面向近期量子处理器的量子神经网络研究进展

吕颜轩¹, 高庆^{1,2*}, 吕金虎^{1,2}, 潘宇³, 董道毅⁴

1. 北京航空航天大学自动化科学与电气工程学院, 北京 100191;

2. 北京航空航天大学大数据与脑机智能高精尖创新中心, 北京 100191;

3. 浙江大学智能系统与控制研究所, 杭州 310027;

4. 新南威尔士大学工程与信息技术学院, 堪培拉 2600, 澳大利亚

* E-mail: gaoqing@buaa.edu.cn

收稿日期: 2020-11-17; 接受日期: 2020-12-17; 网络版发表日期: 2021-05-11

国家自然科学基金(批准号: 61903016, 61803132)和德国亚历山大·冯·洪堡基金资助

摘要 近年来, 人工神经网络在各领域得到了广泛的应用, 展现出强大的计算智能. 与此同时, 量子计算硬件也得到了飞速发展, 近期量子处理器已具备较稳定的计算能力和抑制退相干能力, 多家商用云量子计算机公司已能够为世界各地的学者们提供在线量子计算实验平台. 在诸多技术领域的迅猛进展下, 量子神经网络这一交叉领域也重获科学界的关注, 涌现出了大量新的研究思路并取得了重要的实验进展. 本文首先对早期量子神经网络的研究思路作简要阐述, 然后对基于近期量子处理器的量子神经网络的数据编码、算法流程和量子电路设计等进行综述分析, 最后总结量子神经网络领域的数个关键科学问题, 对未来研究方向做出展望.

关键词 人工智能, 量子神经网络, 近期量子处理器

1 引言

人工神经网络是现代人工智能技术的一个基础模型, 因具有较强的泛化能力和非线性信息处理能力, 在计算机视觉^[1-3]、自然语言处理^[4]、医疗诊断^[5]等领域都取得了巨大的成功, 成为人工智能方向和其他多学科领域的研究热点^[6]. 人工神经网络的基本原理是通过构建生物体神经网络的近似数学模型, 模拟生物处理信息的方式, 实现对输入数据的智能分析. 人工神经网络由具有简单计算功能的神经元单元构成, 利用多层级联结构和非线性激活函数实现对复杂映射关系的拟合, 展现出强大的学习能力. 已经证明, 人工神经

网络可以以任意精度逼近任意非线性函数, 而非线性激活函数则是人工神经网络强大的函数逼近性能的重要因素^[7].

近年来, 随着量子技术的飞速发展, 量子机器学习理论^[8-10], 如量子主成分分析^[11]、量子强化学习^[12]、量子支持向量机^[13]等, 逐步受到了学界的广泛关注. 在量子机器学习的诸多领域中, 量子神经网络融合了量子计算与神经网络模型的各自优势, 有潜力显著提升经典神经网络的运算效率, 以实现更强大的人工智能. 量子计算的设想最早由Feynman于1982年提出^[14], 其指出基于量子力学建造的量子计算机具有经典计算机所无法比拟的优势. 此后, 相继涌现出的许多量子算

引用格式: 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, 2022, 52: 547-564

Lü Y X, Gao Q, Lü J H, et al. Recent advances of quantum neural networks on the near term quantum processor (in Chinese). Sci Sin Tech, 2022, 52: 547-564, doi: 10.1360/SST-2020-0459

法,如Shor算法^[15]、Grover量子搜索算法^[16]等,被证明在特定问题的求解上相比于经典算法具有指数级或平方级的运算效率提升.在量子计算中,量子比特(qubit)作为基本计算单元,其演化满足线性幺正变换.因此,即便搭建最简单的量子感知机神经网络模型,也需要谨慎处理其非线性实现机制^[17].近年来,为了解决量子神经网络的非线性实现机理与量子计算硬件的固有物理属性之间的矛盾,研究者们做出了极大的努力,提出了多种方法.本文将针对相关重要研究进展进行总结与分析.

量子神经网络的早期研究可见于Kak^[18]在1994年的工作.随后,学者们从各个方向广泛地探索了多种可能的量子神经网络模型,主要研究思路包括:(1)基于量子测量中坍缩的非线性与激活函数非线性的相似性,实现量子神经网络;(2)基于交互量子点,通过Feynman路径积分近似神经网络;(3)基于简单耗散量子门实现神经网络的动态特性;(4)经典感知机模型的简单量子类比.在文献^[19]中,作者总结了量子神经网络应该满足的三个条件:(1)系统的初始状态能够编码任何二进制串;(2)计算过程能够反映神经网络计算机制;(3)系统的演化基于量子效应,符合量子力学基本原理.文献^[19]指出当前所提出的量子神经网络模型大多在数学计算层面上探讨量子神经网络,存在着诸如物理可实现性不明确、演化不遵循量子效应以及不具有神经网络计算特点等问题,因而并没有实现真正意义上的“量子神经网络”.

近年来,量子计算硬件技术取得了突破性的进展.2018年,Google公布了当时最大量子比特数——72量子比特的量子处理器Bristlecone.2019年,IBM推出了具有53个量子比特的量子处理器,并可供外部研究人员编程使用,同年,Google宣布其53量子比特的Sycamore量子芯片实现了量子霸权^[20].2020年,霍尼韦尔公司宣布其制造出当前性能最优的量子计算机.学界一般用近期量子处理器(near term quantum processor)^[21]指代在较短时间内(未来五到十年)有望实现的量子处理器,或者称为带噪中等规模量子(noisy intermediate-scale quantum, NISQ)^[22]器件.近期量子处理器上可用的量子比特较少,量子控制易受噪声影响,但是已经具备稳定的计算能力和抑制退相干能力,是目前极为先进的优质量子计算平台.研究面向近期量子处理器的量子神经网络不仅具有相当的现实意义,而且

对未来通用量子计算机的软件构建具有非常重要的理论意义.近年来,相关领域出现了丰富的研究成果,深刻地讨论了各种不同类别的经典神经网络在量子体系架构下的描述,并深入探讨了在量子处理器上的可实现性.

本文首先将介绍量子神经网络的早期研究成果,着重阐述与分析近年来面向近期量子处理器的多种量子神经网络模型,总结量子神经网络领域的数个关键科学问题,并对量子神经网络未来需要重点关注的研究方向作出展望.

2 经典神经网络与量子计算基本原理

2.1 神经网络模型

经典神经网络是一种具有网络结构的仿生运算模型,由多个神经元以特定结构相互连接组成.感知机是一种常用的神经元模型,具有类似生物神经元的阈值激活特性.感知机接收 n 个输入数据 $x_1, x_2, \dots, x_n$,其自身包含 n 个连接的权重参数 $w_1, w_2, \dots, w_n$,以及偏置参数 b .为便于描述,称 $\theta = x_1 w_1 + x_2 w_2 + \dots + x_n w_n + b$ 为感知机的输入信号.经过非线性激活函数 $\sigma(\cdot)$ 的运算,感知机给出输出信号 $z = \sigma(\theta)$,如图1所示.常用的激活函数有阈值函数、sigmoid函数、tanh函数等.

神经元之间不同的连接形式构成了不同的神经网络结构.经典的神经网络结构包括:前馈网络,如BP神经网络;反馈网络,如Hopfield神经网络;此外,还有近期涌现出的多种新的复杂神经网络结构,如卷积网络、残差网络、循环网络等.这些神经网络模型面向不同的学习任务时具有各自的优势.以前馈神经网络为例,图2中给出了其典型结构,包括输入层、隐藏层和输出层,分别对应信息输入、信息传递和信息提取三个信息处理流程.在学习过程中,人工神经网络的结构往往是固定的,神经元的连接权重和阈值根据输入数据进行调整,以拟合输入和输出的复杂映射关系.

量子神经网络的基本思想是通过融合量子计算机制来实现神经网络的基本结构,以期利用量子计算的存储和计算能力优势实现更高效或者更智能的学习能力.在经典神经网络中,信息的输入、传递和提取是非常自然的,通过二进制编码可以方便地对经典信息进行存储和计算.量子神经网络具有类似的信息处理流

图 1 感知机模型

Figure 1 Structure of the perceptron.

图 2 前馈神经网络结构

Figure 2 Structure of the feedforward neural network.

