiv: 1802.06002, 2018

[99] Chefles A. Quantum state discrimination[J]. Contemporary Physics, 2000, 41(6): 401−424

[100] Barnett S M, Croke S. Quantum state discrimination[J]. Advances in Optics and Photonics, 2009, 1(2): 238−278

[101] Čepaitė I, Coyle B, Kashefi E. A continuous variable Born machine[J]. Quantum Machine Intelligence, 2022, 4(1): 6

[102] Coyle B, Mills D, Danos V, et al. The Born supremacy: Quantum advantage and training of an ising Born machine[J]. NPJ Quantum Information, 2020, 6(1): 60

[103] Rudolph M S, Toussaint N B, Katabarwa A, et al. Generation of high-resolution handwritten digits with an iontrap quantum computer[J]. Physical Review X, 2022, 12(3): 031010

[104] Kao P Y, Yang Y C, Chiang W Y, et al. Exploring the advantages of quantum generative adversarial networks in generative chemistry[J]. Journal of Chemical Information and Modeling, 2023, 63(11): 3307−3318

[105] Skolik A, Jerbi S, Dunjko V. Quantum agents in the gym: A variational quantum algorithm for deep Q-learning[J]. Quantum, 2022, 6: 720

[106] Skolik A, Mangini S, Bäck T, et al. Robustness of quantum reinforcement learning under hardware errors[J]. EPJ Quantum Technology, 2023, 10(1): 8

[107] Niu M Y, Zlokapa A, Broughton M, et al. Entangling quantum generative adversarial networks[J]. Physical Review Letters, 2022, 128(22): 220505

[108] Du Yuxuan, Huang Tao, You Shan, et al. Quantum circuit architecture search for variational quantum algorithms[J]. NPJ Quantum Information, 2022, 8(1): 62

[109] Schuld M, Bocharov A, Svore K M, et al. Circuit-centric quantum classifiers[J]. Physical Review A, 2020, 101(3): 032308

[110] Henderson M, Shakya S, Pradhan S, et al. Quanvolutional neural networks: Powering image recognition with quantum circuits[J]. Quantum Machine Intelligence, 2020, 2(1): 2

[111] Chu Cheng, Chia N H, Jiang Lei, et al. QMLP: An error-tolerant nonlinear quantum mlp architecture using parameterized two-qubit gates[C/OL]//Proc of the 29th Int Symp on Low Power Electronics and Design. New York: ACM, 2022[2023-11-16]. https://dl.acm.org/doi/abs/10.1145/35 31437.3539719

[112] Chen S Y C, Huang C M, Hsing C W, et al. An end-to-end trainable hybrid classical-quantum classifier[J]. Machine Learning: Science and Technology, 2021, 2(4): 045021

[113] Pesah A, Cerezo M, Wang S, et al. Absence of barren plateaus in quantum convolutional neural networks[J]. Physical Review X, 2021, 11(4): 041011

[114] Mcclean J R, Boixo S, Smelyanskiy V N, et al. Barren plateaus in quantum neural network training landscapes[J]. Nature Communications, 2018, 9(1): 4812

[115] Monteiro C A, Gustavo Filho I, Costa M H J, et al. Quantum neuron with real weights[J]. Neural Networks, 2021, 143: 698−708

[116] Hu Zhirui, Li Jinyang, Pan Zhenyu, et al. On the design of quantum graph convolutional neural network in the NISQ-era and beyond[C]//Proc of the 40th Int Conf on Computer Design (ICCD). Piscataway, NJ: IEEE, 2022: 290−297

[117] Shepherd D, Bremner M J. Temporally unstructured quantum computation[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2009, 465(2105): 1413−1439

[118] Amin M H, Andriyash E, Rolfe J, et al. Quantum Boltzmann machine[J]. Physical Review X, 2018, 8(2): 021050

[119] Kieferová M, Wiebe N. Tomography and generative training with quantum Boltzmann machines[J]. Physical Review A, 2017, 96(6): 062327

[120] Huijgen O, Coopmans L, Najafi P, et al. Training quantum Boltzmann machines with the β-variational quantum eigensolver[J]. Machine Learning: Science and Technology, 2024, 5(2): 025017

[121] Khoshaman A, Vinci W, Denis B, et al. Quantum variational autoencoder[J]. Quantum Science and Technology, 2018, 4(1): 014001

[122] Huang Changjiang, Ma Hailan, Yin Qi, et al. Realization of a quantum autoencoder for lossless compression of quantum data[J]. Physical Review A, 2020, 102(3): 032412

[123] Cerezo M, Sone A, Volkoff T, et al. Cost function dependent barren plateaus in shallow parametrized quantum circuits[J]. Nature Communications, 2021, 12(1): 1791

