

## 基于参数化量子线路的量子神经网络数据分类

陈超, 闫文杰, 薛桂香

引用本文

陈超, 闫文杰, 薛桂香. 基于参数化量子线路的量子神经网络数据分类[J]. 计算机科学, 2024, 51(11A): 231200112-7.

CHEN Chao, YAN Wenjie, XUE Guixiang. [Parameterized Quantum Circuits Based Quantum Neural Networks for Data Classification](#) [J]. Computer Science, 2024, 51(11A): 231200112-7.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

**Similar articles recommended (Please use Firefox or IE to view the article)**

[基于3D骨架相似性的自适应移位图卷积神经网络人体行为识别算法](#)

Human Action Recognition Algorithm Based on Adaptive Shifted Graph Convolutional Neural Network with 3D Skeleton Similarity

计算机科学, 2024, 51(4): 236-242. <https://doi.org/10.11896/jsjcx.221200120>

[量子原型聚类](#)

Quantum Prototype Clustering

计算机科学, 2023, 50(8): 27-36. <https://doi.org/10.11896/jsjcx.220600124>

[基于Grover算法的图着色问题求解](#)

Solving Graph Coloring Problem Based on Grover Algorithm

计算机科学, 2023, 50(6): 351-357. <https://doi.org/10.11896/jsjcx.220400051>

[基于后状态强化学习的最优订单接受决策](#)

Optimal Order Acceptance Decision Based on After-state Reinforcement Learning

计算机科学, 2022, 49(11A): 210800261-9. <https://doi.org/10.11896/jsjcx.210800261>

[Shor整数分解算法的线路优化](#)

Optimization for Shor's Integer Factorization Algorithm Circuit

计算机科学, 2022, 49(6A): 649-653. <https://doi.org/10.11896/jsjcx.210600149>

# 基于参数化量子线路的量子神经网络数据分类

陈超 闫文杰 薛桂香

河北工业大学人工智能与数据科学学院 天津 300401

(chen\_enda@foxmail.com)

**摘要** 量子神经网络结合了量子计算与经典神经网络模型的优势,为未来人工智能领域的发展提供了一种全新的思路。尽管量子神经网络已被广泛研究,但数据编码方式以及不同训练线路对模型性能的影响尚未得到充分研究。为此,提出一种面向数据分类的量子神经网络新模型,该模型探究了不同数据编码方式和不同结构训练层对分类任务的影响。该方法首先对经典图像进行预处理,采用不同的数据编码方式将其编码到不同的参数化量子线路中进行训练,对模型的输出进行测量,使用参数移位规则更新训练参数完成数据分类。在 MNIST 手写体数据集上的实验结果表明,所提出的模型在数字 {3,6} 分类任务上的分类准确率超过了 97%。与目前的主流方法相比,所提方法在分类准确率上有明显的提升。

**关键词:**量子神经网络;量子计算;数据分类;量子线路;参数移位规则

中图分类号 TP391

## Parameterized Quantum Circuits Based Quantum Neural Networks for Data Classification

CHEN Chao, YAN Wenjie and XUE Guixiang

School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China

**Abstract** Quantum neural network combines the advantages of quantum computing and classical neural network model, and provides a new idea for the development of artificial intelligence in the future. Although quantum neural networks have been widely studied, the impact of data encoding methods and different training circuits on model performance has not yet been fully explored. Therefore, this paper proposes a new quantum neural network model for data classification, which explores the influence of different data encoding methods and different structure training layers on classification tasks. The method first preprocesses the classical image, uses different data encoding methods to encode it into different parameterized quantum circuits for training, measures the output of the model, and uses the parameter shift rule to update the training parameters to complete the data classification. Experimental results on MNIST handwritten dataset show that the proposed model achieves more than 97% classification accuracy on digit {3,6} classification task. Compared with the current mainstream methods, the proposed method has a significant improvement in classification accuracy.

**Keywords** Quantum neural networks, Quantum computing, Data classification, Quantum circuit, Parameter shift rules

## 1 引言

近年来,深度神经网络<sup>[1]</sup>已经成为主流的机器学习模型,在图像分类、目标检测、自然语言处理等应用中取得了成功。其成功的关键是通过加深网络来提高预测精度。然而,存储和计算需求的急剧增加成为经典计算机的瓶颈<sup>[2]</sup>。量子计算具有强大的并行计算能力和希尔伯特空间巨大的存储能力<sup>[3]</sup>,可以作为深度神经网络的量子加速器,有望更高效地解决这些问题。量子神经网络<sup>[4-6]</sup>是一种包含可训练参数的量子线路,是用于机器学习任务的流行算法之一,有望在现有的量子器件上实现量子优势<sup>[7]</sup>。实现参数化量子线路<sup>[8-9]</sup>的量子优势需要足够大的量子比特数量。然而,大量的量子比特被用于实际量子计算机噪声的处理,实际可用于量子线路的量子比特数更少<sup>[10]</sup>。所以,在资源有限的量子比特数下设计量子线路是有必要的。

在文献[11]工作的基础上,本文提出一种基于参数化量子线路的量子神经网络模型。该模型主要有 5 个组成部分:数据准备、数据编码、酉变换、纠缠态测量和经典后处理。首先,通过主成分分析(Principal Component Analysis, PCA)和归一化对 MNIST 数据集进行预处理。然后,将数据编码为量子态。本文讨论了数据在不同编码方式如角度编码、时间演化编码和振幅编码下模型的泛化性能。接着,构建参数化量子线路模型,包括纠缠层<sup>[12]</sup>和叠加层<sup>[13]</sup>。最后,对量子位进行测量,输出的结果采用参数移位规则<sup>[14]</sup>来优化训练参数,使得模型预测值更接近标签值。

本文采用 3 种不同的编码方式将经典数据映射到量子线路,探究了不同数据编码方式对模型泛化性能的影响;设计了 4 种不同结构的强纠缠特性的参数化量子线路模型作为训练参数层,探究了不同量子门电路组合的线路对模型学习能力的影响。实验结果表明,该参数化量子线路模型在较少的量

基金项目:国家自然科学基金(61702157)

This work was supported by the National Natural Science Foundation of China(61702157).

