

# 一种量子衍生神经网络模型<sup>①</sup>

李滨旭<sup>1</sup>, 姚姜虹<sup>2</sup>

<sup>1</sup>(东北石油大学 计算机与信息技术学院, 大庆 163318)

<sup>2</sup>(大庆市油田信息技术公司物联网分公司, 大庆 163318)

**摘要:** 为提高神经网络的逼近能力, 通过在普通 BP 网络中引入量子旋转门, 提出了一种新颖的量子衍生神经网络模型. 该模型隐层由量子神经元组成, 每个量子神经元携带一组量子旋转门, 用于更新隐层的量子权值, 输入层和输出层均为普通神经元. 基于误差反传播算法设计了该模型的学习算法. 模式识别和函数逼近的实验结果验证了提出模型及算法的有效性.

**关键词:** 量子计算; 量子旋转门; 量子衍生神经元; 量子衍生神经网络

## Quantum-Inspired Neural Networks Model and Algorithm

LI Bin-Xu<sup>1</sup>, YAO Jiang-Hong<sup>2</sup>

<sup>1</sup>(School of Computer & Information Technology, Northeast Petroleum University, Daqing 163318, China)

<sup>2</sup>(Daqing Petroleum Information Technology IoT branch Company, Daqing 163318, China)

**Abstract:** To enhance the approximation ability of neural networks, by introducing quantum rotation gates to the traditional BP networks, a novel quantum-inspired neural network model is proposed in this paper. In our model, the hidden layer consists of quantum neurons. Each quantum neuron carries a group of quantum rotation gates which are used to update the quantum weights. Both input and output layer are composed of the traditional neurons. By employing the error back propagation algorithm, the learning algorithms are designed. Simulation-based experiments using two application examples of pattern recognition and function approximation, respectively, illustrates the availability of the proposed model.

**Key words:** quantum computing; quantum rotation gate; quantum-inspired neuron; quantum-inspired neural network

## 1 引言

上世纪 80 年代, Benioff 首次提出了量子计算的概念<sup>[1]</sup>, 1994 年, Shor 提出了大数质因子分解的量子算法<sup>[2]</sup>. 1996 年, Grover 提出了无序数据库搜索的量子算法<sup>[3]</sup>. 目前虽然量子计算的发展还远未成熟, 但量子算法与经典算法的计算机机制是截然不同的, 其高度并行的优良特性已引起国内外学者的广泛关注. 模糊逻辑、进化计算、神经元网络是人工智能领域的三个明珠, 三者统称为智能计算或软计算, 其与量子计算某种程度上存在融合的可行性. 1989 年, Penrose 指出人脑信息处理过程与量子信息处理有某种程度的相似性, 尽管这一理念目前没有被证实, 但有关量子信息处理

的理论探索, 已经成为一个极富应用前景的研究领域. 1995 年, Kak 首次提出了量子神经计算的概念<sup>[4]</sup>. 随后的二十年间, 量子神经网络研究吸引了很多学者的注意, 并已提出一些新颖的模型和算法. 基于量子计算中的叠加态, 文献[5]提出了一种含多个隐层的量子神经网络模型; 文献[6]提出了一种量子门线路的神经网络模型; 文献[7]提出了一种基于量子线路的无权神经网络模型, 该模型基于 Grover 算法训练, 是能在将来的量子计算机上运行的真正意义上的量子神经网络; 文献[8]提出了一种量子门节点神经网络模型和算法, 该模型比传统神经网络包含更多的生物神经系统的特征; 基于单比特量子旋转门和两比特量子受控非

① 基金项目: 东北石油大学研究生创新科研项目(YJSCX2016-030NEPU)

收稿时间: 2016-01-29; 收到修改稿时间: 2016-02-29 [doi:10.15888/j.cnki.csa.005346]

门,文献[9]提出了一种量子BP网络模型和算法。

在量子神经网络的应用方面,已分别应用于频谱感知<sup>[10]</sup>、配电网实时故障定位<sup>[11]</sup>、音频水印<sup>[12]</sup>、岩性识别<sup>[13]</sup>等许多领域的实际问题。

本文研究一种混合量子衍生神经网络模型,该模型为三层结构,隐层为量子神经元,输出层为传统神经元,基于量子计算原理给出了模型的输入输出关系,采用梯度下降法详细设计了该模型的学习算法。模式识别和函数逼近的实验结果表明,提出的模型的收敛率、迭代步数、逼近误差均优于传统BP神经网络。