[124] Dallaire-Demers P L, Killoran N. Quantum generative adversarial networks[J]. Physical Review A, 2018, 98(1): 012324

[125] Tian Jinkai, Sun Xiaoyu, Du Yuxuan, et al. Recent advances for quantum neural networks in generative learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(10): 12321−12340

[126] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Proc of the 28th Conf on Neural Information Processing Systems (NeurIPS). La Jolla, CA: NIPS, 2014, 2672−2680

[127] Carleo G, Cirac I, Cranmer K, et al. Machine learning and the physical sciences[J]. Reviews of Modern Physics, 2019, 91(4): 045002

[128] Hu Ling, Wu Shuhao, Cai Weizhou, et al. Quantum generative adversarial learning in a superconducting quantum circuit[J]. Science Advances, 2019, 5(1): eaav2761

[129] Kim L, Lloyd S, Marvian M. Hamiltonian quantum generative adversarial networks[J]. Physical Review Research, 2024, 6(3): 033019

[130] Du Yuxuan, Hsieh M H, Tao Dacheng. Efficient online quantum generative adversarial learning algorithms with applications[J]. arXiv preprint arXiv: 1904.09602, 2019

[131] Pan Minghua, Wang Bin, Tao Xiaoling, et al. Application of quantum generative adversarial network to the abnormal user behavior detection and evaluation[J]. arXiv preprint, arXiv: 2208.09834, 2022

[132] Watkins C J, Dayan P. Q-learning[J]. Machine Learning, 1992, 8(3/4): 279−292

[133] Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning[J]. arXiv preprint, arXiv: 1312.5602, 2013

[134] Van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double Q-learning[C]//Proc of the 30th AAAI Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2016: 2094−2100

[135] Jerbi S, Gyurik C, Marshall S, et al. Parametrized quantum policies for reinforcement learning[C]//Proc of the 35th Annual Conf on Neural Information Processing Systems (NeurIPS). La Jolla, CA: NIPS, 2021: 28362−28375

[136] Jerbi S, Cornelissen A, Ozols M, et al. Quantum policy gradient algorithms[J]. arXiv preprint, arXiv: 2212.09328, 2022

[137] Wu Shaojun, Jin Shan, Wen Dingding, et al. Quantum reinforcement learning in continuous action space[J]. arXiv preprint, arXiv: 2012.10711, 2020

[138] Ostaszewski M, Trenkwalder L M, Masarczyk W, et al. Reinforcement learning for optimization of variational quantum circuit architectures[C]//Proc of the 35th Conf on Neural Information Processing Systems (NeurIPS). La Jolla, CA: NIPS, 2021: 18182−18194

[139] Wang S, Fontana E, Cerezo M, et al. Noise-induced barren plateaus in variational quantum algorithms[J]. Nature Communications, 2021, 12(1): 6961

[140] Liang Zhiding, Wang Hanrui, Cheng Jinglei, et al. Variational quantum pulse learning[C]//Proc of the 3rd IEEE Int Conf on Quantum Computing and Engineering (QCE). Los Alamitos, CA: IEEE Computer Society, 2022: 556−565

[141] Meitei O R, Gard B T, Barron G S, et al. Gatefree state preparation for fast variational quantum eigensolver simulations[J]. NPJ Quantum Information, 2021, 7(1): 155

[142] Liang Zhiding, Cheng Jinglei, Ren Hang, et al. NAPA: Intermediate-level variational native-pulse Ansatz for variational quantum algorithms[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2024, 43(6): 1834−1847

[143] Rattew A G, Hu S, Pistoia M, et al. A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver[J]. arXiv preprint, arXiv: 1910.09694, 2019

[144] Havlíček V, Córcoles A D, Temme K, et al. Supervised learning with quantum-enhanced feature spaces[J]. Nature, 2019, 567(7747): 209−212

[145] Stilck França D, Garcia-Patron R. Limitations of optimization algorithms on noisy quantum devices[J]. Nature Physics, 2021, 17(11): 1221−1227

[146] Chen Yiting, Farquhar C, Parrish R M. Low-rank density-matrix evolution for noisy quantum circuits[J]. NPJ Quantum Information, 2021, 7(1): 61

[147] Sharma K, Khatri S, Cerezo M, et al. Noise resilience of variational quantum compiling[J]. New Journal of Physics, 2020, 22(4): 043006

[148] Skolik A, Mcclean J R, Mohseni M, et al. Layerwise learning for quantum neural networks[J]. Quantum Machine Intelligence, 2021, 3(1): 5