通信作者:薛桂香(xueguixiang@hebut.edu.cn)

子比特数下达到了更高的分类精度。

## 2 量子计算

量子计算的基本单位是量子比特。相对于经典计算,量子计算的优势在于其叠加性<sup>[12]</sup>和纠缠性<sup>[13]</sup>。

### 2.1 量子比特

量子比特是量子计算的基本单元,具有两个基态,分别为 $|0\rangle$ 和 $|1\rangle$ 。与经典比特只能处于0或1的状态不同,量子位可以处于 $|0\rangle$ 和 $|1\rangle$ 的某种组合的任何状态,称为量子叠加态,如式(1)所示。

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (1)$$

其中,复数 $\alpha$ 和 $\beta$ 分别表示基态 $|0\rangle$ 和 $|1\rangle$ 的概率幅,满足 $|\alpha|^2 + |\beta|^2 = 1$ 。

对于由 $n$ 量子比特构成的量子系统 $|\phi\rangle$ ,其量子态空间是各个量子态空间的张量积,可表示为:

$$|\phi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle \quad (2)$$

### 2.2 量子门电路

量子门电路是通过么正变换来改变量子状态的,是量子计算的基础。

#### 2.2.1 Hadamard 门

Hadamard 门可以将一个量子位 $|0\rangle$ 或 $|1\rangle$ 变为 $|0\rangle + |1\rangle$ 或 $|0\rangle - |1\rangle$ 的均衡叠加态。

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \quad (3)$$

$$H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \quad (4)$$

#### 2.2.2 量子旋转门

量子旋转门作用于量子比特上,可以让相应的量子比特在 Bloch 球面上沿着 $\hat{x}$ ,  $\hat{y}$ ,  $\hat{z}$ 轴进行旋转,改变量子态的相位。常见的量子旋转门有 $R_x$ ,  $R_y$ ,  $R_z$ 。单比特量子旋转门的数学描述为:

$$R_x(\theta) = \begin{bmatrix} \cos(\theta/2) & -i \sin(\theta/2) \\ -i \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \quad (5)$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \quad (6)$$

$$R_z(\theta) = \begin{bmatrix} e^{-i(\theta/2)} & 0 \\ 0 & e^{i(\theta/2)} \end{bmatrix} \quad (7)$$

#### 2.2.3 受控非门

受控非门(C Controlled-NOT, CNOT)是一种量子逻辑门,如图1所示,当控制比特位处于 $|1\rangle$ 时,它会对目标位取反。CNOT 门可以用来创建不同量子位之间的纠缠态<sup>[15]</sup>。

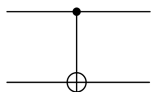


图1 受控非门示意图

Fig. 1 Diagram of controlled NOT gate

### 2.3 量子测量

为了获取量子线路中的信息,需要经过测量将量子态信息转化为经典可观测的信息。测量算符如图2所示。

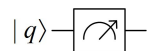


图2 测量的线路表示

Fig. 2 Circuit representation of measurement

在单量子比特 $q$ 的系统中,量子系统可以表示为 $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ 。对量子位 $q$ 进行 Pauli-Z 基测量,对应的测量算子的集合可以表示为 $M_0 = |0\rangle\langle 0|$ ,  $M_1 = |1\rangle\langle 1|$ ,可得到量子位取0或1的概率为:

$$p(0) = \langle \psi | M_0^\dagger M_0 | \psi \rangle = |\alpha|^2 \quad (8)$$

$$p(1) = \langle \psi | M_1^\dagger M_1 | \psi \rangle = |\beta|^2 \quad (9)$$

## 3 量子神经网络数字识别模型设计

量子神经网络是一种结合了量子计算和经典神经网络思想的模型,用于处理和学习量子态数据<sup>[16]</sup>。在分类任务中,量子神经网络可以用于对输入数据进行分类,并输出相应的类别标签。图3是基于参数化量子线路的量子神经网络模型示意图,该量子神经网络包括数据准备、数据嵌入到量子神经网络、神经网络训练参数层的结构设计、优化算法的设计以及分类预测。

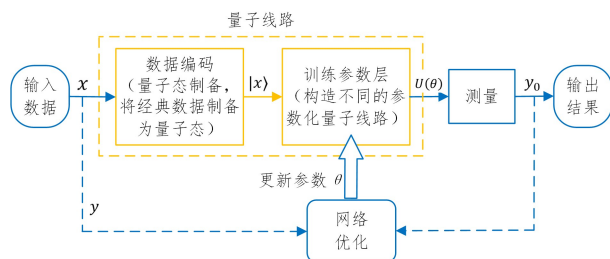


图3 参数化量子线路的量子神经网络模型示意图

Fig. 3 Schematic diagram of quantum neural network model with parameterized quantum circuits