## 2 量子衍生神经网络模型

### 2.1 量子比特和量子旋转门

在量子计算中,量子比特是存储信息的最小单位,用于描述线路的状态。量子比特有两个基态 $|0\rangle$ 和 $|1\rangle$ ,符号“ $\bullet$ ”称为狄拉克符号。量子比特与经典比特的区别在于,量子比特可以处于两个基态的线性叠加态,如下式所示。

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

其中 $\alpha$ 和 $\beta$ 是复数,称为量子比特的概率幅。若对量子比特实时测量,它将以 $|\alpha|^2$ 的概率坍缩到 $|0\rangle$ 态,或者以 $|\beta|^2$ 的概率坍缩到 $|1\rangle$ 态,两个概率幅之间满足如下归一化关系。

$$|\alpha|^2 + |\beta|^2 = 1 \quad (2)$$

因此量子比特也能用概率幅表述为 $[\alpha, \beta]^T$ 。

与单比特系统相似,对于 $n$ 比特量子系统,有 $2^n$ 个基态,该系统也可处于这些基态的线性叠加态,如下式所示。

$$|\phi\rangle = \sum_{x \in \{0,1\}^n} a_x |x\rangle \quad (3)$$

其中 $a_x$ 是基态 $|x\rangle$ 的概率幅,这些概率幅之间也满足归一化关系。

$$\sum_{x \in \{0,1\}^n} |a_x|^2 = 1 \quad (4)$$

在量子计算中,在量子态上执行的一系列酉变换称为量子门,它是执行量子计算的基础。单比特量子旋转门的定义如下式。

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (5)$$

记量子态 $|\phi\rangle = \begin{bmatrix} \cos\theta_0 \\ \sin\theta_0 \end{bmatrix}$ ,  $R(\theta)$ 在 $|\phi\rangle$ 上的作用效果如下式所示。

$$R(\theta)|\phi\rangle = \begin{bmatrix} \cos(\theta_0 + \theta) \\ \sin(\theta_0 + \theta) \end{bmatrix} \quad (6)$$

显然, $R(\theta)$ 在 $|\phi\rangle$ 上执行了相位旋转。

### 2.1 量子衍生神经元模型

一个神经元可描述为一个四元组:输入、权值、激励函数、输出。输入和输出是神经元的外部特性,而权值和激励函数是神经元的内部特性。所以通过修改权值和激励函数可以构造不同类型的神经元模型。根据这种观点,本文提出的量子衍生神经元,权值采用量子比特描述,激励函数采用内积算子。与传统神经元模型的区别在于,每个量子衍生神经元携带一组量子旋转门,用于更新量子权值的相位。模型如图1所示。

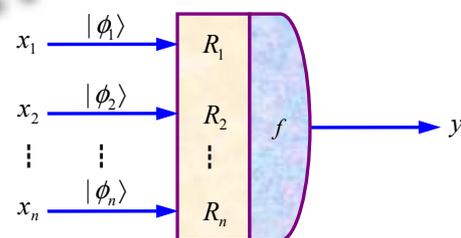


图1 量子衍生神经元模型

在图1中,令权值 $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle = [\alpha, \beta]^T$ ,输入向量 $X = (x_1, x_2, \dots, x_n)^T$ ,量子衍生神经元的输入输出关系可表述为下式。

$$y = f(XR|\phi\rangle) = C \cdot \sum_{i=1}^n x_i R_i |\phi_i\rangle \quad (7)$$

其中“ $\cdot$ ”为内积算子, $f(X) = C \cdot X$ ,  $C = [1, 1]^T$ ,  $R_i$ 是量子旋转门用于修改量子权值的相位。

### 2.3 量子衍生神经网络模型

在量子衍生神经网络(quantum-inspired neural networks, QINN)中,输入、输出、各层链接权都可以是量子比特。QINN的结构和普通神经网络是相同的,也包括输入层、隐层、输出层。很明显,包含图1中量子衍生神经元的网络是量子衍生神经网络。全部采用量子神经元的网络称为正规量子神经网络,同时采用量子神经元和传统神经元的网络称为混合量子神经网络。由于量子神经元采用线性激励函数,其映射能力不够理想,所以本文采用包含一个量子神经元隐层的混合量子神经网络模型(hybrid QINN, HQINN),它融合了量子计算和传统神经网络两者的优点,模型如图2所示。