[149] Volkoff T, Coles P J. Large gradients via correlation in random parameterized quantum circuits[J]. Quantum Science and Technology, 2021, 6(2): 025008

[150] Endo S, Cai Z, Benjamin S C, et al. Hybrid quantumclassical algorithms and quantum error mitigation[J]. Journal of the Physical Society of Japan, 2021, 90(3): 032001

[151] Bilkis M, Cerezo M, Verdon G, et al. A semi-agnostic Ansatz with variable structure for quantum machine learning[J]. arXiv preprint, arXiv: 2103.06712, 2021

[152] Weber M, Liu Nana, Li Bo, et al. Optimal provable robustness of quantum classification via quantum hypothesis testing[J]. NPJ Quantum Information, 2021, 7(1): 76

[153] Du Yuxuan, Hsieh M H, Liu Tongliang, et al. Quantum noise protects quantum classifiers against adversaries[J]. Physical Review Research, 2021, 3(2): 023153

[154] Liu Junyu, Wilde F, Mele A A, et al. Stochastic noise can be helpful for variational quantum algorithms[J]. arXiv preprint, arXiv: 2210.06723, 2022

[155] Gentini L, Cuccoli A, Pirandola S, et al. Noise-resilient variational hybrid quantum-classical optimization[J]. Physical Review A, 2020, 102(5): 052414

[156] Zhang Kaining, Liu Liu, Hsieh M H, et al. Escaping from the barren plateau via Gaussian initializations in deep variational quantum circuits[C]//Proc of the 36th Conf on Neural Information Processing Systems (NeurIPS). La Jolla, CA: NIPS, 2022: 18612−18627

[157] Cervera-Lierta A, Kottmann J S, Aspuruguzik A. Meta-variational quantum eigensolver: Learning energy profiles of parameterized hamiltonians for quantum simulation[J]. PRX Quantum, 2021, 2(2): 020329

[158] Harrow A W, Low R A. Random quantum circuits are approximate 2-designs[J]. Communications in Mathematical Physics, 2009, 291(1): 257−302

[159] Holmes Z, Sharma K, Cerezo M, et al. Connecting Ansatz expressibility to gradient magnitudes and barren plateaus[J]. PRX Quantum, 2022, 3(1): 010313

[160] Sharma K, Cerezo M, Cincio L, et al. Trainability of dissipative perceptron-based quantum neural networks[J]. Physical Review Letters, 2022, 128(18): 180505

[161] Kashif M, Al-Kuwari S. The impact of cost function globality and locality in hybrid quantum neural networks on NISQ devices[J]. Machine Learning: Science and Technology, 2023, 4(1): 015004

[162] Sim S, Johnson P D, Aspuru-Guzik A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms[J]. Advanced Quantum Technologies, 2019, 2(12): 1900070

[163] Nielsen M A, Dawson C M, Dodd J L, et al. Quantum dynamics as a physical resource[J]. Physical Review A, 2003, 67(5): 052301

[164] Jaques S, Rattew A G. Qram: A survey and critique[J]. arXiv preprint, arXiv: 2305.10310, 2023

[165] Phalak K, Li Junde, Ghosh S. Trainable PQC-based QRAM for quantum storage[J]. IEEE Access, 2023, 11: 51892−51899

[166] Fu Xiang, Zheng Yuzhen, Su Xing, et al. A heterogeneous quantum-classical computing system targeting noisy intermediate-scale quantum technology[J]. Journal of Computer Research and Development, 2021, 58(9): 1875−1896 (in Chinese)
(付祥, 郑宇真, 苏醒, 等. 一种面向含噪中尺度量子技术的量子-经典异构计算系统[J]. 计算机研究与发展, 2021, 58(9): 1875−1896)

[167] Verdon G, Pye J, Broughton M. A universal training algorithm for quantum deep learning[J]. arXiv preprint, arXiv: 1806.09729, 2018



Yu Ruiqi, born in 2002. Master candidate. His main research interests include quantum computing and machine learning.
于瑞祺, 2002 年生. 硕士研究生. 主要研究方向为量子计算、机器学习.



Zhang Xinyun, born in 1997. PhD. Member of CCF. His main research interests include machine learning, 3D computer vision, and quantum computing.
张鑫云, 1997 年生. 博士. CCF 会员. 主要研究方向为机器学习、3 维计算机视觉、量子计算.



Ren Shuang, born in 1981. PhD, associate professor, PhD supervisor. Member of CCF. His main research interests include machine learning, quantum computing, and 3D computer vision.
任　爽, 1981 年生. 博士, 副教授, 博士生导师. CCF 会员. 主要研究方向为机器学习、量子计算、3D 计算机视觉.