### 3.1 数据编码

经典数据编码到希尔伯特空间的量子态上,以嵌入到参数化量子线路中。将经典数据编码到量子态的方式有角度编码、时间演化编码、振幅编码等。

#### 3.1.1 角度编码

角度编码是将 $N$ 个特征编码进 $n$ 个量子比特的旋转角度中,属于一对一编码。当 $N \leq n$ 时,输入数据 $x = \{x_1, \dots, x_N\}$ 编码成量子态可以表示为:

$$|x\rangle = \bigotimes_{i=1}^N \cos(x_i) |0\rangle + \sin(x_i) |1\rangle \quad (10)$$

这种编码方式一个量子位一次只能编码一个恒定深度的量子线路。例如,将数据 $x = \{x_0, x_1, x_2, x_3\}$  4个维度的数据点使用角度编码的示意图,如图4所示。

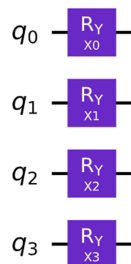


图4 角度编码示意图

Fig. 4 Diagram of angle encoding

#### 3.1.2 时间演化编码

在量子计算中,时间演化可以用 Schrödinger 方程<sup>[17]</sup>来描述:

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle \quad (11)$$

$$U(t) = e^{-i\frac{t}{\hbar}H} \quad (12)$$

其中,  $H$  是一个特殊的可观测量,称为系统的哈密顿量(Hamilton);  $\hbar$  是普朗克常量;时间演化算符为  $U$ 。

时间演化编码是将经典标量  $x \in \mathbb{R}$  与酉矩阵中关于时间的  $t$  联系起来,即  $U(x) = e^{-ixH}$ 。哈密顿量  $H = \frac{1}{2} \delta_i$ , 其中  $\delta_i \in \{\delta_x, \delta_y, \delta_z\}$  为泡利旋转门。其矩阵表示为:

$$\delta_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \delta_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \delta_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (13)$$

时间演化编码的基本思想是利用量子系统的哈密顿量来驱动量子比特的时间演化。通过调整哈密顿量的形式和参数,可以实现对数据的编码。对于经典的数据  $x_i$ , 执行时间演化编码完成  $R_x$  门的特征映射:

$$\begin{aligned} R_x(x_i) |0\rangle &= e^{-i\frac{x_i}{2}\delta_x} |0\rangle \\ &= \begin{bmatrix} \cos(\theta/2) & i \sin(\theta/2) \\ -i \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} |0\rangle \\ &= \cos \frac{\theta}{2} |0\rangle - i \sin \frac{\theta}{2} |1\rangle \end{aligned} \quad (14)$$

一般使用  $N$  量子比特编码  $2^N$  的输入数据,例如归一化的输入数据  $x = [x_0, x_1, \dots, x_{14}, x_{15}]$ , 采用时间演化编码的方式编码到 4 个量子比特,其结构如图 5 所示。

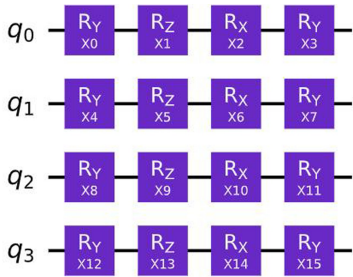


图 5 时间演化编码示意图

Fig. 5 Diagram of time evolution encoding

$$(CNOT) \cdot \left( I \otimes R_y \left( \frac{\pi}{2} \right) \right) \cdot (CNOT) \cdot \left( I \otimes R_y \left( \frac{\pi}{2} \right) \right) \cdot (|0\rangle \otimes |0\rangle)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot$$

$$\left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix} \right) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (18)$$

### 3.2 训练参数层结构设计

对于  $n$  个量子比特的量子系统,在 Ansatz 层(量子线路的一种假设)之前量子态表示为:

$$|\phi(x)\rangle = U(x) |0\rangle^{\otimes n} \quad (19)$$

其中,  $U(x)$  表示经典数据编码为量子态的酉矩阵。

这一节将设计不同结构的 Ansatz 以探寻一种最优的参数化量子线路作为数据分类的新模型,如图 7 所示的 4 种不同的线路结构。整个量子线路 Ansatz 可以表示为一个酉矩阵  $U(\theta)$ 。在量子线路的最后,使用一个 CCCNOT 门,前 3 个量子位为控制位,最后一个量子位为目标位。当控制比特位全为  $|1\rangle$  时,目标位进行取反。对最后一个量子位进行测量,该电路测量的期望值用函数表示为:

对于  $q_0, q_1, q_2, q_3$  这 4 量子比特编码 16 个数据  $x$  可以表示为:

$$\begin{aligned} U(X) &= R_y(x_0) R_z(x_1) R_x(x_2) R_y(x_3) \otimes R_y(x_4) R_z(x_5) \\ &R_x(x_6) R_y(x_7) \otimes R_y(x_8) R_z(x_9) R_x(x_{10}) R_y \\ &(x_{11}) \otimes R_y(x_{12}) R_z(x_{13}) R_x(x_{14}) R_y(x_{15}) \end{aligned} \quad (15)$$