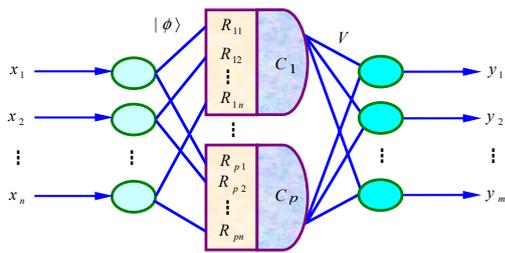


图 2 混合量子衍生神经网络模型

该模型包含三层，输入层和输出层各包含  $n$  和  $m$  个普通神经元，隐层包含  $p$  个量子衍生神经元，输入输出关系如下式所示。

$$\left. \begin{aligned} h_j &= C_j \cdot \sum_{i=1}^n x_i R_{ij} |\phi_{ij}\rangle \\ y_k &= g(\sum_{j=1}^p v_{jk} h_j) \end{aligned} \right\} \quad (8)$$

其中  $i=1,2,\dots,n, j=1,2,\dots,p, k=1,2,\dots,m, v_{jk}$  是隐层第  $j$  个量子衍生神经元和输出层第  $k$  个普通神经元之间的连接权， $g$  是输出层激励函数，本文采用 Sigmoid 函数。

### 2.4 学习算法

对于图 2 所示的 HQINN，若输出层激励函数连续可微，则可采用 BP 算法调整网络参数。输出误差函数定义为下式。

$$E = \frac{1}{2} \sum_{k=1}^m (\hat{y}_k - y_k)^2 \quad (9)$$

其中  $\hat{y}_k$  为期望输出。

根据梯度下降法，输出层权值按下式调整。

$$\begin{aligned} v_{jk}(t+1) &= v_{jk}(t) + \eta \frac{\partial E}{\partial v_{jk}} \\ &= v_{jk}(t) + \eta_1 (\hat{y}_k - y_k(t)) g'(t) h_j(t) \end{aligned} \quad (10)$$

其中  $j=1,2,\dots,p, k=1,2,\dots,m, \eta$  为学习速率， $t$  为当前为迭代步数。

隐层量子权值由量子旋转门按下式调整。

$$R_{ij} = \begin{bmatrix} \cos(\Delta\theta_{ij}) & -\sin(\Delta\theta_{ij}) \\ \sin(\Delta\theta_{ij}) & \cos(\Delta\theta_{ij}) \end{bmatrix} \quad (11)$$

其中  $i=1,2,\dots,n, j=1,2,\dots,p$ 。

根据梯度下降法  $\Delta\theta_{ij}$  按下式计算。

$$\begin{aligned} \Delta\theta_{ij}(t) &= \frac{\partial E}{\partial \theta_{ij}} = -\sum_{k=1}^m \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial h_j} \frac{\partial h_j}{\partial \theta_{ij}} \\ &= -\sum_{k=1}^m (\hat{y}_k - y_k(t)) g'(t) v_{jk}(t) h_j(t) x_i \\ &\quad (\cos(\phi_{ij}(t)) - \sin(\phi_{ij}(t))) \end{aligned} \quad (12)$$

其中  $\phi_{ij}(t)$  是  $|\phi_{ij}\rangle$  的相位。记  $|\phi_{ij}\rangle = [\alpha_{ij}, \beta_{ij}]^T$ ， $|\phi_{ij}\rangle$  具体可按下式修改。

$$\begin{bmatrix} \alpha_{ij}(t+1) \\ \beta_{ij}(t+1) \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta_{ij}) & -\sin(\Delta\theta_{ij}) \\ \sin(\Delta\theta_{ij}) & \cos(\Delta\theta_{ij}) \end{bmatrix} \begin{bmatrix} \alpha_{ij}(t) \\ \beta_{ij}(t) \end{bmatrix} \quad (13)$$

其中  $t$  为当前迭代步数。

关于本文提出的 HQINN 的收敛性，我们给出如下结论：当隐层神经元足够多时，该网络能够逼近  $[0,1]^n$  上的任意 L2 函数。对此结论，我们简要解释如下。

