



# 计算机科学

COMPUTER SCIENCE

## 基于混合量子经典长-短距离特征扩展网络的图像分类

郑毅, 贾星昊, 张骏温, 任爽

引用本文

郑毅, 贾星昊, 张骏温, 任爽. 基于混合量子经典长-短距离特征扩展网络的图像分类[J]. 计算机科学, 2026, 53(4): 277-283.

ZHENG Yi, JIA Xinghao, ZHANG Junwen, REN Shuang. [Image Classification Based on Hybrid Quantum-Classical Long-Short Range Feature Extension Network](#) [J]. Computer Science, 2026, 53(4): 277-283.

---

## 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)**

[自然语言语义表示的范畴论建模: 系统综述与组合机制分析](#)

Category-Theoretic Semantic Representation: Systematic Review and Compositional Mechanism Analysis

计算机科学, 2026, 53(4): 337-346. <https://doi.org/10.11896/jsjcx.251000136>

[NISQ量子线路高频-密集量子门集策略优化算法](#)

High Frequency-Dense Quantum Gate Set Optimization Algorithm for Quantum Circuit in NISQ Era

计算机科学, 2026, 53(4): 112-120. <https://doi.org/10.11896/jsjcx.241200213>

[基于大语言模型和深度网络的认知评估量表自动诊断](#)

Large Language Model and Deep Network Based Cognitive Assessment Automatic Diagnosis

计算机科学, 2026, 53(3): 41-51. <https://doi.org/10.11896/jsjcx.250600034>

[基于变量子量的离散对数求解算法](#)

Variational Quantum Algorithm for Solving Discrete Logarithms

计算机科学, 2026, 53(1): 353-362. <https://doi.org/10.11896/jsjcx.241100181>

[基于神经元覆盖指标的测试用例生成优化研究](#)

Research on Optimization of Test Case Generation Based on Neuron Coverage Index

计算机科学, 2025, 52(11): 339-348. <https://doi.org/10.11896/jsjcx.240900006>

# 基于混合量子经典长-短距离特征扩展网络的图像分类

郑毅 贾星昊 张骏温 任爽

北京交通大学计算机科学与技术学院 北京 100044

(zhengy@bjtu.edu.cn)

**摘要** 经典神经网络的规模、计算时长很难进一步突破,难以兼顾轻量化和高性能,在目前大数据时代下成为了解决海量数据的图像分类问题的瓶颈。而混合量子经典神经网络具有量子计算与经典计算的优势,能够进行高效的并行计算并具有较好的普适性。为此,设计了混合量子经典长-短距离特征扩展网络(Hybrid Quantum-Classical Long-Short Range Feature Extension Neural Network, HQC-LSNet),它是一种包含多个混合模块的多分支结构。通过多种量子旋转门及受控-Z 门构成量子解耦全连接注意力机制,利用量子特性从量子增强特征空间中高效地获取长距离特征;与此同时,采用经典卷积模块获取短距离特征,并以组合特征图的方式进行特征扩展。在 MNIST 的十分类以及 CIFAR-10 数据集上的三分类这两个图像多分类任务上测试其准确率分别为 99.42% 和 91.42%,相较于对应的经典模型及混合量子经典模型均有提升,而且该模型的参数量与时间复杂度相较于经典模型均有所减小。

**关键词**: 量子机器学习; 图像分类; 量子神经网络; 量子计算; 混合量子经典神经网络

中图分类号 TP183

## Image Classification Based on Hybrid Quantum-Classical Long-Short Range Feature Extension Network

ZHENG Yi, JIA Xinghao, ZHANG Junwen and REN Shuang

School of Computer Science & Technology Engineering, Beijing Jiaotong University, Beijing 100044, China

**Abstract** It is difficult to further break through the scale and computing time of classical neural networks, making it difficult to simultaneously achieve lightweight design and high performance. This has become a bottleneck in solving large-scale image classification problems in the era of big data. In contrast, hybrid quantum-classical neural networks combine the advantages of quantum and classical computing, enabling efficient, parallel computation and strong generalizability. To address this, this paper proposes the hybrid quantum-classical long-short range feature extension neural network (HQC-LSNet), a multi-branch architecture composed of multiple hybrid modules. It incorporates a quantum decoupled fully connected attention mechanism built with various quantum rotation gates and controlled-Z gates to efficiently extract long-range features from a quantum-enhanced feature space. Simultaneously, a classical convolutional module is employed to capture short-range features, and feature extension is performed by combining the resulting feature maps. The model achieves classification accuracies of 99.42% on the ten-class MNIST dataset and 91.42% on a three-class CIFAR-10 dataset, outperforming corresponding classical and hybrid quantum-classical models. Additionally, HQC-LSNet reduces both the parameter count and time complexity compared to purely classical models.