#### 3.1.3 振幅编码

振幅编码是基于量子叠加态原理和幺正演化的特性。假设有  $n$  个量子比特,每个量子比特可以处于两个状态 0 和 1, 总共有  $2^n$  个可能的状态。具体来说,对于  $N$  维归一化的经典数据  $x$ , 一个  $n$  量子比特量子态的振幅  $|\phi_x\rangle$  可以表示为:

$$|\phi(x)\rangle = \sum_{i=1}^N x_i |i\rangle \quad (16)$$

其中,  $N=2^n$ ,  $x_i$  是  $x$  的第  $i$  元素,  $|i\rangle$  是第  $i$  个计算基态。

一个  $n$  比特的量子系统可以用一个复数数组表示其振幅,数组长度为  $2^n$ , 每个数组元素表示对应状态的振幅值。给定一个向量  $x = (x_0, x_1, \dots, x_{2^n-1})$ , 使用振幅编码是将这  $2^n$  的向量编码到  $n$  量子比特量子态的振幅上,如式(17)所示:

$$|\phi\rangle = x_0 |00\dots 0\rangle + x_1 |00\dots 1\rangle + \dots + x_{2^n-1} |11\dots 1\rangle \quad (17)$$

其中,  $x_i \in \mathbb{R}$ , 并且满足  $|x_0|^2 + |x_1|^2 + \dots + |x_{2^n-1}|^2 = 1$ 。为了计算方便,四维归一化的经典数据  $x = (0, 1, 0, 0)$  使用两个量子比特进行编码。振幅编码量子线路如图 6 所示。

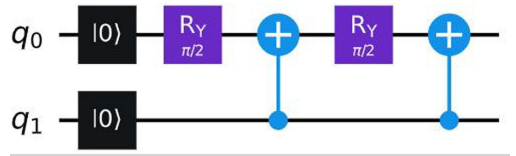


图 6 二量子位编码经典数据  $x = (0, 1, 0, 0)$

Fig. 6 Two qubits encode classical data  $x = (0, 1, 0, 0)$

图 6 量子线路的原理是利用量子门电路的幺正运算来编码。对该量子线路的幺正运算推导出输入数据  $x$ :

$$f(x; \theta) = \langle \phi(x) | U^\dagger(\theta) M U(\theta) | \phi(x) \rangle \quad (20)$$

其中,  $\phi(x)$  表示对输入数据  $x$  编码后的酉矩阵,  $U(\theta)$  为 Ansatz 层的矩阵表示,  $U^\dagger(\theta)$  为  $U(\theta)$  的共轭转置,  $M$  取  $\delta_z$  表示在  $z$  方向上进行测量的测量算子。测量后的量子位会坍塌,将测量后得到的经典信息保存在经典寄存器中。

图 7 设计的 4 种不同的 Ansatz 都选用  $\theta = [\theta_0, \theta_1, \dots, \theta_{15}]$  作为训练参数。 Ansatz 1 将  $\theta$  全部编码到  $R_y$  门中,并将每个量子位之间用 CNOT 门纠缠。 Ansatz 2 中将一般的训练参数  $\theta$  用  $R_z$  编码。 Ansatz 3 网络层与层之间没有使用 CNOT 将量子比特位之间纠缠,只在测量前将各量子位聚合到最后一个量子位进行测量。 Ansatz 4 在参数层前每个量子位添加一个 Hadamard 门,用于创建量子叠加态。

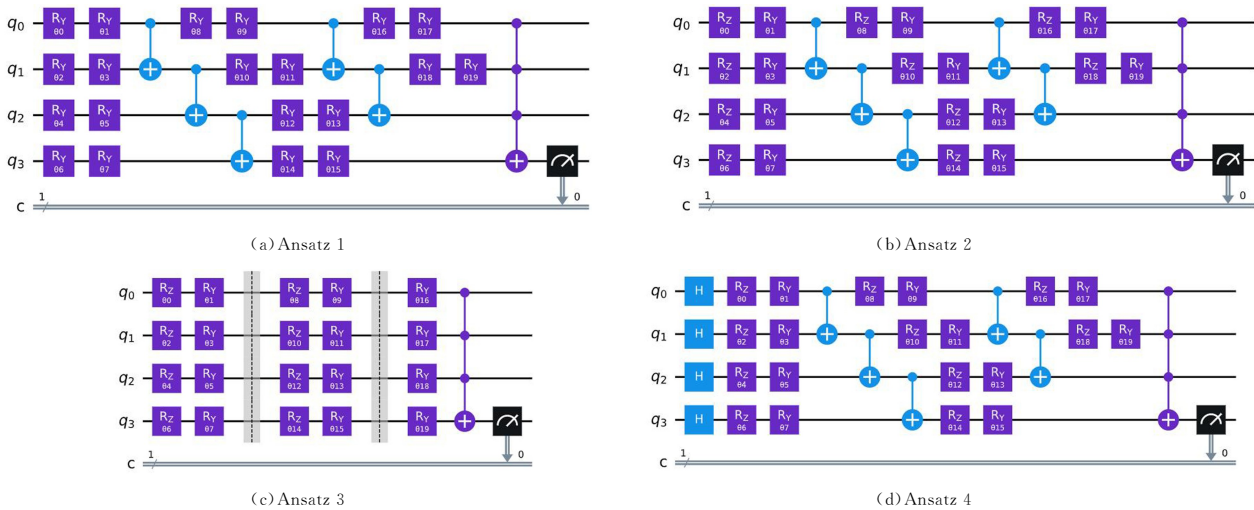


图 7 4 种不同的 Ansatz

Fig. 7 Four different Ansatz

### 3.3 优化算法

经典神经网络中通常使用梯度下降<sup>[18]</sup>等基于梯度的优化技术来更新网络权重参数,从而最小化损失函数。类似地,量子线路的参数优化也可以使用基于梯度的优化来达到参数更新的目的。图 8 是基于参数化量子线路的量子神经网络的