程, 如图3所示, 但是量子力学的独特性质却使得相应过程具有截然不同的特点: 量子神经网络只能对量子数据进行处理, 所以需要将经典信息用量子比特进行编码, 转化为量子数据; 量子门对量子信息的变换与经典逻辑门对经典信息的变换不同, 是么正(unitary)的、线性的; 量子数据中的信息处于叠加态, 对其进行测量会造成随机坍缩, 造成部分信息丢失。这些独特的性质既构成了量子机器学习的理论基石, 也为量子神经网络的研究带来了科学挑战。下面将分别介绍经典信息的量子态描述、量子门以及量子测量等量子力学相关基础, 并描述它们在量子电路中的实现方法^[23]。

2.2 经典信息的量子态描述

在经典电子计算机中, 信息存储和操作的基本单元是比特(bit), 其值可取为0或者1, 通常采用电平的高低或者磁性颗粒的磁场方向作为信息载体。在量子计算中, 信息的基本单元是量子比特(qubit)。量子比特也有两个基本状态, 对应二能级量子系统的基态、激发态或者光子的两个相互垂直的偏振方向等多种物理形式, 分别用 $|0\rangle$ 和 $|1\rangle$ 来表示。不同于经典比特在任一时刻只能取一个值, 量子比特可以处于 $|0\rangle$ 和 $|1\rangle$ 的任意叠加态, 表示为

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad (1)$$

其中, 复数 α, β 表示概率幅, 满足 $|\alpha|^2 + |\beta|^2 = 1$ 。选择 $\{|0\rangle, |1\rangle\}$ 作为一组基, 则任意量子比特态 $|\psi\rangle$ 可用如下 $\mathbb{C}^2$ 中的复向量表示:

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (2)$$

在量子力学里, 我们称一个与环境完全没有交互作用的量子系统为孤立量子系统。对于 n 个孤立的量子比特, 其量子态分别记作 $|\psi^1\rangle, \dots, |\psi^n\rangle$, 则由 n 个量子比特构成的复合系统量子态为 $|\Psi\rangle = |\psi^1\rangle \otimes \dots \otimes |\psi^n\rangle$ 。为方便起见, 一般将 $|\Psi\rangle$ 写作 $|\psi^1 \dots \psi^n\rangle$ 。显然, 任意量子态 $|\Psi\rangle$ 所属的线性空间拥有一组基: $\{|00 \dots 0\rangle, |00 \dots 1\rangle, \dots, |11 \dots 1\rangle\}$ 。因此, 任意量子态 $|\Psi\rangle$ 均可表达为如下的叠加态:

$$|\Psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle, \alpha_i \in \mathbb{C}, \quad (3)$$

其中, $|i\rangle$ 表示对整数 i 的二进制描述所对应的量子态,

图 3 量子神经网络信息处理流程

Figure 3 Information processing of the quantum neural network.

如 $|7\rangle = |111\rangle$. 式(3)中的系数同样满足归一化性质

$\sum_{i=0}^{2^n-1} |\alpha_i|^2 = 1$. 同样的, 选定基向量后, $|\Psi\rangle$ 可以用如下

$\mathbb{C}^{2^n}$ 中的复向量唯一表示:

$$|\Psi\rangle = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{2^n-1} \end{bmatrix}. \quad (4)$$

如果量子比特之间发生了纠缠, 则复合量子系统的状态不能写成各量子比特量子态的张量积的形式,

如Bell态: $|\phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$. 这里我们不针对量子纠缠做过多讨论.

量子处理器可以直接处理量子数据, 当面对经典信息时, 则需要将经典数据预先转化为量子数据. 基于量子态的叠加性质, 对经典信息的编码主要有两种方式, 分别是量子比特编码和幅值编码. 量子比特编码将量子比特视作经典比特, 与经典比特的二进制编码方式类似. 量子态 $|0\rangle$ 表示数字0, 量子态 $|1\rangle$ 表示数字1, 可以用 n 位量子比特表示一个范围为 $[0, 2^n-1]$ 的整数或者一定精度的浮点数. 幅值编码则利用量子态的概率幅来编码经典数据. 对于 n 个量子比特, 其量子态包含 2^n 个概率幅. 一种幅值编码方式是利用这些系数存储 2^n 个实数, 但是通常制备这样的量子态需要很高的代价, 甚至可能抵消量子计算的速度优势. 另一种方式是将数据分别编码到单个量子比特的概率幅上, 这样 n 个量子比特只能编码 n 个实数, 但是制备效率大大提高. 在量子神经网络中, 针对不同的网络结构和数据类型会采用不同的编码方式.

2.3 量子门

经典计算中, 可以通过逻辑门对经典比特的值进行变换. 类似的, 在量子计算中, 量子逻辑门也能够对量子比特进行操作. 量子力学中, 孤立量子系统的演化可以用幺正变换 U 来描述, U 满足 $UU^\dagger = I$, 其中 $U^\dagger$ 表示 U 的共轭转置, I 是单位矩阵. 量子门就是通过对量子比特施加幺正变换来实现的.

对于初始状态为 $|\psi_0\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle$ 的量子系统, 量子

门 U 将初始态 $|\psi_0\rangle$ 变换为

$$U|\psi_0\rangle = U \sum_{i=0}^{2^n-1} \alpha_i |i\rangle = \sum_{i=0}^{2^n-1} \beta_i |i\rangle. \quad (5)$$

同样的, 选择 $\{|i\rangle\}$ 作为基矢态, 上述操作可以写为如下矩阵乘法操作:

$$U\bar{\alpha} = \bar{\beta}, \quad (6)$$

其中, $\bar{\alpha} = [\alpha_0, \dots, \alpha_{2^n-1}]^T$, $\bar{\beta} = [\beta_0, \dots, \beta_{2^n-1}]^T$, U 是量子门 U 的矩阵形式. 可见, 单个量子门进行运算时, 同时对量子态所有概率幅进行运算, 即相当于同时进行了 2^n 次运算.

量子门的幺正变换是可逆的, 对于给定的量子态 $|\psi_i\rangle$ 和施加的幺正变换 U_i , 可得到初始时刻的量子态 $|\psi_0\rangle = U_i^\dagger |\psi_i\rangle$, 所以量子门的信息传递是可逆的无损信息传递. 量子门的输入和输出的量子比特数相等, 否则变换将不是幺正的. 由于可逆计算过程没有能量损失, 量子计算原理上是无功耗的.

在文献[24]中, 作者证明任意的幺正变换都可以用一组有限数目的简单幺正变换的组合来实现. 本文主要关注其中的单量子比特门和受控门(见表1), 其他量子门可以由这些基本量子门的组合来实现. 基本的单量子比特门包括量子非门(X门)、Hadamard门和泡利-Z门(Z门). 非门使得量子比特的值反转, Hadamard门使得基态 $|0\rangle$ 和 $|1\rangle$ 变为均匀叠加态 $|+\rangle$ 和 $|-\rangle$, 泡利-Z门则改变量子态的相位. 受控门接收两个及以上的量子比特, 其中一个或多个量子比特作为控制比特, 是否对其余受控量子比特施加变换取决于控制比特的值. 受控非门(CNOT)是最简单的受控门, 其操作两个量子比特, 仅当控制量子比特的值为 $|1\rangle$ 时, 对受控量子比特取反. 进一步可扩展得到受控U门(C-U), 其仅当控制量子比特为 $|1\rangle$ 时, 对其余多个量子比特施加幺正变换 U .

2.4 量子测量

量子电路中的测量如图4所示. 在标准基矢 $\{|0\rangle, |1\rangle\}$ 下, 投影测量操作对应的矩阵表示分别为

表 1 基本量子门及受控多量子比特门的变换及表示

Table 1 Definitions and matrix representations of basic quantum gates and the multi-target-qubit controlled gates

量子门	变换	矩阵	量子电路
X 门	$X 0\rangle = 1\rangle$ $X 1\rangle = 0\rangle$	$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$ \psi_i\rangle \xrightarrow{X} \psi_o\rangle$
Hadamard 门	$H 0\rangle = +\rangle = \frac{ 0\rangle + 1\rangle}{\sqrt{2}}$ $H 1\rangle = -\rangle = \frac{ 0\rangle - 1\rangle}{\sqrt{2}}$	$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	$ \psi_i\rangle \xrightarrow{H} \psi_o\rangle$
Z 门	$Z 0\rangle = 0\rangle$ $Z 1\rangle = - 1\rangle$	$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$ \psi_i\rangle \xrightarrow{Z} \psi_o\rangle$
CNOT 门	$ 00\rangle \mapsto 00\rangle$ $ 01\rangle \mapsto 01\rangle$ $ 10\rangle \mapsto 11\rangle$ $ 11\rangle \mapsto 10\rangle$	$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	$ \psi_i\rangle \xrightarrow{\text{CNOT}} \psi_o\rangle$
C-U 门	$ 0\rangle \psi\rangle \mapsto 0\rangle \psi\rangle$ $ 1\rangle \psi\rangle \mapsto 1\rangle U \psi\rangle$	$\text{C-U} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & U \end{bmatrix}$	$ \psi_i\rangle \xrightarrow{\text{C-U}} \psi_o\rangle$

图 4 量子电路中的测量

Figure 4 Measurement in quantum circuits.

$$M_0 = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, M_1 = |1\rangle\langle 1| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (7)$$

对量子态为 $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ 的量子比特进行测量时, 量子态投影到基矢态 $|0\rangle$ 上的概率为

$$P(0) = \langle \psi | M_0^\dagger M_0 | \psi \rangle = [\alpha \quad \beta] \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = |\alpha|^2. \quad (8)$$

测量过后, 量子态坍缩为基矢态 $|0\rangle$. 同理, 计算可得, 量子测量后得到基矢态 $|1\rangle$ 的概率为 $|\beta|^2$. 这与经典比特每次测量返回确定的 0 或 1 有所不同.

由于测量结果是随机性的, 每次测量只能以一定概率得到其中一个基态, 所以不能通过单次测量精确提取量子态信息 $(|\alpha|^2, |\beta|^2)$. 为准确获知量子态, 需要进行多次重复测量, 用频率来近似概率. 而且, 测量会破坏量子态, 测量之后叠加态消失, 量子态中不再包

含 α 和 β 的信息. 想要重复测量, 需要重新制备初始量子态, 重复上述操作.

一个完整的量子电路包含量子比特、量子门和量子测量, 如图 5 所示. 在量子电路中, 部分量子比特在初始时编码输入数据. 量子门按照操作顺序自左向右放置. 最终运算结果通过对部分量子比特施加量子测量获得.