根据 HQINN 隐层量子神经元输出式(8)可知， $h_j$  为量子旋转门  $R_{ij}$  的函数，由式(11)知， $R_{ij}$  为  $\Delta\theta_{ij}$  正余弦函数；因此  $h_j$  为  $\Delta\theta_{ij}$  的复合函数。由于正余弦函数及其复合函数均连续可微，所以  $h_j = f(\cdot)$  连续可微；由量子衍生神经元输出式(7)可知， $h_j$  必有界。因此量子神经元可视传统神经元在激励函数、聚合方式两方面的推广，即将传统的 Sigmoid 函数改为  $f(\cdot)$ ，而将对输入的加权聚合改为相位移动。HQINN 的输出层就是普通神经元。因此 HQINN 可视为一种特殊形式的传统两层神经网络。根据文献[14]的结论“具有两层连接权值的误差反向传播神经网络，在隐层节点足够多的情况下能够逼近  $[0,1]^n$  上的任意 L2 函数”可知，本文提出的 HQINN 是收敛的。

综上所述，HQINN 学些算法可描述如下。

Step 1: 初始化网络参数。包括：各层节点数、学习速率、限定误差  $\epsilon$ 、限定步数  $Max$ 、置当前步数  $t=0$ 。

Step 2: 初始化网络权值。隐层： $\phi_{ij}(t) = 2\pi \times rnd, |\phi_{ij}\rangle = \begin{bmatrix} \cos(\phi_{ij}(t)) \\ \sin(\phi_{ij}(t)) \end{bmatrix}$ ，输出层： $v_{jk} = rnd - 0.5$ ，其中  $rnd$  为  $(0,1)$  内均匀分布的随机数， $i=1,2,\dots,n, j=1,2,\dots,p, k=1,2,\dots,m$ 。

Step 3: 根据式(8)计算网络输出，根据式(10)和式(13)修改网络权值。

Step 4: 根据式(9)计算网络输出误差，如果  $E < \epsilon$  或  $t > Max$  转步 5，否则  $t = t + 1$  转步 3。

Step 5: 保存权值，结束。

### 3 对比实验

为验证 HQINN 的性能，本文设计了两类实验并与传统 BP 网络对比。为使对比公平，两种模型采用相同的结构和参数。

### 3.1 平面点集分类

在图 3 所示平面上有两类 25 个样本点, 试用 HQINN 对其分类. 这是一个典型的两类模式识别问题. 可以看做异或问题的推广.

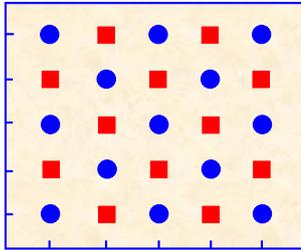


图 3 平面点集分类问题样本

网络结构取 2-10-1 型, 学习速率取 0.8. 采用图 3 中的 25 个样本, HQINN 和普通 BP 网络各训练 50 次, 然后统计平均迭代步数和收敛率, 训练结果对比如表 1 所示, 当限定误差取 0.05 时, 收敛曲线对比如图 4 所示.

表 1 平面点集分类问题训练结果对比

限定误差	限定步数	HQINN		BP	
		迭代步数	收敛率	迭代步数	收敛率
0.10	1000	583	90%	1000	不收敛
0.05	2000	1131	84%	2000	不收敛
0.01	10000	8893	82%	10000	不收敛

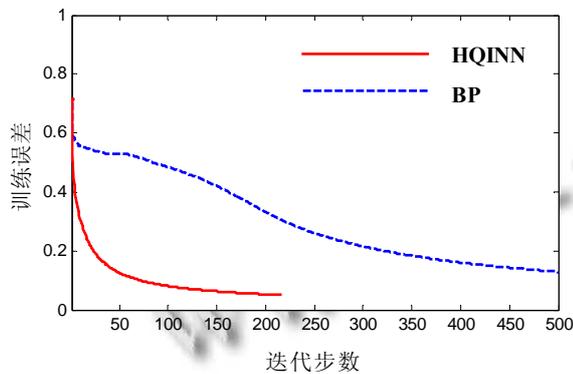


图 4 平面点集分类问题收敛曲线对比

### 3.2 双螺旋线分类

对于图 5 所示的双螺旋曲线, 每条曲线上分别取 25 个点, 组成包含 50 个样本的训练样本集, 试用 HQINN 对这些点分类. 网络结构却 2-20-1 型, 学习速率取 0.7, HQINN 和普通 BP 网络各训练 30 次, 然后统计平均迭代步数和收敛率, 训练结果对比如表 2 所示, 当限定误差取 0.10 时, 收敛曲线对比如图 6 所示.