**Keywords** Quantum machine learning, Image classification, Quantum neural network, Quantum computing, Hybrid quantum-classical neural network

## 1 引言

近年来,随着神经网络的发展,使用神经网络进行图像分类的准确率超过传统算法,甚至可比肩人类水平;但由于图像数据量的不断增大,采用传统的基于经典神经网络的方法进行图像分类时,在处理大规模数据及缩短处理时间方面具有一定的局限性<sup>[1-2]</sup>。

针对以上问题,学者提出了许多改善方式。其中一种主要方式为轻量化网络模型,有代表性的包括 obileNet 系列<sup>[3-4]</sup>、GhostNet 系列<sup>[5-6]</sup>、IremulbNet<sup>[7]</sup>以及 FMGNet<sup>[8]</sup>等。上述轻量化网络模型大多运用了较低的参数量、较低的复杂度的网络结构及计算操作等来实现模型加速的目的。同时也有研究对经典网络使用各种模型压缩的方法,如剪枝、量化、知识蒸馏等来降低参数的数量与模型的计算量,提高模型的

到稿日期:2025-06-17 返修日期:2025-09-13

基金项目:国家自然科学基金(62072025)

This work was supported by the National Natural Science Foundation of China(62072025).

通信作者:任爽(sren@bjtu.edu.cn)

运行效率。但是经典的框架具有较难突破的限制,所以优化后的模型依然会有十万到百万数量级以上的参数,并且在优化模型的规模和复杂度的同时,可能降低了准确率。针对这样的情况,寻找并实现更优的海量数据处理方法仍是重要的任务。

与此同时,诞生了混合量子经典神经网络(Hybrid Quantum-Classical Neural Network, HQCNN),其具有量子 and 经典神经网络的优点:一方面,可以利用量子叠加及纠缠特性实现少量参数下的高速度运行,甚至是某些问题上的指数加速<sup>[9-10]</sup>;另一方面,可利用经典模块补充非线性操作并弥补复杂量子模块对量子硬件要求较高的缺陷。同时,运用量子模块能够使得模型拥有更大的特征空间,更有利于找到全局最优解,能够在保证模型简化的同时得到更高的准确率<sup>[11-12]</sup>。HQCNN 应用于有海量数据需要处理的图像分类任务的优势显著,在近几年的研究中越发热门。

但是,目前的 HQCNN 用于图像分类时存在分类准确度低,算法与数据尺寸和维度耦合的问题。本文参考了 GhostNet v2 中介绍的方法<sup>[6]</sup>,提出了混合量子经典长-短距离特征扩展网络,以及量子解耦全连接注意力机制(Quantum Decoupled Fully Connected Attention, QDFC Attention),目的在于通过量子计算提高图像分类的精度以及运算速度。本文的主要贡献如下:

1)提出了 HQC-LSNet 神经网络模型,基于经典卷积层、池化层以及利用混合量子经典特征扩展模块形成的多分支结构,使用的混合模块采用量子解耦全连接注意力机制和经典卷积,从长距离和短距离对信息进行抽取与特征扩增。

2)提出了量子解耦全连接注意力机制模块,采用参数化量子电路,通过振幅编码及多种量子门相结合的方式构造,在基于滑动窗口计算方法的基础上分别从左到右、从上至下两个方向得到该图像区域的全局注意力特征图,并且此模块中的特征图大小与模块本身的设计解耦,能够方便调整与扩展。

3)将 HQC-LSNet 应用于图像分类任务,在 MNIST 十分类、CIFAR-10 三分类任务中实验准确率分别达到了 99.42% 和 91.42%,相比同规模经典神经网络模型,在准确率上更高,在参数量和计算量上更少;相较于之前已提出的其他混合量子经典神经网络模型,其在参数量、计算量和准确率之间有较好的平衡。

## 2 量子计算基础





量子比特是量子计算和量子存储的基本单位。量子比特有叠加属性, $n$  个量子比特就足以生成一个  $2^n$  维的状态空间,这使得量子算法即使只有少量的量子比特也能表示巨大的数据量;量子比特间还具有叠加属性,量子比特发生变换就会相互牵制、联动,所以信息可以以低耗费被传输,全局处理也能以廉价的方式实现。

量子门作为量子计算机中操控量子比特状态的基本单元,可以在高维希尔伯特空间上实现搜索和计算,以使量子算法获得较大的解空间,为量子算法取得更高的性能奠定基础<sup>[13]</sup>。量子门可以看成是一个满足  $U^\dagger U = U U^\dagger = \mathbf{I}$  的酉矩阵

$U, U^\dagger$  为  $U$  的共轭转置, $\mathbf{I}$  为单位矩阵。本文中涉及的量子门及其对应的表示如表 1 所列。

表 1 基础量子门及其表示

Table 1 Basic quantum gates and their representations

量子门	符号表示	矩阵表示	图形表示
RX 门 (旋转 X 门)	$R_X$	$\begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right) \\ -i\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix}$	
RY 门 (旋转 Y 门)	$R_Y$	$\begin{bmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{bmatrix}$	
RZ 门 (旋转 Z 门)	$R_Z$	$\begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$	
受控-Z 门	$C_Z$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	

参数化量子电路(Parameterized Quantum Circuit, PQC)是目前常用的一种量子神经网络的实现方式。具体来说,PQC 的核心思路是利用有参和无参两种类型的量子门组合成多层的量子拟设,然后用经典优化算法来进行学习训练,此思路在现在的量子比特数相当有限又存在量子噪声干扰的含噪声的中尺度量子计算(Noisy Intermediate-Scale Quantum, NISQ)时代非常适用。PQC 构建的量子神经网络一般包含输入编码、参数化操作和测量模块 3 个基本部分。