3 早期量子神经网络

在量子神经网络的研究初期, 即 20 世纪 90 年代左右, 世界上尚未出现成熟的量子计算机模型. 早期的量子神经网络大多基于量子计算的相关物理过程提出, 并没有针对量子比特、量子电路等具体量子处理器结构进行描述, 存在诸如参数训练过程不符合量子系统演化的规律、非线性量子器件尚不具备等问题, 因而难以在量子计算机上实现. 早期量子神经网络大致可归纳为四类: (1) 基于多世界诠释的量子神经网络; (2) 基于交互量子点的量子神经网络; (3) 基于简单耗散量子门的量子神经网络; (4) 类比经典感知机的早期量子感知机模型. 下面将对这些模型分别进行介绍.

图5 量子电路
Figure 5 A quantum circuit.

3.1 基于多世界诠释的量子神经网络

多世界诠释是量子力学诠释的一种, 对于测量一个处于叠加状态的粒子所引起的坍缩现象, 其认为, 这是由于测量时分离出了许多平行宇宙, 每个宇宙中包含粒子的一个确定的状态. 基于多世界诠释, 早期学者对量子神经网络进行了原理上的构想. Kak^[18]最早提出量子神经网络的概念, 他将类Hopfield网络的稳定态诠释为量子系统的特征值. 网络权重的更新对应于通过量子测量实现系统态矢量的一个特征向量选取. Kak指出量子力学中没有对应sigmoid函数的非线性特性, 但并没有继续讨论这个重要问题. 文献[25]提出量子启发神经网络, 从多世界诠释出发, 把量子神经网络的叠加诠释为每个宇宙储存了一个模式的模型参数. 进行测量将使得神经网络的参数坍缩到其中一个模式上.

3.2 基于交互量子点的量子神经网络

量子点是微小的约束电荷岛, 可以用作量子比特, 通常通过与周围光子或场的交互进行演化. 文献[26]提出采用单个量子点在不同时间段的演化来模拟多个神经元, 构成时间序列神经网络. 神经元的演化由Feynman路径积分来描述, 通过Feynman路径积分中的指数函数自然地引入非线性, 符合神经网络耗散、不可逆的特性, 可实现经典神经网络的前向和反向传播. 但是, 其运算形式相比经典感知机存在较大差异, 因此对其是否反映神经计算机制存在争议.

3.3 基于耗散门的量子神经网络

文献[27]提出一种非线性、不可逆的耗散量子门的设想, 并基于该耗散门构建了一种量子神经网络模型. 在该量子神经网络的量子计算操作中, 每个么正门都伴随一个耗散门, 耗散门依据量子态是否超过阈值将其映射为一个复数或者0, 从而模拟感知机的激

活函数. 耗散门的非线性特性不符合量子力学演化的么正性, 但是, 该文献对具有类似非线性关系的量子系统进行了分析, 指出这样的耗散量子门是可能实现的.

3.4 早期量子感知机模型

文献[28]最早提出了量子感知机模型, 其具有类似经典感知机的状态更新函数和权重学习规则:

$$|y(t)\rangle = F \sum_{i=1}^m \bar{w}_i(t) |x_i\rangle, \quad (9)$$

$$\bar{w}_i(t+1) = \bar{w}_i(t) + \eta(|d\rangle - |y(t)\rangle)\langle x_i|, \quad (10)$$

其中, F 和 $\bar{w}_i$ 为么正变换, $|d\rangle$ 是目标量子态, $|x_i\rangle$ 和 $|y(t)\rangle$ 是感知机在 t 时刻的输入和输出量子态, η 为学习率. 可见, 经典感知机中的输入和输出向量被量子态所取代, 而感知机权值和激活函数被替换为量子门. 文献[28]指出, 这样的具有非线性映射的权重学习规则是非么正的, 参数更新不能保留系统全部的概率; 如果采用么正的学习规则, 则无法反映神经网络学习的非线性和耗散特性. 该文献没有解决上述矛盾, 但是后面介绍的面向近期量子处理器的量子-经典混合计算框架为量子感知机的学习提供了新的思路.

4 面向近期量子处理器的量子神经网络

近年来, 量子技术的发展从硬件与算法两方面推动了量子神经网络的发展. 在硬件技术方面, 近期量子处理器尽管量子比特数量较少, 远未达到通用量子计算系统的要求, 但是为当前探索量子算法提供了可靠的硬件测试平台. 在算法研究方面, 量子-经典混合算法引入了经典计算过程, 使得量子神经网络的搭建和训练更加灵活, 为充分利用当前有限的量子比特资源提供了一种可行的方案.

相比早期研究成果, 近期提出的量子神经网络具有更宽泛的含义. “神经网络”一词更多地被用于指代通过量子电路或量子系统实现的具有网络化结构、参数可训练的计算模型^[29]. 此外, 量子神经网络模型的研究也更加强调模型的物理可实现性.

下面我们从基本的量子感知机及更复杂的神经网络模型两方面来介绍近年来的主要研究进展.

4.1 量子感知机

感知机是神经网络的常见基本组成单元, 结构相对简单, 因而设计量子神经网络的一个自然的想法是构建量子感知机. 量子感知机通常追求模拟经典感知机的前向传播和非线性激活函数. 下面将介绍几类近期提出的量子感知机模型.

4.1.1 基于RUS电路的量子感知机

文献[30]提出一种基于RUS (repeat until success) 量子电路^[31]的感知机, 可以在量子比特的相位上实现近似的阈值激活函数, 并且完全保持量子演化的么正特性. 感知机的基本结构如图6所示.

该感知机包含 n 个输入量子比特、1个辅助量子比特以及1个输出量子比特. 输入数据 $x_1, x_2, \dots, x_n \in \{0, 1\}$ 用输入量子比特的量子态 $|x\rangle = |x_1 x_2 \dots x_n\rangle$ 编码, 其中, $|x_i\rangle = |0\rangle$ 表示 $x_i = 0$, $|x_i\rangle = |1\rangle$ 表示 $x_i = 1$. 该编码方式能够编码所有可能的 n 维二值输入 $x \in \{0, 1\}^n$. 辅助量子比特和输出量子比特的初始状态均为 $|0\rangle$. 量子门包括受控旋转门 $C-R_y(t)$ 和受控Y门, 其中,

$$R_y(t) = \exp(-itY/2), \quad (11)$$

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (12)$$

为模拟经典感知机计算输入信号 $\theta = \sum_{i=1}^n x_i w_i + b$, 分别以量子比特 $|x_i\rangle$ 作为控制比特对辅助量子比特施加受控旋转门 $C-R_y(2w_i)$, 以及旋转门 $R_y(2b)$. 上述量子门序列等价于对辅助量子比特施加 $R_y(2\theta)$ 的么正变换. 系统量子态变换为

$$|x\rangle |0\rangle |0\rangle \mapsto |x\rangle e^{-i\theta Y} |0\rangle |0\rangle. \quad (13)$$

进一步, 对 θ 作非线性映射, 继续施加C-Y门、旋

图6 基于RUS电路的量子感知机, 修改自参考文献[30]
Figure 6 A RUS-based quantum perceptron, modified from ref. [30].

转门 $R_z(-\pi/2)$ 和受控旋转门 $C-R_y^\dagger(2\theta)$, 则系统的量子态为

$$|\psi\rangle = \frac{1+i}{\sqrt{2}} |x\rangle (\cos^2(\theta) |00\rangle + \sin^2(\theta) |01\rangle - \sin(\theta)\cos(\theta)(|10\rangle + |11\rangle)). \quad (14)$$

对辅助量子比特进行测量, 如果得到量子态 $|0\rangle$, 则系统量子态坍缩为

$$|\psi_0\rangle = |x\rangle |0\rangle (\cos^2(\theta) |0\rangle + \sin^2(\theta) |1\rangle), \quad (15)$$

等价于对输出量子比特施加了 $R_y(2q(\theta))$ 的么正变换, 其中 $q(\theta) = \arctan(\tan^2(\theta))$ 类似阈值激活函数. 如果测量得到量子态 $|1\rangle$, 则系统的量子态坍缩为

$$|\psi_1\rangle = |x\rangle |1\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle), \quad (16)$$

等价于对输出量子比特施加了 $R_y\left(-\frac{\pi}{2}\right)$ 的么正变换. 这样无法得到所需的非线性映射, 需要重复运行该电路直至得到期望量子态. 该量子电路可作进一步扩展以实现更高阶非线性函数:

$$q^{\circ k}(\theta) = \arctan(\tan^{2^k}(\theta)), k = 1, 2, \dots, \quad (17)$$

如图7所示, 从而具有更接近阈值函数的映射关系. 在该RUS电路中, 当输入信号 $\theta > \frac{\pi}{4}$ 时, 输出量子比特的状态接近 $R_y(\pi)|0\rangle = |1\rangle$; 当 $\theta < \frac{\pi}{4}$ 时, 输出量子比特状态接近 $R_y(0)|0\rangle = |0\rangle$, 实现了经典感知机中阈值激活函数的效果.

该量子感知机可组成前馈网络和Hopfield网络. 数值仿真显示, 所组成的量子前馈神经网络能够以叠加的形式学习训练数据, 实现经典神经网络无法比拟的训练效率, 体现量子计算并行性的优势. 该感知机结构也可实现其他激活函数, 如ReLU等, 并且适用于多种机器学习框架, 包括监督学习、无监督学习和强化学习等.