示意图,在训练参数层后对  $q_3$  量子位进行测量,该线路测量值的期望值可以表示为电路函数  $f(x; \theta)$ ; 选取交叉熵损失函数<sup>[19]</sup>作为测量期望值与实际标签值的判别依据;求损失函数  $L$  对  $\theta_i$  的梯度  $\frac{\partial L}{\partial \theta_i}$ ; 使用参数移位来计算每个参数的梯度。

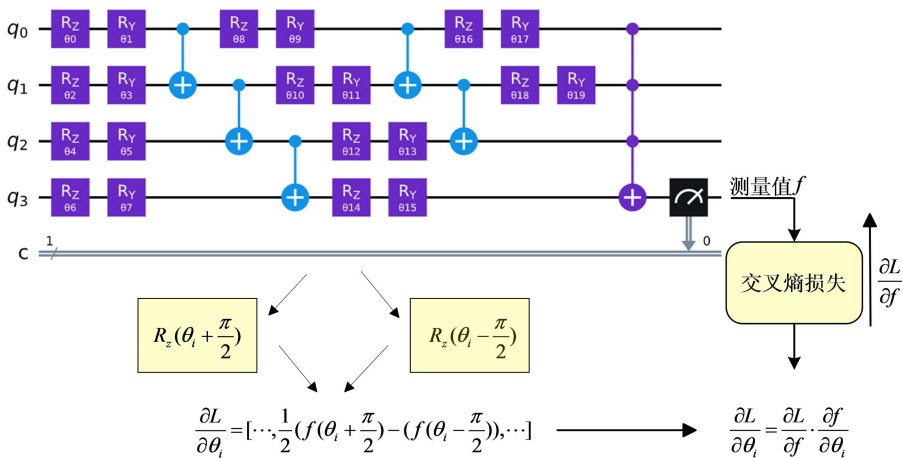


图 8 基于参数化量子线路的量子神经网络模型示意图

Fig. 8 Schematic of quantum neural network model based on parameterized quantum circuits

基于参数化量子线路的量子全连接神经网络解决现实图像的分类问题,其训练过程可以概括为以下步骤。

1) 输入经典数据集  $\mathcal{R} = \{(x^d, y^d)\}_{d=1}^D$ , 首先将每个特征向量  $x$  归一化为单位长度并使用酉矩阵  $U(x)$  编码为模型的输入量子态  $|\psi(x)\rangle$ , 其中  $|\psi(x)\rangle = U(x)|0\rangle^{\otimes N}$ 。

2) 执行一系列 ansatz 酉矩阵  $U(\theta)$ , 可以得到量子系统  $U(x)U(\theta)|0\rangle^{\otimes N}$ 。

3) 测量量子系统, 得到样本预测值。对线路的最终输出态进行哈密顿算子  $M$  测量, 使用 pauli-z 基测量量子系统。用期望值  $f(x; \theta)$  表示测量后的量子系统。输出值为:

$$f(x; \theta) = \langle \psi(x) | U^\dagger(\theta) M U(\theta) | \psi(x) \rangle \quad (21)$$

4) 计算标签预测值  $f(x_i; \theta)$  与真实标签  $y_i$  的损失。损失函数  $L$  如式(22)所示, 设置样本 batchsize 为 1, 即  $N=1$ 。得到损失函数  $L$ :

$$L = -y \log f - (1-y) \log(1-f) \quad (22)$$

5) 计算损失函数  $L$  对训练参数  $\theta_i$  的微分  $\frac{\partial L}{\partial \theta_i}$ 。  $\frac{\partial L}{\partial \theta_i}$  的表达式如下:

$$\frac{\partial L}{\partial \theta_i} = \frac{\partial L}{\partial f} \cdot \frac{\partial f}{\partial \theta_i} \quad (23)$$

使用混合经典量子算法计算  $\frac{\partial L}{\partial \theta_i}$ ,  $\frac{\partial L}{\partial f}$  在经典计算机上完成, 得到  $\frac{\partial L}{\partial f} = \frac{f-y}{f(1-y)}$ 。使用 parameter-shift rule 计算量子线路中  $\frac{\partial f}{\partial \theta}$  的梯度:

$$\frac{\partial f(x; \theta_0, \theta_1, \dots, \theta_k)}{\partial \theta_j} = \frac{f(x; \theta_0, \theta_1, \dots, \theta_j + \pi/2, \dots, \theta_k)}{2} - \frac{f(x; \theta_0, \theta_1, \dots, \theta_j - \pi/2, \dots, \theta_k)}{2} \quad (24)$$

其中,  $k$  为 Ansatz 层参数个数。得到损失函数  $L$  对  $\theta_i$  的

微分  $\frac{\partial L}{\partial \theta_j}$ :

$$\frac{\partial}{\partial \theta_j} L = \frac{f(x_i; \theta) - y_i}{f(x_i; \theta)(1 - f(x_i; \theta))} \frac{\partial}{\partial \theta_j} f(x_i; \theta) \quad (25)$$