## 3 相关工作

混合量子经典神经网络是将经典模块与量子模块组合形成的混合神经网络,它同时具有量子神经网络的高速度与经典神经网络的普适性,能够展现出更强的表达能力与计算性能。

Henderson 等<sup>[14]</sup>使用随机量子电路实现了一种量子卷积层,并将其加入到经典神经网络框架中使用,取得了相较于经典卷积神经网络更高的准确率。Gupta 等<sup>[15]</sup>提出了量子非局部神经网络,结合经典卷积神经网络以及量子非局部模块并行地计算大量输入特征的远距离依赖关系,在图像二分类任务上取得高准确率的同时将参数量控制在十万以内。Senokosov 等<sup>[16]</sup>提出了一个具有并行量子电路的混合量子神经网络(HQNN-Parallel),使用多个可以并行执行的量子模块进行计算,从而实现减少总计算时间的效果。Rao 等<sup>[17]</sup>提出一种自定义卷积神经网络,通过引入量子分类器,不仅实现了超过一般经典模型分类准确率,而且显著降低了内存消耗和模型体积。Xiong 等<sup>[18]</sup>提出了一种轻量化的混合量子卷积神经网络,并将其应用在恶意代码检测任务中,取得了较优的检测性能,但相比目前先进的深度学习模型还有一定差距。另外,文献[19-21]也研究了混合量子经典神经网络,并证明了混合量子经典神经网络能够达到与结构相似的经典网络相近或更高的性能,文献[21-22]的研究也证明了量子模块能够降低参数量。

## 4 混合量子经典长-短距离特征扩展网络

本文参考 GhostNet v2<sup>[6]</sup>的工作,提出了 HQC-LSNet,如图 1 所示,它是一种多分支并行的网络结构。输入图像首先

进入每一分支的一组经典卷积层与最大池化完成降维和局部特征提取;之后利用 QDFC 注意力机制与经典卷积单元构成的混合量子经典特征扩展模块,在长、短距离上分别抽取特征;分支的最后利用一层经典卷积将不同尺度的特征图统一为固定的大小与维数。对经过多个分支处理得到的结果进行合并,再利用经典卷积层做进一步处理,最终在全连接层进行

分类。HQC-LSNet 通过并行分支结构可以同时得到、处理多组特征图,实现了特征图的扩展。其中的量子模块虽然可以提供更高的计算效率和更小的参数量,但无法实现激活函数等非线性操作,且当输入数据为高维度大尺寸时需要大量的量子资源,所以此处利用的经典结构可以对上述问题进行修正。

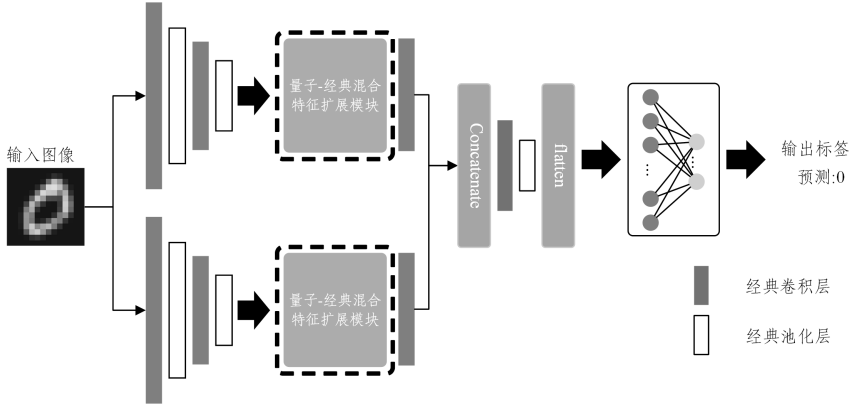


图 1 HQC-LSNet 模型的整体结构

Fig. 1 Overall structure of HQC-LSNet model

#### 4.1 QDFC 注意力机制

注意力机制可以通过引入长距离的空间依赖性来提高模型对时空信号理解的效果,但由于注意力机制的计算复杂度随特征图大小呈二次增长,很难以低成本处理海量的图像数据。虽然经典方法已经针对注意力机制进行了大量尝试性的加速,如采用窗口划分策略的方法<sup>[23-24]</sup>,但由于本身复杂度上就存在下限,因此只能减缓,无法从根本上突破注意力机制的复杂度瓶颈。而现阶段对量子注意力网络的研究<sup>[25-26]</sup>大多还是沿用了经典的注意力网络结构,因此不容易扩展与应用。但是解耦全连接注意力机制(Decoupled Fully Connected Attention,DFC Attention)通过把注意力过程拆开,并利用全连接层做计算,显著简化了算法结构,降低了计算复杂度。在此基础上,本文进一步提出其在量子计算框架下的实现方案,以实现低成本的长距离特征提取。

给定一个特征图  $\mathbf{Z} \in \mathbb{R}^{H \times W \times C}$ , 可将其视为  $HW$  个 token  $\mathbf{z}_i \in \mathbb{R}^C$  的组合,即  $\mathbf{Z} = \{\mathbf{z}_{11}, \mathbf{z}_{12}, \dots, \mathbf{z}_{HW}\}$ 。则对特征图  $\mathbf{Z}$  的