4.1.2 基于量子相位估计算法的量子感知机

基于量子相位估计(quantum phase estimation, QPE)算法^[21], 文献[32]提出一种实现阈值激活函数的量子感知机模型, 基本思想是将输入信号写到量子态

的概率幅上, 然后通过逆量子傅里叶变换以一定精度读取输入信号.

感知机的电路如图8所示, 包含 τ 位输出量子比特和 n 位输入量子比特. 输入限定为二值, 即 $x_i \in \{-1, 1\}$, n 维输入变量用 n 个量子比特编码为 $|x_1 x_2 \cdots x_n\rangle$, 其中 $|x_i\rangle = |1\rangle$ 表示 $x_i = 1$, 而 $|x_i\rangle = |0\rangle$ 表示 $x_i = -1$. 此外, n 个权重值被编码在么正变换参数中.

在量子电路中, 施加于 $|x_k\rangle$ 的么正变换定义为

$$U_k(w_k) = \begin{pmatrix} e^{-2\pi i w_k \Delta\phi} & 0 \\ 0 & e^{2\pi i w_k \Delta\phi} \end{pmatrix}, \quad (18)$$

其中, 实数 w_k 代表权值, $\Delta\phi = \frac{1}{2n}$. 施加所有受控么正变换后, 前 τ 位量子比特的量子态为

$$|\psi^\tau\rangle = \frac{1}{\sqrt{2}^\tau} \sum_{j=0}^{2^\tau-1} e^{2\pi i j \varphi} |j\rangle, \quad (19)$$

其中 $\varphi = \frac{\bar{x} \cdot \bar{w}}{2n}$ 为神经元输入信号的归一化值. 该量子

图7 近似的激活函数

Figure 7 An approximation activation function.

图8 量子相位估计神经网络, 修改自参考文献[32]

Figure 8 A QPE-based neural network, modified from ref. [32].

态与量子傅里叶变换后的结果在形式上高度近似.

类似量子相位估计算法, 施加逆量子傅里叶变换, 可在前 τ 位量子比特(输出量子比特)上得到以二进制形式编码的 φ 的近似值. 如果 φ 可表示为

$$\varphi = J_1 \frac{1}{2} + \cdots + J_\tau \frac{1}{2^\tau}, J_i \in \{0, 1\}, \quad (20)$$

则输出量子比特的值为

$$|j\rangle = |J_1 \cdots J_\tau\rangle. \quad (21)$$

对第一位量子比特 $|J_1\rangle$ 进行测量, 所得结果即反映输入信号是否大于0.5, 从而实现阈值函数. 如果 φ 不能用二进制编码准确表示, 则输出量子比特处于叠加态, 测量后有一定概率得到正确结果. 为提高算法的准确率, 可增加输出量子比特位数.

对于 n 维输入, 该量子感知机具有与经典神经元相近的 $O(n)$ 计算复杂度. 其优势在于, 利用量子计算的并行性, 神经元能够以叠加的方式同时处理所有训练数据. 在此基础上, 研究人员可以进一步发展基于叠加态的高效学习框架.

考虑到上述感知机模型中, 每个神经元所接受的输入是一维的, 而且取值限制为 $\{-1, 1\}$, 文献[33]提出了改进的多维输入的量子感知机模型, 如图9所示. 该感知机模型允许更复杂的输入数据, 每个输入 x_i 可以由多个量子比特编码, 如整数值的编码: $7 \mapsto |7\rangle = |111\rangle$. 它的量子电路结构与文献[32]类似, 但是所使用的受控量子门需要变换为式(22)的形式, 其中参数 ϕ 和 θ 均作为神经网络的权重. 同样的, 该模型通过测量最高

图9 多维输入的量子相位估计神经网络, 修改自参考文献[33]

Figure 9 A QPE-based neural network with multidimensional inputs, modified from ref. [33].

位输出量子比特的值实现阈值激活函数.

$$W_{n_k} = \begin{pmatrix} \frac{i\phi_n}{e^{2^k}} & 0 \\ 0 & \frac{i\theta_n}{e^{2^k}} \end{pmatrix}. \quad (22)$$

受文献[32]中实现阈值激活函数方法的启发, 文献[17]提出一种在量子电路中以一定精度近似任意非线性函数的量子感知机, 量子电路如图10所示. 在该模型中, 神经元的输入 $x_1, \dots, x_n$ 和权重 $w_1, \dots, w_n$ 为整数, 采用量子比特编码. 为便于描述, 设分别用 p, q 位量子比特编码每个输入值和权重值, 如 $|x_i\rangle = \underbrace{|0 \dots 1\rangle}_p$. 另外, τ 位输出量子比特用于存放输出结果.

设传播函数和激活函数分别为 $f(\cdot)$ 和 $g(\cdot)$, 则对于任意输入信号 $\bar{x}$ 及感知机权重 $\bar{w}$, 感知机的输出信号为 $g(f(\bar{x}, \bar{w}))$. 定义函数:

$$\Phi(\tau, g, f) = e^{(2\pi i/2^\tau)g(f(\bar{x}, \bar{w}))}, \quad (23)$$

构建么正算子 N :

$$N_\tau(p, q, g, f) = \sum_{x_1=0}^{2^p-1} \dots \sum_{x_n=0}^{2^p-1} \sum_{w_1=0}^{2^q-1} \dots \sum_{w_n=0}^{2^q-1} \Phi(\tau, g, f) |\bar{x}, \bar{w}\rangle \langle \bar{x}, \bar{w}|, \quad (24)$$

则感知机任意可能的输入向量 $(\bar{x}, \bar{w})$ 是么正算子 N 的特征向量, 相位为 $\frac{g(f(\bar{x}, \bar{w}))}{2^\tau}$, 通过量子相位估计算法可将相位值, 即感知机的对应运算结果, 提取到输出量子比特中.

这种量子电路结构并不依赖具体的非线性函数, 依据所需非线性映射关系构建么正矩阵, 即可在量子

感知机上执行相应运算. 值得注意的是, 感知机中包含 $m = n \cdot (p + q)$ 个编码输入和权重的量子比特, 对应的么正变换 U 包含 $O(2^m)$ 数量级的参数. 具有指数级参数的么正变换在实际中能否高效实现是该模型面临的重要问题.

前述量子感知机模型仅适用于基于量子比特编码的量子数据, 所需量子比特数目较大. 文献[34]提出适用于幅值编码量子数据的量子感知机, 通过swap-test电路[35]估计输入量子态 $|x\rangle$ 和权重量子态 $|w\rangle$ 的内积 $\langle x | w \rangle$, 并通过量子相位估计将该值存入量子态 $|u\rangle$. 进一步, 通过近似非线性函数的量子电路作非线性映射 $g(\cdot)$, 得到感知机的运算结果:

$$|u\rangle |0\rangle^{\otimes m} \mapsto |u\rangle |g(u)\rangle. \quad (25)$$

该模型能够模拟sigmoid, ReLU等多种激活函数, 并且所需量子比特和量子门的数量相比文献[17]具有指数级下降.

4.1.3 二值量子感知机

文献[36]提出一种二值的量子感知机并首先在真实的量子处理器上运行. 感知机的经典信息输入 $\bar{x} = (x_0, x_1, \dots, x_{n-1})$ 和经典权重 $\bar{w} = (w_0, w_1, \dots, w_{n-1})$ 均为二值向量 $\{-1, 1\}^n$. 感知机的量子电路结构如图11所示, 主要包括量子态制备电路和内积电路, 分别对应么正变换 U_x 和 U_w . 在运算过程中, 首先对初始量子态 $|0\rangle^{\otimes n}$ 施加么正变换 U_x , 将经典输入向量 $\bar{x}$ 编码到量子态中:

$$U_x |0\rangle^{\otimes n} = |\psi_x\rangle, \quad (26)$$

其中, $|\psi_x\rangle$ 满足

图 11 量子二值感知机模型, 修改自参考文献[36]

Figure 11 A quantum binary perceptron, modified from ref. [36].

图 10 实现任意激活函数的量子相位估计神经网络, 修改自参考文献[17]

Figure 10 A QPE-based neural network implementing any non-linear activation function, modified from ref. [17].

$$|\psi_x\rangle = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} x_j |j\rangle. \quad (27)$$

可见, U_x 仅要求矩阵第一列的值为向量 x , 这样的么正变换必定存在. 然后, 通过内积电路计算感知机输入向量和权重的内积. 定义权值量子态 $|\psi_w\rangle$:

$$|\psi_w\rangle = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} w_j |j\rangle, \quad (28)$$

以及么正变换 U_w , U_w 满足

$$U_w |\psi_w\rangle = |1\rangle^{\otimes N} = |n-1\rangle, \quad (29)$$

则施加么正变换 U_w 后, 系统量子态变为

$$U_w |\psi_x\rangle = \sum_{j=0}^{n-1} c_j |j\rangle = |\phi_{x,w}\rangle. \quad (30)$$

考虑输入向量 $|\psi_x\rangle$ 和权值 $|\psi_w\rangle$ 的内积, 如式(31), 该值是感知机的输入信号, 同时也是式(30)中第 n 项的系数.

$$\langle \psi_w | \psi_x \rangle = \langle \psi_w | U_w^\dagger U_w | \psi_x \rangle = \langle n-1 | \phi_{x,w} \rangle = c_{n-1}. \quad (31)$$

进一步对辅助量子比特 $|0\rangle_a$ 施加多受控非门, 则系统量子态演化为

$$|\phi_{x,w}\rangle |0\rangle_a \mapsto \sum_{j=0}^{n-2} c_j |j\rangle |0\rangle_a + c_{n-1} |n-1\rangle |1\rangle_a. \quad (32)$$

这样在辅助量子比特上进行测量, 将以 $|c_{n-1}|^2$ 的概率得到结果 $|1\rangle$, 从而实现感知机的非线性映射.