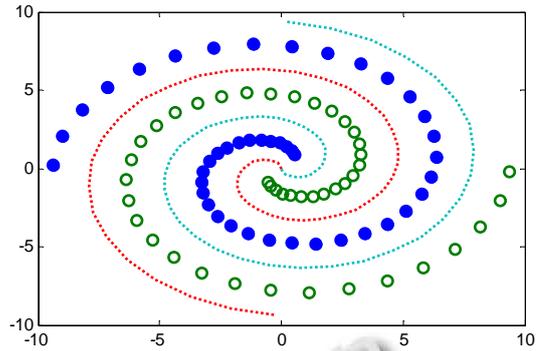


图 5 双螺旋先分类问题样本

表 2 双螺旋线分类问题训练结果对比

限定误差	限定步数	HQINN		BP	
		迭代步数	收敛率	迭代步数	收敛率
0.10	2000	871	82%	2000	不收敛
0.05	10000	6893	92%	8631	76%

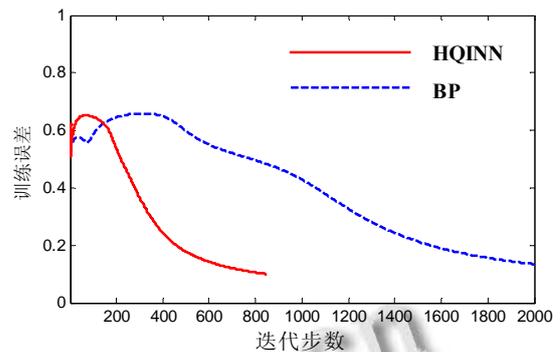


图 6 双螺旋线分类问题收敛曲线对比

### 3.3 函数逼近问题

本实验考虑如下函数,

$$y = (1.0 + \sqrt{x_1} + 1/x_2 + x_3^{-1.5})^2 \quad (14)$$

其中  $x_1, x_2, x_3 \in \{1, 2, 3, 4, 5\}$ .

为逼近上述非线性函数, 我们随机取 40 组离散数据, 前 20 组用于训练网络, 使其逼近式(14)所示的非线性函数关系, 余下的 40 组作为测试集, 检验网络的泛化推广能力. 最大迭代步数取 5000. 通过改变隐层节点数和学习速率, 我们详细考察了 HQINN 的性能, 并与其它方法进行了对比.

首先考察不同隐层节点数下的性能对比. 将学习速率固定为 0.8, 隐层节点数分别取 6 和 8, 分别用 HQINN、BP 网络、文献[12]中的网络独立训练 30 次, 并分别测试网络的泛化推广能力, 取 30 次平均误差作为对比指标, 实验结果对比如表 3 所示.

表 3 函数逼近问题实验结果对比

算法	6 个隐层节点		8 个隐层节点	
	训练误差	测试误差	训练误差	测试误差
HQINN	0.035236	0.635218	0.011836	1.008653
BP	0.083605	3.268351	0.029605	2.831532
文献[12]	0.019852	1.775031	0.019324	4.302806

下面考察学习速率对收敛率的影响. 网络结构取 3-5-1 型, 限定误差取 0.1, 限定步数取 5000, 学习速率分别取, 对于学习速率的每种取值, 分别用 HQINN 和 BP 网络独立训练 30 次, 当学习速率变化时, HQINN 的收敛率均为 100%, 而 BP 网络的收敛率变化很大, 最高为 80%, 最低只有 10%. 对比结果如图 7 所示.

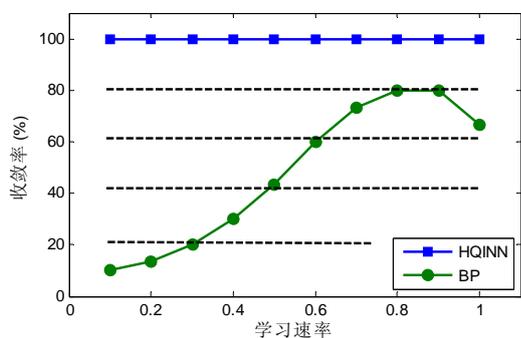


图 7 收敛率和学习速率的关系对比

对比结果表明, HQINN 在收敛率和鲁棒性两方面, 均明显优于 BP 网络.