全连接注意力  $\mathbf{a}_{hw}$  可表示为式(1)的形式:

$$\mathbf{a}_{hw} = \sum_{h',w'} \mathbf{P}_{hw,h'w'} \odot \mathbf{z}_{h'w'} \quad (1)$$

由于特征图可视为低秩矩阵<sup>[27]</sup>,可将全连接注意力按照式(2)一式(4)的形式分解为水平和垂直两个方向上的特征矩阵,从而进行解耦的全连接注意力计算。

$$\mathbf{a}'_{hw} = \sum_{h'=1}^H \mathbf{P}_{h,h'w}^H \odot \mathbf{z}_{h'w'}, h=1,2,\dots,H, w=1,2,\dots,W \quad (2)$$

$$\mathbf{a}_{hw} = \sum_{w'=1}^W \mathbf{P}_{w,hw'}^W \odot \mathbf{z}_{hw'}, h=1,2,\dots,H, w=1,2,\dots,W \quad (3)$$

$$\mathbf{a}_{hw} = \mathbf{a}'_{hw} \times \mathbf{a}_{hw} \quad (4)$$

其中,  $\mathbf{P}^H$  和  $\mathbf{P}^W$  为全连接层中的可学习权重。

对 DFC 注意力的特征图进行划分,并对划分出的每个小窗口内的所有特征点使用参数共享且基于 PQC 构建的全局特征提取器(Global Feature Extractor, GFE)来进行特征提取,能够进一步减少模型的参数量、计算量,摆脱输入特征图大小对算法的影响,使其具有良好的输入适应性和结构可扩充性。GFE 的整体量子电路图如图 2(a)所示。

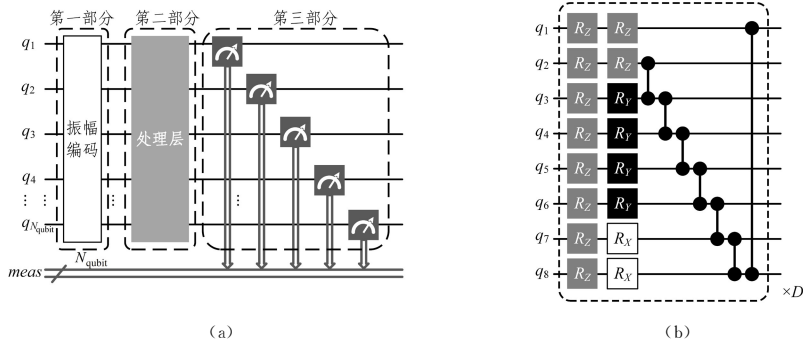


图 2 全局特征提取器的量子电路图

Fig. 2 Quantum circuit of the global feature extractor

GFE 包括 3 个部分,分别对应处理流程中的 3 个步骤。第一是采用振幅编码的方式将大小为  $N \times 1 \times C$  的滑动窗口中选取的经典特征向量  $\mathbf{X} = \{x_{11}, x_{12}, \dots, x_{NC}\} \in \mathbb{R}^{N \times 1 \times C}$  映射

到  $N_{\text{qubit}} = \log_2(NC)$  个量子比特上,得到一个初始量子态。使用振幅编码方式将原始经典数据转换成多维的叠加态,实现了输入数据的降维,减少了对量子比特的需求和对电路的要

求,使基于模型的模型参数大幅减少。

如图 2(b)所示,在第二阶段,该阶段堆叠的一个  $D$  层量子电路是由首先施加在全部量子比特上的  $RZ$  门,之后对不同的量子比特各自应用的  $RX, RY$  和  $RZ$  旋转门所构成的,这样保证了该量子电路拥有灵活的酉变换拟合能力。而这个灵活性可以通过旋转参数  $\theta$  在学习过程中不断调节,并且组合多个旋转门操作扩大特征空间有助于增强模型性能。在每一层的最后,还利用受控- $Z$  门来实现不同量子比特间的纠缠,通过这种跨越量子比特的连接方式构成一种循环的纠缠环路,以增强量子态之间的信息交流,从而实现全局相关性建模,充分捕获窗口内的远距离关联。式(5)~式(7)是上述整个过程的数学描述,其中  $\theta$  为旋转门操作中的相应参数。

$$|\phi_1\rangle_0 = (R_Z(\theta_1)_{q_1} \cdots R_Z(\theta_D)_{q_{N_{\text{qubit}}}}) |\psi_0\rangle \quad (5)$$

$$|\phi_1\rangle_1 = (R_Z(\theta_{D+1})_{q_1} \cdots R_Y(\theta_{D+d})_{q_d} \cdots R_X(\theta_{2D})_{q_{N_{\text{qubit}}}}) |\phi_1\rangle_0 \quad (6)$$

$$|\phi_1\rangle_2 = (C_Z(q_2 q_3) \cdots C_Z(q_{N_{\text{qubit}}-1} q_{N_{\text{qubit}}})) C_Z(q_{N_{\text{qubit}}} q_1) |\phi_1\rangle_1 \quad (7)$$