相比前述两类量子感知机, 该模型处理 n 维输入仅需 $N = \lceil \log_2 n \rceil$ 个量子比特, 能够更高效地利用硬件资源. 但是, 该模型采用的量子编码方式在处理大规模样本数据时会具有高昂的计算耗费. 针对这一问题, 文献[36]提出基于超图态的量子态制备方法提高制备效率. 网络训练采用量子-经典混合训练算法, 损失函数和权重更新的计算由经典计算机完成, 从而大大简化了量子神经网络的训练过程.

4.1.4 线性量子感知机

单个么正变换能够对输入信息做线性变换, 完成经典感知机中线性模型的功能. 对于一类较简单的量子机器学习任务, 如拟合未知么正变换^[37], 可以构建功

能上较简单的线性量子感知机, 以拟合么正映射关系.

文献[38]提出将么正算符作为量子感知机, 并提出相应的高效训练算法. 该量子感知机是作用于 $m+n$ 个量子比特的任意么正变换, 其中包含 m 个输入量子比特和 n 个输出量子比特, 分别存储输入和输出量子态. 通过调整么正变换参数, 该感知机可拟合输入与输出量子态间的线性映射关系.

多层的量子感知机构成量子前馈神经网络, 如图12. 其中感知机作了简化, 仅含1个输出量子比特. 对于具有 L 个隐藏层的量子神经网络, 其变换可由式(33)表示:

$$\rho^{\text{out}} = \text{tr}_{\text{in,hid}}(U(\rho^{\text{in}} \otimes |0 \cdots 0\rangle_{\text{hid,out}} \langle 0 \cdots 0|)U^\dagger). \quad (33)$$

输入量子比特的初始态为 ρ^{in} , 输出量子比特的初始态为 $|0 \cdots 0\rangle$. 经过么正变换, 得到输出量子比特的量子态 ρ^{out} . U 是该量子神经网络的么正变换, 如式(34):

$$U = U^{\text{out}} U^L U^{L-1} \cdots U^1. \quad (34)$$

其中, $U^l (l=1, 2, \dots, L)$ 表示第 l 层量子感知机的么正变换, 由作用在 $l-1$ 层到 l 层量子比特的么正变换组成, 该变换可进一步分解为单输出量子比特的么正变换的组合. 由于感知机是任意的么正变换, 通常是非互易的, 所以施加么正变换的顺序不可替代.

网络的训练算法是对反向传播算法的量子类比. 在量子神经网络的训练过程中需要评估学习效果, 即衡量网络对于输入量子态 $|\phi^{\text{in}}\rangle$ 的输出 ρ^{out} 与正确标签 $|\phi^{\text{out}}\rangle$ 的相近程度, 为此需要衡量两个量子态的相近程度. 以保真度作度量并在所有训练数据 $\{|\phi_i^{\text{in}}\rangle, |\phi_i^{\text{out}}\rangle\}$ ($i=1, \dots, N$)上取平均值, 定义损失函数如下:

图12 么正算符感知机, 修改自参考文献[38]

Figure 12 A perceptron of unitary operator, modified from ref. [38].

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^N \langle \phi_i^{\text{out}} | \rho_i^{\text{out}} | \phi_i^{\text{out}} \rangle. \quad (35)$$

训练过程中, 每次依据损失函数值迭代更新感知机的幺正矩阵: $U \rightarrow e^{i\varepsilon K} U$, 其中, ε 是设定的步长, 矩阵 K 的取值使得损失函数下降最快.

该量子感知机所组成的量子神经网络被应用于通过少量的训练数据学习未知的幺正变换, 数值实验验证其具有良好的泛化性能和噪声鲁棒性. 该量子神经网络减少了用于存储中间量子态的相干量子比特的数目, 仅需要网络宽度数量级的量子比特. 在即将到来的NISQ量子计算时代, 该量子神经网络模型对量子比特资源的高效利用使得其有望得到实际应用.

文献[39]提出对经典感知机进行量子泛化, 构建执行线性运算的量子感知机. 经典神经元运算的不可逆、非线性特性与量子门的幺正、线性特性之间存在矛盾. 为将经典神经元扩展为量子神经元, 一个自然的思路是先将其扩展为经典可逆门, 再扩展为幺正的量子门. 第一步, 通过添加辅助位使经典神经元可逆, 将神经元原本的输入输出关系 $(\text{in}_1, \dots, \text{in}_n) \rightarrow \text{out}$ 转化为 $(\text{in}_1, \dots, \text{in}_n, 0) \rightarrow (\text{in}_1, \dots, \text{in}_n, \text{out})$, 这实际上将不可逆映射转化为一个置换矩阵. 此时, 经典神经元的功能被限制为仅能进行线性运算. 第二步, 将该可逆运算转化为量子的幺正变换 $|\psi_{\text{in}}\rangle_{1,2,\dots,n} |0\rangle \rightarrow |\psi_{\text{in}}\rangle_{1,2,\dots,n,\text{out}}$. 相比置换矩阵, 幺正矩阵有着更强的表达能力. 此时泛化的经典神经网络就可以通过量子门来实现.

神经网络还涉及感知机输出的“扇出”, 即将一个神经元的输出传输到下一层的每个神经元. 一个简单的想法是通过幺正变换来实现 $b \rightarrow (b, b)$, 但是量子不可克隆原理指出这样的幺正操作无法实现. 文献[39]提出将CNOT门扩展为两量子比特的扇出幺正变换, 以实现 $|b\rangle |0\rangle \rightarrow U_F |b\rangle |0\rangle$. 这样的幺正变换实际上不是复制信息, 而是将信息由一个量子比特传输到多个量子比特, 该幺正变换的数值可以在训练中学习.

图13所示为具有 M 个输入量子比特和 N 个扇出量子比特的量子感知机, 其中, $U^{[M+1]}$ 表示接收 M 个输入的幺正量子门, 而 $U^{[N]}$ 表示对感知机计算结果进行扇出的扇出门. 数值实验中, 该量子感知机构建的量子自编码器能够较好地拟合量子态压缩的映射关系.

4.2 量子神经网络模型

面向近期量子处理器的量子神经网络, 其结构可一般性地表示为对输入量子态的一系列幺正变换和对部分输出量子比特的测量. 文献[40]提出一种通用量子神经网络框架, 如图14所示, 该模型可以对经典数据和量子数据进行监督学习. 这一量子神经网络模型采用一系列参数化的基础幺正变换, 作用在由输入量子比特 $|\psi\rangle$ 与辅助量子比特 $|1\rangle$ 共同组成的输入量子态 $|\psi, 1\rangle$ 上. 施加幺正变换序列 $U(\theta) = U_L(\theta_L)U_{L-1}(\theta_{L-1}) \cdots U_1(\theta_1)$ 后, 对输入量子比特进行 Y 测量, 结果为

$$I(\psi) = \langle \psi, 1 | U^\dagger(\theta) Y U(\theta) | \psi, 1 \rangle, \quad (36)$$

其取值为1或者-1. 取多次测量结果的平均值作为最终判别结果. 由于仅采用量子电路中的基础幺正变换, 此模型易于物理实现.

依据Reed-Muller表达式, 任何二进制函数可以表示为Reed-Muller项的模二加法的形式. 因此, 通过适当选取幺正变换参数, 该量子神经网络模型能够表达 n 维输入的所有布尔标签函数. 采用梯度下降训练

图13 感知机的量子泛化, 修改自参考文献[39]

Figure 13 Quantum generalization of the perceptron, modified from ref. [39].

图14 一般结构的量子神经网络, 修改自参考文献[40]

Figure 14 A quantum neural network with general structure, modified from ref. [40].

算法找到正确的么正变换参数, 该网络模型可逼近训练数据中隐藏的映射关系. 文献[40]提出如下损失函数:

$$\text{Loss}(\theta, \psi) = 1 - I(\psi) \langle \psi, 1 | U^\dagger(\theta) Y_{n+1} U(\theta) | \psi, 1 \rangle. \quad (37)$$

显然, 判别结果与标签值越一致, 损失函数越接近于0.

为了验证网络性能, 文献[40]在奇偶校验的二进制函数拟合、手写数字识别、量子态识别等任务上分别进行了数值实验, 证明了该网络可以通过随机梯度下降方式进行训练, 而且误差较小. 受限于经典计算机的模拟能力, 该文献的验证实验限于17位量子比特, 尽管并未显示出量子神经网络相比于经典神经网络的优势, 但其工作证明了量子神经网络可以用于对数据进行分类.