6)重复执行 2)-5)直到  $\theta$  收敛,最后训练出来的线路可以给出正确的分类结果。

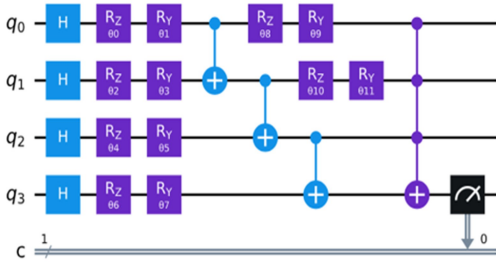
## 4 实验

### 4.1 实验环境

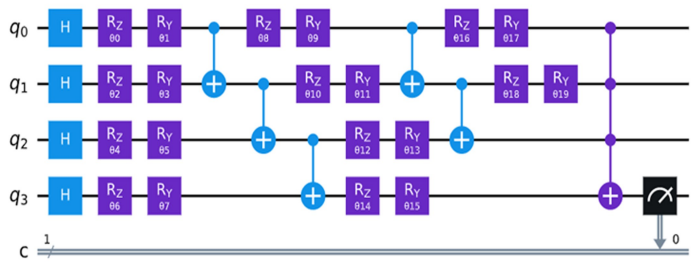
实验中参数化量子线路模型均基于 Qiskit 设计,模型的训练和测试都是在深度学习框架 PyTorch 下进行的。系统硬件环境为 Intel(R) Xeon(R) CPU E5-2697 v3 @2.60 GHz 处理器,NVIDIA GeForce RTX 1080Ti GPU。软件环境为 PyTorch 1.10.0 + cu102, qiskit 0.38.0,并在 qiskit-aer-gpu<sup>[20]</sup> 上进行模拟。

### 4.2 实验数据集和预处理

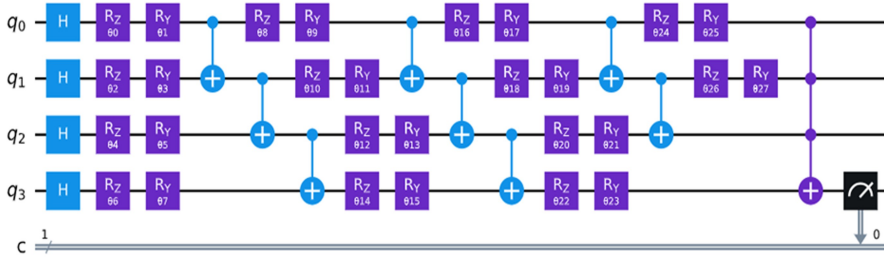
实验数据集为 MNIST 手写体数据集,MNIST 数据集总



(a)12-parameters circuit



(b)20-parameters circuit



(c)28-parameters circuit

图 9 不同训练参数的 Ansatz

Fig. 9 Ansatz of different training parameters

### 4.4 实验结果

在四量子位的量子神经网络中,基于角度编码的数据编码方式是一一对一编码,所使用的数据集为 PCA 选取的 4 个最重要的特征向量。基于时间演化编码和振幅编码所选的数据集均为 PCA 选取的 16 个最重要的特征向量。每次实验学习

共包含了 70000 张 0-9 的手写数字图像。每个图像都是灰度图像,大小为  $28 \times 28$  像素,像素值的范围是 0 到 255,表示了图像中每个像素的灰度强度。选用感兴趣的数字 3 和 6,其对应的特征和标签一共有 14017 个样本。由于目前可用的量子比特的数量极其有限,将  $28 \times 28$  的 784 个特征维度像素值大小为 0 到 255 的图像展开为一维向量,然后使用 PCA 将其降低到 16 个特征向量。实验选取的 3 和 6 数据集均为 600 个样本,总样本数量为 1200 个,按照 4 : 1 的比例划分训练集和测试集。

### 4.3 模型参数设置

实验使用角度编码、时间演化编码和振幅编码来探究不同编码方式对模型学习能力和性能的影响。在训练参数层设计了 4 个不同的 Ansatz,分别为 circuit 1—circuit 4。在训练层中采用不同的网络深度,分别为 12 个训练参数、20 个训练参数和 28 个训练参数,网络结构如图 9 所示。相对而言 28 个训练参数比 20 个训练参数多一层隐藏层的网络深度。

率  $\gamma$  均为 0.01,经典计算部分选用 SGD 算法,量子线路优化选用 Parameter-Shift Relus,每一次 batch 后对参数  $\theta_j$  分别移动  $+2/\pi$  和  $-2/\pi$ ,训练进行 100 次迭代。

MNIST 预处理后数字 3 和 6 在模型上的部分测试结果如图 10 所示。

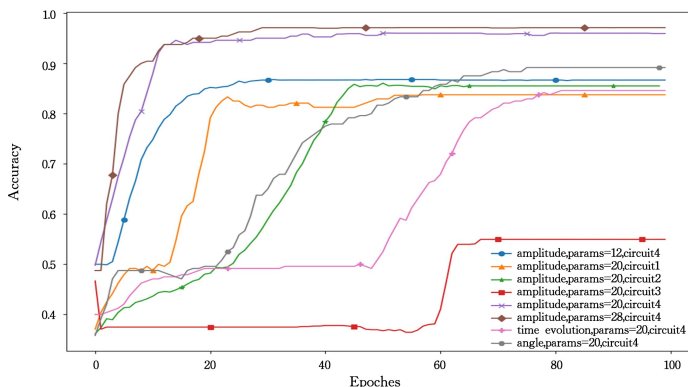


图 10 MNIST 预处理后数字 3 和 6 在模型上的部分测试结果

Fig. 10 Partial test results of digits 3 and 6 on the model after MNIST preprocessing