堆叠  $D$  层后的整个过程可以用酉算子表示为式(8)、式(9)的形式:

$$U(\theta) = |\phi_1\rangle_2 \quad (8)$$

$$|\phi_1\rangle = (U(\theta))^D |\phi_0\rangle \quad (9)$$

$|\phi_1\rangle$ 即为通过第二部分得到的最终量子态。

最后,GFE的第三部分负责对被第一、第二部分影响后的量子态进行测量,输出经典数据。

至此,得到了完整的 GFE,它通过量子全局操作实现滑动窗口内的全局特征捕获,可实现与经典卷积以及全连接层相似的效果,且参数量只有  $2DN_{\text{qubit}}$ 。由于 QDFC 注意力机制需要在水平和垂直两个方向上使用滑动窗口对整个特征图进行处理,因此同时应用了水平和垂直两个 GFE,最终整个 QDFC 注意力机制的参数量为  $4DN_{\text{qubit}}$ 。由于引入振幅编码缩小了数据规模,其时间复杂度为  $O(\frac{K_H}{H} \log_2(H^2 WC^2) +$

$\frac{K_W}{W} \log_2(HW^2 C^2))$ ,其中  $K_H$  和  $K_W$  为滑动窗口的大小。相比于经典自注意力机制的  $O(H^2 W^2 C + HWC^2)$ <sup>[28]</sup>以及经典 DFC 注意力机制的  $O(K_H HWC^2 + K_W HWC^2)$ <sup>[6]</sup>,QDFC 实现了指数级的加速。

#### 4.2 混合量子经典特征扩展模块

图 3 是混合量子经典特征扩展模块的结构示意图,它有两个平行的子分支,分别建模图像长距离的全局特征以及短距离的局部特征。QDFC 注意力机制分支处输入的特征图经过均值池化及经典卷积降维后,在输出分支之前再经由上采样恢复;经典卷积分支利用两个经典的卷积层从前往后提取图像的不同层次局部特征,将两次卷积层得到的两张不同特征图合并在一起,实现了多尺度信息的相互补充、融合;最后两个分支所得的结果相乘即可采用注意力机制突出特征,着重加强高重要性的特征。

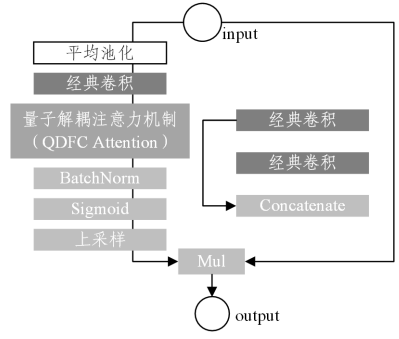


图 3 混合量子经典特征扩展模块结构

Fig. 3 Structure of the hybrid quantum-classical feature extension module

## 5 实验与结果分析

### 5.1 实验环境与评价指标

本研究使用了在图像处理领域广泛使用的 MNIST 数据集以及 CIFAR-10 数据集进行图像分类实验。对于 MNIST 数据集,所有类别进行十分类实验,且每个类别包含 400 个样本。对于 CIFAR-10,则进行了类别为 2,4,8 的三分类实验,每个类别包含的样本数均为 10000。两个数据集均按照 4:1 的比例划分为训练集和测试集。在将图像输入模型之前,使用每个数据集的全局平均值和标准偏差对其进行了归一化处理。

本研究在经典计算机上对量子模型进行模拟实验,虽然这会造成其在处理时间上的损失。实验中使用的具体硬件设备、软件版本以及各种训练参数如表 2 所列。

表 2 实验环境配置

Table 2 Experimental environment configuration

硬件/软件/参数	型号/版本/参数取值
操作系统	Ubuntu 18.04.6
CPU	Intel <sup>(R)</sup> Xeon <sup>(R)</sup> Silver 4214 2.20 GHz
Python	3.11
Cuda	12.1
PyTorch	2.2.2
Qiskit	1.2.4
优化器	Adam
学习率	0.01~0.005
量子比特数 $N_{\text{qubit}}$	8
QDFC 注意力中的堆叠层数 $D$	2

本研究选择分类任务常用的准确率(Accuracy)、精确率(Precision)作为性能评价指标,且使用宏平均来进行整体评估,如式(10)~式(12)所示:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$precision = \frac{TP}{TP + FP} \quad (11)$$

$$Metric_{\text{macro}} = \frac{1}{k} \sum_{i=1}^k metric_i \quad (12)$$

$$metric_i = accuracy_i \text{ 或 } precision_i$$

另外,本研究还使用参数量与每秒浮点运算次数(FLOPs)来评估算法规模与计算速度。

### 5.2 模型具体参数

HQC-LSNet 模型中的经典结构和具体参数与不同的输入输出要求相关。在 MNIST 数据集以及 CIFAR-10 数据集的设置分别如图 4(a)、图 4(b)所示。