文献[41]中给出了表征任意二进制映射关系的另一类通用量子电路结构, 如图15所示. 这是一种仅使用多受控非门的通用量子电路, 可以通过调整参数学习任何二进制函数. 该网络由 2^n 个可调整的量子门 G_u 组成, 对于 n 维二进制输入, 每个门可取单位阵 I 或受控异或门 C_u , 所以有 2^{2^n} 种不同的电路配置. 对于给定的任意二进制函数 $f \in \mathbb{B}^n$ (从 $\mathbb{B}^n$ 到 $\mathbb{B}$ 的映射), 学习算法总能找到正确的电路配置来表达这个函数. 在训练中, 如果以叠加态的形式同时学习 2^n 种可能的输入, 最多通过 $n+1$ 次更新即可完成训练. 需要注意的是, 该量子电路中并不包含测量操作, 训练时需要将该电路接入其他判别电路, 对输出量子态进行识别.

尽管文献[40,41]中的量子神经网络由基础么正变换构建, 但在实际的量子计算机中, 量子比特的空间位置关系等现实条件对实际可执行的么正变换会造成限制. 文献[42]将约束条件下的量子神经网络建模为有向无环图 $G = (V, S)$, 其中么正变换作为图的节点, 么正变换间的量子链接作为图的边. 结合量子-经典混合信息处理框架, 该文献给出了量子神经网络的一般性信息通路构建算法.

上述文献^[40~42]对量子神经网络的一般性模型进行了描述, 但未考虑具体的神经网络结构, 以下将分别介绍几种典型结构的量子神经网络模型.

4.2.1 量子多层前馈神经网络

在文献[36]中, 作者提出一类简单的二值量子感知机. 进一步, 在文献[43]中, 作者将量子感知机扩展

为量子多层前馈神经网络, 并在IBM-Q超导量子比特计算机上进行实验. 文献[43]提出两种量子神经网络框架, 分别是量子-经典混合模型和完全量子模型. 量子-经典混合模型如图16(a)所示, 图中单线表示量子态信息, 双线表示测量得到的经典信息. 各个量子神经元独立完成由输入到输出的计算, 并通过对量子比特的测量实现非线性的阈值函数, 测量值直接送入下一层量子神经元. 完全量子模型如图16(b)所示, 则通过受控么正变换连接不同神经元, 在完成所有么正变换之后执行测量操作.

量子-经典混合模型对量子计算机的规模要求较低, 各量子神经元可以分别完成计算或者在同一量子比特上先后完成计算, 但是引入经典控制可能降低运行效率. 完全量子模型则要求所有量子比特保持相干, 在同一量子处理器上完成计算, 同时需要施加更大规模的么正变换. 两种模型各有优劣.

受限于当前量子计算技术, 该算法在IBM-Q量子计算机上运行时并不包含训练过程, 网络的权重是人工选定的, 实验中仅验证了量子神经网络的前向传播. 文献[43]指出该神经网络结构存在潜在的指数级的训练代价, 未来可以探索量子-经典混合训练算法, 如变

图15 通用二进制量子电路, 修改自参考文献[41]

Figure 15 A general boolean quantum circuit, modified from ref. [41].

图16 量子多层前馈网络, 修改自参考文献[43]. (a) 量子-经典混合前馈神经网络; (b) 完全量子前馈神经网络

Figure 16 Quantum feedforward neural networks, modified from ref. [43]. (a) A hybrid quantum-classical feedforward neural network; (b) a quantum coherent feedforward neural network.

分量子特征值求解算法.

4.2.2 量子张量神经网络

经典深度神经网络中的多层结构使其具有更强的表达能力. 与之相似, 量子力学中的张量网络也具有类似的分级结构, 能够提供量子态的高效近似表达并且简化量子电路描述. 最初, 学者们的研究集中于受张量网络启发的具备类似结构的机器学习算法. 矩阵积态(matrix product states, MPS)原本用于描述一维的纠缠量子态, 文献[44]提出利用MPS来参数化非线性核学习模型, 作为监督学习任务的优化方法, 在MNIST手写数字识别任务上进行了实验. 文献[45]提出基于MPS张量网络的无监督学习模型, 用于建模给定数据内在的概率分布. 文献[46]提出基于二维树张量网络(tree tensor network, TTN)结构的神经网络以及从多尺度纠缠重整拟设(multi-scale entanglement renormalization ansatz, MERA)导出的训练方法. 该算法用于分类手写数字图像, 以及将图像信息编码在量子多体态中. 类似经典的卷积神经网络, 在该TTN张量网络中也展现出了特征抽象逐层递增的特点.

文献[47]首次探索面向近期量子处理器的量子张量神经网络. 在文献[47]中, 作者提出一种具有张量网络结构的量子神经网络, 可用于判别任务和生成任务. 该神经网络模型具有固定的树状结构量子电路, 其中幺正变换的参数初始不固定, 训练算法为量子-经典混合算法, 即在经典算法的辅助下搜索合适的参数.

判别模型中, 神经网络的输入数据 $x = (x_1, x_2, \dots, x_n)$, $x_i \in [0, 1]$ 编码为 n 个量子比特的乘积态, 作为输入量子态, 如式(38). 该量子态的制备可以通过对基态 $|0\rangle^{\otimes n}$ 施加单量子比特旋转门来高效实现.

$$|x\rangle = \left[\begin{array}{c} \cos\left(\frac{\pi}{2}x_1\right) \\ \sin\left(\frac{\pi}{2}x_1\right) \end{array} \right] \otimes \left[\begin{array}{c} \cos\left(\frac{\pi}{2}x_2\right) \\ \sin\left(\frac{\pi}{2}x_2\right) \end{array} \right] \otimes \dots \otimes \left[\begin{array}{c} \cos\left(\frac{\pi}{2}x_n\right) \\ \sin\left(\frac{\pi}{2}x_n\right) \end{array} \right]. \quad (38)$$

判别模型的量子电路类似多层树状结构, 如图17(a)所示. 施加幺正变换后, 一半的量子比特被忽略, 其余量子比特继续参与下一层节点的运算. 最后, 一个或者多个量子比特作为输出量子比特, 最可能的测量结果

作为神经网络对输入数据的判别结果. 训练过程中, 采用经典算法对判别结果和真实标签进行比较, 根据误差更新电路参数. 数值试验结果显示, 判别模型能够较好地处理MNIST图像的二分类问题.

生成模型采用与判别模型几乎相反的结构, 如图17(b)所示. 新加入的量子比特与原有量子比特组合, 参与下层幺正变换节点的计算. 生成所需数量的量子比特后, 通过测量得到量子神经网络模型生成的数据信息. 训练网络时, 同样根据生成结果和标签的误差通过经典算法调整量子电路参数.

文献[47]指出该量子张量神经网络能够高效利用量子比特. 图18是一种硬件高效的生成网络实现方式, 它将参与计算后被忽略的量子比特重新初始化, 再加入后续运算过程, 极大地降低了对量子比特的数目要求.

4.2.3 量子生成对抗网络

经典的生成对抗网络包含两部分, 判别器和生成器, 其中生成器尽量模仿真实数据集的概率分布, 生成人工数据; 而判别器则尽量区分真实数据和生成器生成的人工数据. 量子生成对抗网络将两部分或者其中之一替换为量子神经网络, 利用量子计算天然的概

图 17 量子张量网络, 修改自参考文献[47]. (a) 判别式网络; (b) 生成式网络

Figure 17 Quantum tensor networks, modified from ref. [47]. (a) The discriminative network; (b) the generative network.

图 18 硬件高效的树张量网络, 修改自参考文献[47]

Figure 18 A qubit-efficient tree tensor network, modified from ref. [47].

率特性, 生成特定量子态或者符合一定概率分布的数据.

量子生成对抗学习最早由文献[48]进行了理论研究. 该文献分析了多种情形下量子生成对抗学习的运行效率, 包括训练数据为经典数据或者量子数据, 以及判别器和生成器是否采用量子处理器运行等情况, 指出在训练数据为量子数据的情况下, 量子对抗网络相比于经典对抗网络可能展现出指数级的优势.

文献[49]提出一种生成随机经典数据的量子生成对抗网络, 包括由参数化的量子门组成的量子生成器, 以及由经典前馈神经网络组成的经典判别器。量子生成器包含两类量子电路, 其中一类由多层幺正变换组成, 每层幺正变换由单量子比特旋转门和两量子比特受控相位门堆叠而成。该网络生成 N 维数据需要 N 位量子比特; 另一类则采用MPS张量网络结构, 生成数据的维数仅与网络深度相关, 适合在量子比特数受限的近期量子处理器上运行。该网络采用了量子-经典混合结构, 避免了输入输出的量子-经典数据转换瓶颈。其中, 经典判别器接收经典数据, 并对量子生成器生成的量子态施加测量来得到生成数据。相比于经典GAN, 该量子GAN的量子特性使其能够生成离散数据, 避免梯度消失问题。

文献[50]提出利用量子-经典混合的量子生成对抗网络来高效、近似地加载量子数据，包括量子生成网络和经典判别网络。量子生成网络采用变分子子电路的形式，包含重复的旋转门和受控Z门，通过对输出量子比特进行测量来得到输出数据。经典的生成网络通常需要引入随机输入来产生不同的输出，但是，该量子神经网络能够利用量子测量的天然随机性，对于固定的输入也产生不同的输出。判别网络为经典的人工神经网络，将生成网络的输出数据与经典数据作对比，通过误差更新网络参数。对于 n 位量子比特的量子态，精确制备需要 $O(2^n)$ 数量级的量子门，而采用该算法通过生成对抗网络来学习和加载数据背后的一般概率分布，仅需要 $O(\text{poly}(n))$ 数量级的量子门，具有潜在的量子优势。

文献[51]提出了量子亥姆霍兹机, 也借鉴了生成对抗式的网络结构. 经典亥姆霍兹机是一种特殊的人工神经网络, 可以通过训练来提取原始数据的底层结构, 并创建近似原始数据的生成模型. 经典亥姆霍兹机的学习算法依赖变分估计或者马尔科夫链蒙特卡罗

采样, 具有极高的计算复杂度. 但是, 量子计算机能够高效地实现采样, 所以量子版的亥姆霍兹机可以显著地降低训练过程的计算复杂度.