表 1 列出了 Ansatz 层 circuit 1—circuit 4 采用角度编码、时间演化编码和振幅编码的数据编码方式,以及不同深度 Ansatz 层的测试结果。

表 1 不同数据编码层和训练层在 MNIST 数据集 3 和 6 上的测试结果

Table 1 Test results of different data encoding layers and training layers on MNIST datasets 3 and 6					
模型 参数数量	编码层结构	训练层结构			
		circuit 1	circuit 2	circuit 3	circuit 4
20	角度编码	0.8150	0.830	0.525	0.892
20	时间演化编码	0.8050	0.835	0.565	0.846
20	振幅编码	0.8375	0.855	0.655	0.960
12	振幅编码	0.8170	0.825	0.384	0.867
28	振幅编码	0.8500	0.910	0.704	0.971

#### 4.4.1 不同数据编码方式下模型结果分析

从图 10 可以看出,模型参数数量一定,并且训练电路相同时,在四量子比特位下振幅编码的编码方式的性能最好。在训练参数数量一定时,线路 circuit 4 性能是最优的。具体而言,从表 1 可以看出,在相同训练参数和网络结构下,使用振幅编码的数据编码方式比使用时间演化编码方式的泛化性能更强,其中在 20 个训练参数下,基于振幅编码的 circuit 4 精度达到了 0.96,高于基于时间演化编码的 0.846 分类精度。分析原因是,振幅编码将输入的经典数据转化为量子比特的幅度上,每个经典数据样本被映射为一个量子态,其中量子比特的幅度对应于该数据样本的相对频率。而时间演化编码是通过量子门的么正运算,利用量子系统在演化过程中的相位变化和幅度变化来实现特定的编码目标。例如编码 256 个特征的经典数据,使用振幅编码仅需 8 量子比特,而时间演化编码每个量子比特需要堆叠 32 个量子比特门编码经典数据,空间复杂度较大。角度编码是对经典数据一对一编码,当前量子比特资源十分有限,这种方式需要大量量子比特,而量子线路的优势是基于量子力学中的叠加性,所以对于经典数据编码到量子态,选用振幅编码是不错的选择。

#### 4.4.2 不同 Ansatz 下模型性能分析

本实验设计了 4 种不同的 Ansatz,探究分类任务中最优的参数化量子线路。

1) circuit 1 相比于 circuit 2,将训练参数从全部的  $R_Y$  门替换为使用  $R_Z R_Y$  相结合的方式,在整个 Bloch 球体中,将数据耗散到希尔伯特的整个空间对分类有很大的帮助;

2) 对于 circuit 2 和 circuit 3,如果量子位之间不使用 CNOT 门,那么量子位之间就无法成为一个复合系统,每个量子位会是一个孤立的量子系统而无法建立联系,从表 1 中的数据可以看到对于二分类问题,该模型分类精度仅为 0.525~0.655,并没有分类能力;

3) 比较量子电路 circuit 2 和 circuit 4,探究在电路中添加 Hadamard 门是否影响线路时,选取较为简单的量子线路作为演示,如图 11 所示。

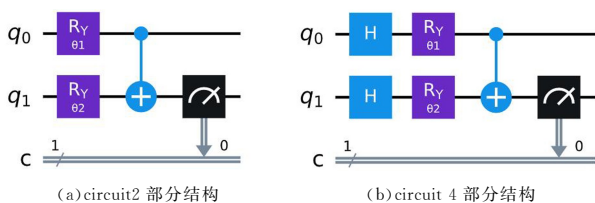


图 11 Hadamard 门对线路的影响

Fig. 11 Effect of Hadamard gate on circuit

对于图 11 中的两个量子线路,  $q_0$  和  $q_1$  两个初始态为  $|0\rangle$  的量子线路,测量前的量子态可以表示为式(26)、式(27):

$$(CONT) \cdot (R_Y(\theta_1) \otimes R_Y(\theta_2)) \cdot (|0\rangle \otimes |0\rangle) = \begin{bmatrix} \cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \\ \cos \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \\ \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \\ \sin \frac{\theta_1}{2} \cos \frac{\theta_2}{2} \end{bmatrix} \quad (26)$$

$$(CONT) \cdot (R_Y(\theta_1) \otimes R_Y(\theta_2)) \cdot (H \otimes H) (|0\rangle \otimes |0\rangle) = \frac{1}{2} \begin{bmatrix} \left( \cos \frac{\theta_1}{2} + \sin \frac{\theta_1}{2} \right) \left( \cos \frac{\theta_2}{2} + \sin \frac{\theta_2}{2} \right) \\ \left( \cos \frac{\theta_1}{2} + \sin \frac{\theta_1}{2} \right) \left( \cos \frac{\theta_2}{2} - \sin \frac{\theta_2}{2} \right) \\ \left( \cos \frac{\theta_1}{2} - \sin \frac{\theta_1}{2} \right) \left( \cos \frac{\theta_2}{2} - \sin \frac{\theta_2}{2} \right) \\ \left( \cos \frac{\theta_1}{2} - \sin \frac{\theta_1}{2} \right) \left( \cos \frac{\theta_2}{2} + \sin \frac{\theta_2}{2} \right) \end{bmatrix} \quad (27)$$