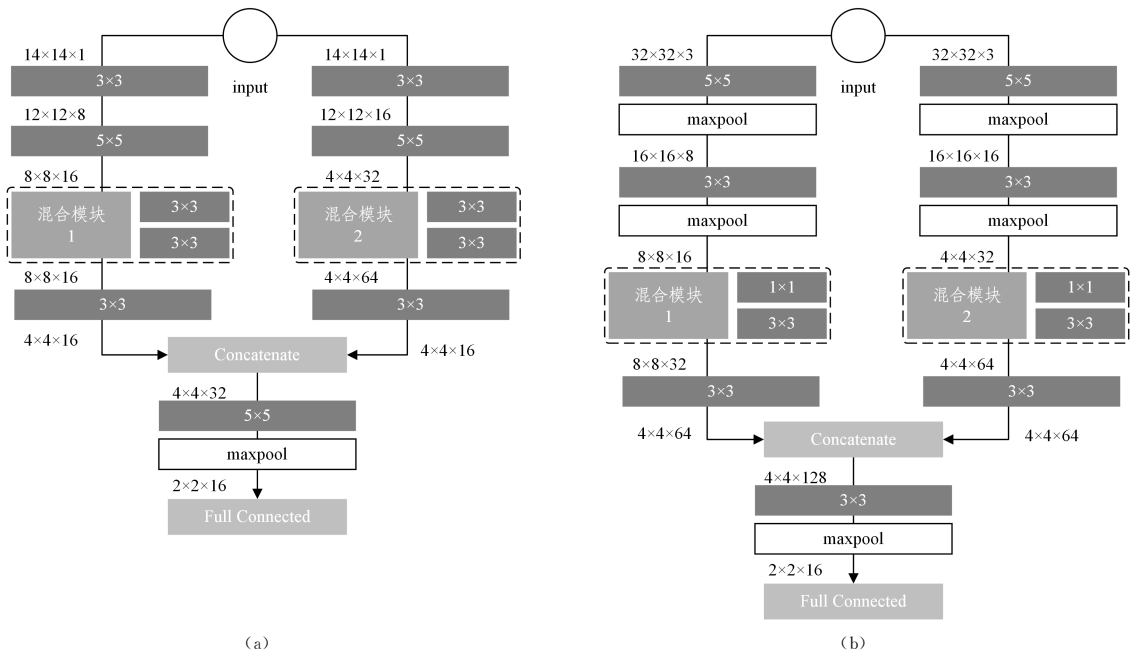


图4 HQC-LSNet的具体结构以及输入每个模块的特征图尺寸

Fig. 4 Specific structure of HQC-LSNet and size of the feature map input for each module

### 5.3 实验结果

为了评估本文模型的性能,分别和经典神经网络 CNN, GhostNet v2<sup>[6]</sup>, MobileNet v3<sup>[4]</sup> 以及混合量子经典神经网络 QCNN-Hen<sup>[14]</sup>, QCNN-Liu<sup>[29]</sup>, MERA<sup>[30]</sup>, HQNN-Parallel<sup>[16]</sup>, HQNN-Quanv<sup>[16]</sup> 模型进行了对比实验。在经典的网络实验中,把经典的网络参数控制为与本文模型相近的数量。在实验中,我们通过调整经典网络的整体结构,使其参数量与本文的混合模型保持接近,以确保对比公平。表3、表4为本次对比实验的结果,图5(a)为本文模型与其他参考模型在 MNIST 数据集上对于十分类任务训练过程中损失的变化情况,图5(b)为在 CIFAR-10 上关于三分类任务训练损失变化情况。

表3 MNIST上的对比实验结果

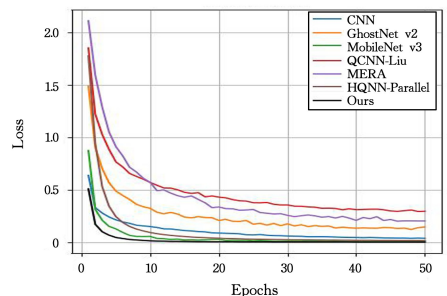
Table 3 Comparative experimental results on the MNIST

模型	参数量	FLOPs	准确率	精确率
CNN	44634	16.31	0.9860	0.9297
GhostNet v2	43278	13.09	0.9910	0.9560
MobileNet v3	43298	15.92	0.9890	0.9441
QCNN-Hen	1978	0.04	0.9305	0.6479
QCNN-Liu	44034	11.94	0.9737	0.8678
MERA	40122	13.38	0.9920	0.9597
HQNN-Parallel	50714	1.50	0.9933	0.9658
HQNN-Quanv	1986	0.03	0.8838	0.2121
Ours	43178	13.15	<b>0.9942</b>	<b>0.9716</b>

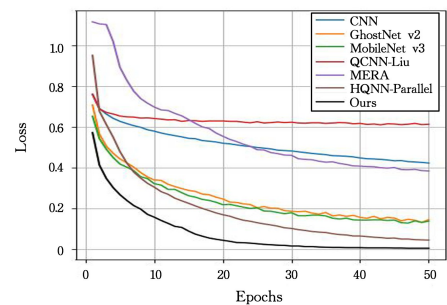
表4 CIFAR-10上的对比实验结果

Table 4 Comparative experimental results on the CIFAR-10

模型	参数量	FLOPs	准确率	精确率
CNN	88083	91.51	0.8329	0.7485
GhostNet v2	96747	64.92	0.8991	0.8482
MobileNet v3	89119	67.56	0.8856	0.8273
QCNN-Hen	22355	14.34	0.6922	0.5374
QCNN-Liu	104263	64.80	0.8011	0.7004
MERA	75587	62.90	0.8856	0.8302
HQNN-Parallel	262567	76.46	0.9118	0.8690
HQNN-Quanv	22363	14.34	0.6947	0.5499
Ours	89555	59.62	<b>0.9142</b>	<b>0.8718</b>