4.2.4 量子卷积神经网络

文献[52]提出一种基于量子电路的量子卷积神经网络QCNN，实现了类似经典卷积神经网络的卷积和池化变换，用于处理一维量子数据(而不是处理二维图像数据)。QCNN的量子电路结构如图19所示，包括重复的卷积层和池化层：卷积层施加有限深度的拟局部么正变换，同一卷积层内部各么正变换参数相同；池化层对部分量子比特进行测量，并在相邻量子比特上施加经典受控么正变换，么正变换的参数取决于邻近量子比特的测量结果。经过多层卷积与池化变换，在剩余量子比特较少时，施加作为全连接层的么正变换，并对指定的量子比特做测量，得到网络的判别结果。

该网络每个卷积层和池化层内部共享同一参数的幺正变换，所以对于 n 位输入量子比特，其仅具有 $O(\log(n))$ 数量级的参数，从而能够高效训练和在近期量子计算机上部署。相比于文献[40]中一般结构的量子电路，该网络模型的参数量双指数级下降。另外，池化层能够降低数据的维度并且通过部分测量引入一定的非线性。值得注意的是，随着卷积层的增加，卷积操作跨越的量子比特数也会增加，所以QCNN模型在量子计算机上的部署要求能够在各种距离上实施两量子比特门和投影测量。

类似于经典卷积神经网络, QCNN具有很强的可

图 19 (网络版彩图)量子卷积神经网络, 修改自参考文献 [52]

Figure 19 (Color online) A quantum convolutional neural network, modified from ref. [52].

扩展性,有望作为量子机器学习的通用基础框架.针对多种任务,QCNN可做不同扩展:图19所示的网络结构仅能处理一维量子数据,对于更高维输入量子态,可以探索处理高维输入的量子卷积神经网络结构;结合量子态编码技术,QCNN有望用于加速经典机器学习任务;针对更复杂的输入输出映射关系,可放宽平移不变性的限制,从而网络具有 $O(n\log(n))$ 数量级的参数;此外,也可探索其他网络训练方法,例如经典机器学习中的反向传播.

5 未来展望

现阶段,虽然大规模的通用量子计算机尚未真正实现,但是日趋成熟的近期量子处理器技术已为各类量子算法的简单验证提供了条件.近几年,获益于IBM等公司研发的商用量子计算机,研究工作者可以通过英特网远程操纵十数个量子比特,搭建简单的量子电路,实现小规模量子网络系统.在作者看来,这一方面为各类量子神经网络模型与学习算法提供了简单的实验验证平台,另一方面也为量子神经网络理论研究规范了严格的体系框架:量子神经网络模型及其学习算法必须面向真实的量子电路,在量子力学体系下进行严格设计.在此种意义下,量子神经网络的研究工作仍然任重而道远,如下数个关键的科学问题亟待进一步的研究.

(1) 量子神经网络与经典神经网络的科学比较问题.受限于当前量子计算硬件,量子神经网络仅能针对低维度小样本问题进行实验,难以验证相比经典神经网络的优势.针对这一关键问题,一方面,需要建立统一的定量指标和计算模型,对量子神经网络和经典神经网络的运行复杂度和资源需求进行精确比较,严格证明量子计算相比经典计算的优越性.另一方面,

需要在大型基准数据集上严格验证量子神经网络的预测精度和泛化性能是否具有优势.目前,相关研究较少.文献[53~55]对面向近期量子处理器的量子优化算法相较于经典优化算法的优越性做了深入探讨,具有重要的参考价值.

(2) 中等规模量子神经网络的训练机制.基于近期量子处理器的中等规模量子神经网络参数训练面临着数个本质困难.一方面,量子神经网络指数量级的参数空间容易导致“贫瘠高原”现象^[56,57],这使得参数训练容易陷入局部极小值,造成欠拟合现象;另一方面,由于量子系统的特殊性,实际中的训练算法的迭代步骤多依赖于量子观测期望的精确估计,需要进行大量的重复性量子测量,提高了计算复杂度.基于量子最优控制理论^[58,59],研究基于量子电路采样的量子-经典混合优化框架^[60,61],是发展高性能近似优化训练算法的一种重要的研究思路.此外,从软件和算法方面,需要进一步研究量子神经网络的可学习性^[62],近似能力^[63~65],算法收敛性等.

(3) 近期量子处理器的硬件缺陷.近期量子处理器普遍存在噪音大、难控制等缺陷,这给量子神经网络的稳定可靠运行造成了本质困难.从控制理论的角度出发,对噪音及控制约束对神经网络设计和精度的影响进行建模和分析^[66,67],并研究噪音干扰下量子神经网络的鲁棒控制问题^[68,69],是一种可能的解决途径.

此外,作为一种特殊的复杂网络系统^[70,71],量子神经网络也可以被用来模拟在实际中广泛存在的具有网络结构特性的量子系统,解决一些具有挑战性的量子物理问题.经典神经网络辅助量子问题的求解已被广泛探索,比如量子层析^[72~74]和量子系统仿真^[75,76]等.对于具有更高维度和更复杂网络结构的量子系统,如何借助量子神经网络强大的计算智能来求解相应的物理难题,具有重要的科学价值.

参考文献

- 1 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th Advances in Neural Information Processing Systems. Lake Tahoe, 2012. 1097–1105
- 2 Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vis*, 2015, 115: 211–252
- 3 He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016. 770–778
- 4 Devlin J, Chang M W, Lee K, et al. Bert: pre-training of deep bidirectional transformers for language understanding. arXiv: [1810.04805](https://arxiv.org/abs/1810.04805)
- 5 Litjens G, Kooi T, Bejnordi B E, et al. A survey on deep learning in medical image analysis. *Med Image Anal*, 2017, 42: 60–88

- 6 Nielsen M A. Neural Networks and Deep Learning. San Francisco: Determination Press, 2015. 3–11
- 7 Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks*, 1989, 2: 359–366
- 8 Biamonte J, Wittek P, Pancotti N, et al. Quantum machine learning. *Nature*, 2017, 549: 195–202
- 9 Huang Y M, Lei H, Li X Y. A survey on quantum machine learning (in Chinese). *Chin J Comput*, 2018, 41: 145–163 [黄一鸣, 雷航, 李晓瑜. 量子机器学习算法综述. *计算机学报*, 2018, 41: 145–163]
- 10 Lu S C, Zheng Y, Wang X T, et al. Quantum machine learning (in Chinese). *Control Theory Appl*, 2017, 34: 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: 1429–1436]
- 11 Lloyd S, Mohseni M, Rebentrost P. Quantum principal component analysis. *Nat Phys*, 2014, 10: 631–633
- 12 Dong D Y, Chen C L, Li H X, et al. Quantum reinforcement learning. *IEEE Trans Syst Man Cybern B*, 2008, 38: 1207–1220
- 13 Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big data classification. *Phys Rev Lett*, 2014, 113: 130503
- 14 Feynman R P. Simulating physics with computers. *Int J Theor Phys*, 1982, 21: 467–488
- 15 Shor P W. Algorithms for quantum computation: Discrete logarithms and factoring. In: *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*. Santa Fe, 1994. 124–134
- 16 Grover L K. A fast quantum mechanical algorithm for database search. In: *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*. Philadelphia Pennsylvania, 1996. 212–219
- 17 de Paula Neto F M, Ludermir T B, de Oliveira W R, et al. Implementing any nonlinear quantum neuron. *IEEE Trans Neural Netw Learning Syst*, 2020, 31: 3741–3746
- 18 Kak S C. Quantum neural computing. *Adv Imaging Electron Phys*, 1995, 94: 259–313
- 19 Schuld M, Sinayskiy I, Petruccione F. The quest for a quantum neural network. *Quantum Inf Process*, 2014, 13: 2567–2586
- 20 Arute F, Arya K, Babbush R, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 2019, 574: 505–510
- 21 Córcoles A D, Kandala A, Javadi-Abhari A, et al. Challenges and opportunities of near-term quantum computing systems. arXiv: [1910.02894](https://arxiv.org/abs/1910.02894)
- 22 Preskill J. Quantum computing in the NISQ era and beyond. *Quantum*, 2018, 2: 79
- 23 Chung I L, Nielsen M A. Quantum Computing and Quantum Information. New York: Cambridge University Press, 2000. 61–73
- 24 Deutsch D. Quantum computational networks. *Proc R Soc Lond A*, 1989, 425: 73–90
- 25 Menneer T, Narayanan A. Quantum-inspired neural networks. Technical Report. Department of Computer Science, University of Exeter, 1995. 329
- 26 Behrman E C, Nash L R, Steck J E, et al. Simulations of quantum neural networks. *Inf Sci*, 2000, 128: 257–269
- 27 Gupta S, Zia R K P. Quantum neural networks. *J Comput Syst Sci*, 2001, 63: 355–383
- 28 Altaisky M V. Quantum neural network. arXiv: [quant-ph/0107012](https://arxiv.org/abs/quant-ph/0107012)
- 29 Killoran N, Bromley T R, Arrazola J M, et al. Continuous-variable quantum neural networks. *Phys Rev Res*, 2019, 1: 033063
- 30 Cao Y D, Guerreschi G G, Aspuru-Guzik A. Quantum neuron: An elementary building block for machine learning on quantum computers. arXiv: [1711.11240](https://arxiv.org/abs/1711.11240)
- 31 Paetzniack A, Svore K M. Repeat-until-Success: Non-deterministic decomposition of single-qubit unitaries. *Quantum Inf Comput*, 2014, 14: 1277–1301
- 32 Schuld M, Sinayskiy I, Petruccione F. Simulating a perceptron on a quantum computer. *Phys Lett A*, 2015, 379: 660–663
- 33 Yamamoto A Y, Sundqvist K M, Li P, et al. Simulation of a multidimensional input quantum perceptron. *Quantum Inf Process*, 2018, 17: 128
- 34 Yan S, Qi H, Cui W. Nonlinear quantum neuron: A fundamental building block for quantum neural networks. *Phys Rev A*, 2020, 102: 052421
- 35 Zhao J, Zhang Y H, Shao C P, et al. Building quantum neural networks based on a swap test. *Phys Rev A*, 2019, 100: 012334
- 36 Tacchino F, Macchiavello C, Gerace D, et al. An artificial neuron implemented on an actual quantum processor. *npj Quantum Inf*, 2019, 5: 26
- 37 Verdon G, Pye J, Broughton M. A universal training algorithm for quantum deep learning. arXiv: [1806.09729](https://arxiv.org/abs/1806.09729)
- 38 Beer K, Bondarenko D, Farrelly T, et al. Training deep quantum neural networks. *Nat Commun*, 2020, 11: 808
- 39 Wan K H, Dahlsten O, Kristjánsson H, et al. Quantum generalisation of feedforward neural networks. *npj Quantum Inf*, 2017, 3: 36
- 40 Farhi E, Neven H. Classification with quantum neural networks on near term processors. arXiv: [1802.06002](https://arxiv.org/abs/1802.06002)
- 41 Ngoc V P, Wiklicky H. Tunable quantum neural networks for boolean functions. arXiv: [2003.14122](https://arxiv.org/abs/2003.14122)
- 42 Gyongyosi L, Imre S. Training optimization for gate-model quantum neural networks. *Sci Rep*, 2019, 9: 12679
- 43 Tacchino F, Barkoutsos P, Macchiavello C, et al. Quantum implementation of an artificial feed-forward neural network. arXiv: [1912.12486](https://arxiv.org/abs/1912.12486)