最后考察学习速率对迭代步数的影响. 网络结构取 3-5-1 型, 限定误差取 0.1, 学习速率分别取, 对于学习速率的每种取值, 分别用 HQINN 和 BP 网络独立训练 30 次, 当学习速率变化时, HQINN 的平均迭代步数最大为 958 步, 最小仅为 496 步, 平均为 693 步; 然而对于 BP 网络, 最大为 4619 步, 最小为 1532 步, 平均为 2489 步. 因此当学习速率变化时, HQINN 不仅平均迭代步数明显少于 BP 网络, 而且波动范围也较小. 对比结果如图 8 所示. 对比结果表明, HQINN 在平均迭代步数和鲁棒性两方面均明显优于 BP 网络.

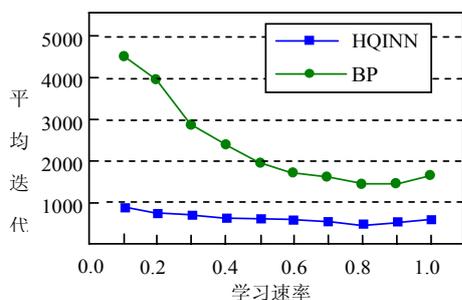


图 8 迭代步数和学习速率的关系对比

## 4 结论

HQINN 是对量子计算和神经计算的模拟, 兼具量子计算的高效性和神经计算的连续性、逼近能力和推广能力等优点. 该模型的隐层权值采用量子比特描述, 采用量子旋转门更新量子比特的两个概率幅, 借助概率幅三角函数的周期性, 可使调整过程更为精细, 从而有效增强了网络的逼近能力. 实验结果表明, 提出的模型明显优于普通 BP 网络.

### 参考文献

- Benioff P. Quantum mechanical hamiltonian models of Turing machines. *Journal of Statistical Physics*, 1982, 3: 515-546.
- Shor PW. Algorithm for quantum computing: Discrete logarithms and factoring. *Proc. of the 35th Annual Symp on Foundations of Computer Science*. New Mexico: IEEE Computer Society Press. 1994. 124-134.
- Grover LK. A fast quantum mechanical algorithm for database search. *Proc. of the 28th Annual ACM Symposium on the Theory of Computing*. Pennsylvania. 1996. 212-221.
- Kak S. On quantum neural computing. *Information Sciences*, 1995, 83: 143-160.
- Gopathy P, Nicolaos BK. Quantum Neural networks (QNN's): Inherently fuzzy feedforward neural networks. *IEEE Trans. on Neural Networks*, 1997, 8: 679-693.
- Li PC, Xiao H. Model and algorithm of quantum inspired neural network with sequence input based on controlled rotation gates. *Applied Intelligence*, 2014, 40(1): 107-126.
- Adenilton J, Wilson R, Teresa B. Classical and superposed learning for quantum weightless neural network. *Neurocomputing*, 2012, 75: 52-60.
- Li PC, Xiao H. Sequence input-based quantum-inspired neural networks with applications. *Neural Processing Letters*. 2014, 40(2): 143-168.
- Li PC, Li SY. Learning algorithm and application of quantum BP neural networks based on universal quantum gates. *Journal of Systems Engineering and Electronics*, 2008, 19: 167-174.
- 赵阳, 孙学斌, 周正. 一种改进的量子神经网络频谱感知算法. *通信系统与网络技术*, 2015, 41(2): 7-11.
- 马亮, 杨萍萍, 高建宇. 基于动态云—量子神经网络群的配电网实时故障定位方法. *工矿自动化*, 2014, 40(11): 71-75.
- 张翼鹏, 陈亮, 郝欢. 采用量子神经网络的音频水印新算法. *信号处理*, 2013, 29(6): 684-690.
- 王团. 量子神经网络在雷家地区岩性识别中的应用. *长春大学学报*, 2015, 25(4): 56-59.
- Hecht-Nielsen R. Theory of the back propagation neural network. *Proc. of the Int Joint Conf on Neural Networks*. New York: IEEE Communications Society. 1989. 593-605.