(a)



(b)

图5 训练损失值变化

Fig. 5 Changes in training loss

从经典网络对比实验中可以看出,在参数规模接近的前提下,本文模型可以取得最高的准确率与精确率,除在 MNIST 实验中的 GhostNet v2 外还具有最低的 FLOPs。这说明本文模型在图像分类任务中,无论是分类性能还是加速能力,都比常用的经典轻量化模型有切实的优势。在混合模型的对比实验中,本文模型的参数量与 FLOPs 均多于 QCNN-Hen, HQNN-Quanv,且在 MNIST 实验中也不能全面超过 QCNN-Liu, MERA, HQNN-Parallel,但本文模型取得了最高的准确率与精确率,除 QCNN-Hen, HQNN-Quanv 以外参数量与计算量并没有落后太多,在 CIFAR-10 的任务中还实

现了更低的参数量与计算量。这说明本文模型相比其他混合模型,在分类性能与轻量化性能间实现了较好的平衡。同时,由图 5 可知,本文模型具有相对更快的收敛速度。以上对比实验表明,本文模型在单通道的灰度图像和 RGB 彩色图像上的分类任务都有较强的多分类能力,还能有效降低算法规模与计算量。

本文另外进行了其他实验来进一步验证本文模型构型的有效性。具体来说,在针对并行构型的实验中,将原本的并行结构前后串联,改为串行结构进行实验;在针对量子模块的实验中,将所设计的量子模块替换为经典卷积实现的 DFC 注意力机制<sup>[6]</sup>、量子硬件高效电路<sup>[31]</sup>、量子非局部电路<sup>[15]</sup>;在针对 GFE 中的参数  $D$  实验中,分别设置  $D$  为 1,3,4。这部分的实验结果如表 5 和表 6 所列。

表 5 MNIST 上相似结构模型的实验结果

Table 5 Experimental results of similar structure models on the MNIST

模型	参数量	FLOPs	准确率	精确率
串行模型	54570	21.08	0.9860	0.9420
经典替换	153546	24.63	0.9885	0.9293
硬件高效电路替换	43114	13.15	0.9857	0.9420
量子非局部电路替换	43178	13.15	0.9305	0.6479
$D=1$	43114	13.15	0.9942	0.9714
$D=3$	43242	13.15	0.9942	0.9715
$D=4$	43306	13.15	0.9930	0.9644
Our Model	43178	13.15	<b>0.9942</b>	<b>0.9716</b>

表 6 CIFAR-10 上相似结构模型的实验结果

Table 6 Experimental results of similar structure models on the CIFAR-10

模型	参数量	FLOPs	准确率	精确率
串行模型	113955	42.69	0.8893	0.8341
经典替换	199923	71.09	0.9102	0.8662
硬件高效电路替换	89491	59.62	0.9098	0.8642
量子非局部电路替换	89555	59.62	0.9127	0.8693
$D=1$	89491	59.62	0.9073	0.8603
$D=3$	89619	59.62	0.9102	0.8647
$D=4$	89683	59.62	0.9078	0.8637
Our Model	89555	59.62	<b>0.9142</b>	<b>0.8718</b>

对实验结果进行分析可知,在 MNIST 十分类任务以及 CIFAR-10 三分类任务中,本文模型在性能上的表现全面优于其他构型,但是参数量略微大于硬件高效电路替换的模型。串行模型在较为简单的 MNIST 十分类任务上表现较好,但在较为复杂的 CIFAR-10 三分类任务上性能下降,这可能是由量子模块使用的振幅编码、多层池化等对数据的过度压缩导致的,并行结构则通过多分支分离解决了这个问题。另外可以看出,使用经典模块进行替换后的模型不仅参数量增加,准确率也出现下降;硬件高效电路以及量子非局部电路的表现则接近本文模型。最后, $D$  值的不同也会影响模型的性能,在  $D$  为 4 时可取得与为 2 时接近的准确率,但  $D$  的增大必然带来复杂度的上升,因此本文的实验均使用  $D$  为 2 的模型。

**结束语** 本文提出了基于经典卷积层、全连接层等经典部分及混合量子经典特征扩展模块的量子经典长-短距离特征扩展网络,并在整体上表现为一种分支结构。混合量子经典特征扩展模块中,通过参数化量子电路以及滑动窗口实现了量子解耦全连接注意力机制,能充分利用量子计算的优势

高效地提取长距离特征,压缩模型参数量,保证模块扩展性。除量子模块外,该混合模块还增加了负责提取短距离特征的经典卷积层,并利用合并操作得到扩展特征。在 MNIST 和 CIFAR-10 上的分类实验表明,所提出模型可用于解决图像分类问题并具有更高的准确率,且与规模相当的经典模型相比具有更低的计算量;与目前部分其他混合量子经典神经网络相比,则可以实现性能与轻量化的平衡,适用于处理海量图像数据的分类问题。

当然,本文模型也存在着一些局限性:1)目前的方法还基于经典计算方法,如果数据集足够大、模型足够复杂,则该方法可能会不够高效,需探索更高效的编码方式及量子模块,同时,尽量减少经典结构去实现全量子神经网络能很大程度上提高模型的效率;2)因为现阶段处于 NISQ 时代,量子算法很容易被噪声影响,所以需测试本文所提出算法的抗噪能力,并依据实际实验情况采取改进措施。