- 44 Stoudenmire E, Schwab D J. Supervised learning with tensor networks. In: Proceedings of the 29th Advances in Neural Information Processing Systems. Barcelona, 2016. 4799–4807
- 45 Han Z Y, Wang J, Fan H, et al. Unsupervised generative modeling using matrix product states. *Phys Rev X*, 2018, 8: 031012
- 46 Liu D, Ran S J, Wittek P, et al. Machine learning by unitary tensor network of hierarchical tree structure. *New J Phys*, 2019, 21: 073059
- 47 Huggins W, Patil P, Mitchell B, et al. Towards quantum machine learning with tensor networks. *Quantum Sci Technol*, 2019, 4: 024001
- 48 Lloyd S, Weedbrook C. Quantum generative adversarial learning. *Phys Rev Lett*, 2018, 121: 040502
- 49 Situ H Z, He Z M, Wang Y Y, et al. Quantum generative adversarial network for generating discrete data. 2018, arXiv: [1807.01235](#)
- 50 Zoufal C, Lucchi A, Woerner S. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Inf*, 2019, 5: 103
- 51 van Dam T J, Neumann N M P, Phillipson F, et al. Hybrid Helmholtz machines: A gate-based quantum circuit implementation. *Quantum Inf Process*, 2020, 19: 174
- 52 Cong I, Choi S, Lukin M D. Quantum convolutional neural networks. *Nat Phys*, 2019, 15: 1273–1278
- 53 Farhi E, Harrow A W. Quantum supremacy through the quantum approximate optimization algorithm. arXiv: [1602.07674](#)
- 54 Guerreschi G G, Matsuura A Y. QAOA for Max-Cut requires hundreds of qubits for quantum speed-up. *Sci Rep*, 2019, 9: 6903
- 55 Zhou L, Wang S T, Choi S, et al. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. arXiv: [1812.01041](#)
- 56 McClean J R, Boixo S, Smelyanskiy V N, et al. Barren plateaus in quantum neural network training landscapes. *Nat Commun*, 2018, 9: 4812
- 57 Cerezo M, Sone A, Volkoff T, et al. Cost-function-dependent barren plateaus in shallow quantum neural networks. arXiv: [2001.00550](#)
- 58 Li J, Yang X, Peng X, et al. Hybrid quantum-classical approach to quantum optimal control. *Phys Rev Lett*, 2017, 118: 150503
- 59 Choquette A, Di Paolo A, Barkoutsos P K, et al. Quantum-optimal-control-inspired ansatz for variational quantum algorithms. arXiv: [2008.01098](#)
- 60 Harrow A, Napp J. Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms. arXiv: [1901.05374](#)
- 61 Sweke R, Wilde F, Meyer J J, et al. Stochastic gradient descent for hybrid quantum-classical optimization. *Quantum*, 2020, 4: 314
- 62 Du Y X, Hsieh M H, Liu T L, et al. On the learnability of quantum neural networks. arXiv: [2007.12369](#)
- 63 Hubregtsen T, Pichlmeier J, Bertels K. Evaluation of parameterized quantum circuits: On the design, and the relation between classification accuracy, expressibility and entangling capability. arXiv: [2003.09887](#)
- 64 Nakaji K, Yamamoto N. Expressibility of the alternating layered ansatz for quantum computation. arXiv: [2005.12537](#)
- 65 Sim S, Johnson P D, Aspuru-Guzik A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Adv Quantum Tech*, 2019, 2: 1900070
- 66 Gu H B, Lü J H, Lin Z L. On PID control for synchronization of complex dynamical network with delayed nodes. *Sci China Tech Sci*, 2019, 62: 1412–1422
- 67 Gentini L, Cuccoli A, Pirandola S, et al. Noise-assisted variational hybrid quantum-classical optimization. arXiv: [1912.06744](#)
- 68 Dong D, Xing X, Ma H, et al. Learning-based quantum robust control: Algorithm, applications, and experiments. *IEEE Trans Cybern*, 2020, 50: 3581–3593
- 69 Kosut R L, Grace M D, Brif C. Robust control of quantum gates via sequential convex programming. *Phys Rev A*, 2013, 88: 052326
- 70 Wang X, Gu H B, Wang Q Y, et al. Identifying topologies and system parameters of uncertain time-varying delayed complex networks. *Sci China Tech Sci*, 2019, 62: 94–105
- 71 Wang Y F, Wu X Q, Feng H, et al. Topology inference of uncertain complex dynamical networks and its applications in hidden nodes detection. *Sci China Tech Sci*, 2016, 59: 1232–1243
- 72 Qi B, Hou Z, Li L, et al. Quantum state tomography via linear regression estimation. *Sci Rep*, 2013, 3: 3496
- 73 Palmieri A M, Kovlakov E, Bianchi F, et al. Experimental neural network enhanced quantum tomography. *npj Quantum Inf*, 2020, 6: 20
- 74 Torlai G, Mazzola G, Carrasquilla J, et al. Neural-network quantum state tomography. *Nat Phys*, 2018, 14: 447–450
- 75 Gao X, Duan L M. Efficient representation of quantum many-body states with deep neural networks. *Nat Commun*, 2017, 8: 662
- 76 Torlai G, Mazzola G, Carleo G, et al. Precise measurement of quantum observables with neural-network estimators. *Phys Rev Res*, 2020, 2: 022060

Recent advances of quantum neural networks on the near term quantum processor

LÜ YanXuan¹, GAO Qing^{1,2}, LÜ JinHu^{1,2}, PAN Yu³ & DONG DaoYi⁴

¹ School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;

² Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing 100191, China;

³ Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China;

⁴ School of Engineering and Information Technology, University of New South Wales, Canberra 2600, Australia

In recent years, artificial neural networks have been widely used in various fields and have shown strong computational intelligence. At the same time, quantum computing hardware has been developed rapidly. In particular, near term quantum processors have relatively stable computing capabilities and robustness against decoherence. With the aid of this powerful quantum technique, nowadays many commercial cloud quantum computing companies can provide online quantum computing experimental platforms for scholars around the world. Various technical developments have promoted the emergence of many new research ideas and experimental progress in the cross-field of quantum neural networks. In this paper, the research ideas of early quantum neural networks are summarized first. The data encoding, algorithm flows, and quantum circuit design of quantum neural networks based on recent quantum processors are then reviewed and analyzed. Finally, several key scientific problems in the field of quantum neural networks are summarized, and prospective future research directions are presented.

artificial intelligence, quantum neural network, near term quantum processor

doi: [10.1360/SST-2020-0459](https://doi.org/10.1360/SST-2020-0459)