在每个量子位添加 Hadamard 门使得量子电路更加复杂,表达能力更强,便于找到最优的参数。

#### 4.4.3 线路深度对模型性能的影响

神经网络的隐藏层数过少,网络的表达能力会受到限制,无法有效地学习到数据的特征。因此,在参数化量子线路的训练层中将训练参数  $\theta_i$  分别设置为 12, 20 和 28 个,相当于设置了 3 种不同的隐藏层层数,具体的训练层结构如图 9 所示。从表 1 中可以看出,在同一编码模式下,训练层的层数越多,对于分类任务的准确率越高。使用振幅编码对输入数据进行编码,同时采用 circuit 4 的模型结构,当训练参数为 28 个时,模型在 MNIST 数字 3 和 6 的分类任务上达到了 97% 以上的准确率。

**结束语** 为探究量子计算与机器学习结合的可能性,本文针对经典神经网络进行量子电路层嵌入,构成量子神经网络,并提出了 3 种不同的数据编码方式。针对训练层网络结构,提出了 4 种强纠缠的参数化量子线路的量子神经网络模型,并且详细阐述了网络模型的内部结构和理论基础。对每个模型在超参数一定的情况下进行实验比较。实验结果表明,基于振幅编码的参数化量子线路模型在  $\{3, 6\}$  分类任务中的分类准确率超过了 97%,高于文献[11]模型 88% 的分类精度。在  $\{3, 8\}$  分类任务中,本文模型分类准确率达到 91.6%,证明了基于参数化量子线路构建量子神经网络的可行性,未来可以将模型扩展到多分类任务。对于泛化能力更强的卷积神经网络,是否可以将参数化量子线路用于编码卷积核构成量子卷积神经网络<sup>[21]</sup>,仍需进一步研究。

## 参考文献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [2] XU X, DING Y, HU S X, et al. Scaling for edge inference of deep neural networks[J]. Nature Electronics, 2018, 1(4): 216-222.
- [3] BRAVYI S, DIAL O, GAMBETTA J M, et al. The future of quantum computing with superconducting qubits[J]. Journal of Applied Physics, 2022, 132(16).

- [4] BEER K, BONDARENKO D, FARRELLY T, et al. Training deep quantum neural networks [J]. *Nature Communications*, 2020, 11(1): 808.
- [5] JIE G J, ZHUANG Z Q. Quantum Neural Networks [J]. *Computer Science*, 2001(7): 1-6.
- [6] ZHANG L F, ZHANG X P. Network Traffic Prediction based on BP Neural network optimized by Quantum Genetic Algorithm [J]. *Computer Engineering and Science*, 2016, 38(1): 114-119.
- [7] HUANG H Y, BROUGHTON M, COTLER J, et al. Quantum advantage in learning from experiments [J]. *Science*, 2022, 376(6598): 1182-1186.
- [8] SIM S, JOHNSON P D, ASPURU-GUZI K A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms [J]. *Advanced Quantum Technologies*, 2019, 2(12): 1900070.
- [9] ZHENG J, GAO Q, LV Y X, et al. Model and application of quantum Convolutional Neural Network Based on Parameterized quantum Circuit [J]. *Control Theory & Applications*, 2021, 38(11): 1772-1784.
- [10] PRESKILL J. Quantum computing in the NISQ era and beyond [J]. *Quantum*, 2018, 2: 79.
- [11] WANG H, LI Z, GU J, et al. QOC: quantum on-chip training with parameter shift and gradient pruning [C] // *Proceedings of the 59th ACM/IEEE Design Automation Conference*. 2022: 655-660.
- [12] CHEN K, ALBEVERIO S, FEI S M. Entanglement of formation of bipartite quantum states [J]. *Physical Review Letters*, 2005, 95(21): 210501.
- [13] HUANG C H. Analysis and discussion on superposition principle of quantum states and its measurement [J]. *University Physics*, 2013, 32(4): 22-24, 27.
- [14] CROOKS G E. Gradients of parameterized quantum gates using the parameter-shift rule and gate decomposition [J]. *arXiv: 1905.13311*, 2019.
- [15] QIAN C. Quantum Entanglement and Quantum Computation [J]. *Computer Science*, 2006(12): 230-234.
- [16] CORTESE J, BRAJE T. System and technique for loading classical data into a quantum computer; U. S. Patent 11, 113, 621 [P]. 2021-09-07.
- [17] GANDHI V, PRASAD G, COYLE D, et al. Quantum neural network-based EEG filtering for a brain-computer interface [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, 25(2): 278-288.
- [18] CHAPPELLE O, WU M. Gradient descent optimization of smoothed information retrieval metrics [J]. *Information Retrieval*, 2010, 13: 216-235.
- [19] KOBAYASHI T. Large Margin In Softmax Cross-Entropy Loss [C] // *BMVC*. 2019: 3.
- [20] SHAIK E H, RANGASWAMY N. Implementation of quantum gates based logic circuits using IBM Qiskit [C] // *2020 5th International Conference on Computing, Communication and Security (ICCCS)*. IEEE, 2020: 1-6.
- [21] CONG I, CHOI S, LUKIN M D. Quantum convolutional neural networks [J]. *Nature Physics*, 2019, 15(12): 1273-1278.



**CHEN Chao**, born in 1995, postgraduate. His main research interests include quantum machine learning, machine learning theory, and its applications.



**XUE Guixiang**, born in 1979, Ph.D., associate professor. Her main research interests include machine learning, deep learning, and its application in smart city and intelligent transport systems.