## 参考文献

- [1] CHEN F, LI S, HAN J, et al. Review of lightweight deep convolutional neural networks[J]. Archives of Computational Methods in Engineering, 2024, 31(4): 1915-1937.
- [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] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861, 2017.
- [4] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 1314-1324.
- [5] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 1580-1589.
- [6] TANG Y, HAN K, GUO J, et al. GhostNetv2: Enhance cheap operation with long-range attention[J]. Advances in Neural Information Processing Systems, 2022, 35: 9969-9982.
- [7] SU T, LIU A, SHI Y, et al. IremulbNet: Rethinking the inverted residual architecture for image recognition[J]. Neural Networks, 2024, 172: 106140.
- [8] ZHANG H, MA Y, ZHANG K, et al. FMGNet: An efficient feature-multiplex group network for real-time vision task[J]. Pattern Recognition, 2024, 156: 110698.
- [9] SCHULD M, FINGERHUTH M, PETRUCCIONE F. Implementing a distance-based classifier with a quantum interference circuit[J]. Europhysics Letters, 2017, 119(6): 60002.
- [10] WANG J, ZHANG R, JIANG N. Survey on Quantum Machine Learning[J]. Journal of Software, 2024, 35(8): 3843-3877.
- [11] LLOYD S, MOHSENI M, REBENTROST P. Quantum algorithms for supervised and unsupervised machine learning[J]. arXiv:1307.0411, 2013.
- [12] FEDOROV D A, PENG B, GOVIND N, et al. VQE method: a

- short survey and recent developments[J]. *Materials Theory*, 2022,6(1):2.
- [13] HUR T,KIM L,PARK D K. Quantum convolutional neural network for classical data classification[J]. *Quantum Machine Intelligence*,2022,4(1):3.
- [14] HENDERSON M,SHAKYA S,PRADHAN S,et al. Quantvolutional neural networks: powering image recognition with quantum circuits[J]. *Quantum Machine Intelligence*,2020,2(1):2.
- [15] GUPTA S,KONAR D,AGGARWAL V. A Scalable Quantum Non-local Neural Network for Image Classification[J]. arXiv: 2407.18906,2024.
- [16] SENOKOSOV A,SEDYKH A,SAGINGALIEVA A,et al. Quantum machine learning for image classification[J]. *Machine Learning: Science and Technology*,2024,5(1):015040.
- [17] RAO G V E, RAJITHA B, SRINIVASU P N, et al. Hybrid framework for respiratory lung diseases detection based on classical CNN and quantum classifiers from chest X-rays[J]. *Biomedical Signal Processing and Control*,2024,88:105567.
- [18] XIONG Q B, MIAO Q G, YANG T. Malicious Code Detection Method Based on Hybrid Quantum Convolutional Neural Network [J]. *Computer Science*,2025,52(3):385-390.
- [19] FAN F,SHI Y,GUGGEMOS T,et al. Hybrid quantum-classical convolutional neural network model for image classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024,35(12):18145-18159.
- [20] LIU J, LIM K H, WOOD K L, et al. Hybrid quantum-classical convolutional neural networks[J]. *Science China Physics, Mechanics & Astronomy*,2021,64(9):290311.
- [21] FENG Y,LI Y,LI J,et al. Variational Shadow Quantum Circuits Assisted Quantum Convolutional Neural Network[J]. *Advanced Quantum Technologies*,2025,8(8):2400510.
- [22] WEI S J,CHEN Y H,ZHOU Z R,et al. A quantum convolutional neural network on NISQ devices[J]. *AAPPS Bulletin*, 2022,32:1-11.
- [23] LIU Z,LIN Y,CAO Y,et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021: 10012-10022.
- [24] MEHTA S,RASTEGARI M. Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer[J]. arXiv:2110.02178,2021.
- [25] SHI J,ZHAO R X,WANG W,et al. QSAN: A near-term achievable quantum self-attention network[J]. *IEEE Transactions on Neural Networks and Learning Systems*,2025,38(8): 13995-14008.
- [26] ZHAO R X,SHI J,LI X. QKSAN: A Quantum Kernel Self-Attention Network [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2024,46(12):10184-10195.
- [27] JADERBERG M,VEDALDI A,ZISSERMAN A. Speeding up convolutional neural networks with low rank expansions[J]. arXiv:1405.3866,2014.
- [28] VASWANI A,SHAZEER N,PARMAR N,et al. Attention is all you need[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017: 6000-6010.
- [29] LIU J, LIM K H, WOOD K L, et al. Hybrid quantum-classical convolutional neural networks[J]. *Science China Physics, Mechanics & Astronomy*,2021,64(9):290311.
- [30] GRANT E,BENEDETTI M,CAO S,et al. Hierarchical quantum classifiers[J]. *NPJ Quantum Information*,2018,4(1):65.
- [31] 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.



**ZHENG Yi**, born in 2000, postgraduate. His main research interests include quantum computing and machine learning.



**REN Shuang**, born in 1981, Ph.D, associate professor, Ph.D supervisor. His main research interests include machine learning, quantum computing and 3D computer vision.

(责任编辑:柯颖)