Methodology in AI
Literature Survey on Quantum Neural Networks

Seeram Sumanth (S5652588) Kyriakos Antoniou (S5715881)

February 2nd, 2024

1 Introduction

Quantum computing has been the center of the focus in the past couple of years, with IBM, Google, and Amazon leading from the front. Over the years, quantum algorithms have solved certain problems more efficiently than their classical counterparts, algorithms such as Shor's algorithm [7] were able to break through the RSA encryption protocol in a reasonable time.

Quantum Neural Networks are an emerging class of neural networks that integrates the principles of quantum computing into artificial neural networks.QNNs are based on parameterized quantum circuits that can be trained in a variational manner using classical optimizers. These circuits contain a feature map (with input parameters) and an ansatz (with trainable weights).

In this survey, we explore the concept of Quantum Neural Networks, assess whether they can outperform classical neural networks, and understand the benefits and pitfalls they offer compared to classical neural networks. Furthermore, we assess whether their advancement is still in the early phase.

2 Background

Quantum Computing harnesses the principles of quantum mechanics to perform complex computations for classical computers. The basic units of computation in Quantum Computing are termed qubits.

2.1 Qubits and Superposition

Unlike classical computing bits that have fixed 0 and 1 values, qubits can be in any linear combination of 0 and 1, known as a qubit's superposition. The state of a qubit can be represented as:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where $|\psi\rangle$ represents the state of the qubit, and α and β are complex numbers that determine the probability amplitudes of the qubit being in the state $|0\rangle$ and $|1\rangle$, respectively.

2.2 Entanglement

Entanglement is a concept that allows for efficient computations in quantum machines. It allows two or more qubits to be linked together so that they are interdependent. The phenomenon can be described using Alice and Bob, Alice is an observer for system A, and Bob is an observer for system B. If in the entangled state given above Alice makes a measurement, there are two possible outcomes, occurring with equal probability:

$$\text{If Alice measures 0, the system collapses to } |0\rangle_A \otimes |1\rangle_B$$

$$\text{If Alice measures 1, the system collapses to } |1\rangle_A \otimes |0\rangle_B$$

If the former occurs, then any subsequent measurement performed by Bob on the same basis will always return 1. If the latter occurs (Alice measures 1), then Bob's measurement will return 0 with certainty. Thus, Alice has altered system B by performing a local measurement on system A. This remains true even if the systems A and B are spatially separated

2.3 Feature Map

Quantum feature map encodes classical data to the quantum state space by using a quantum circuit, such that quantum algorithms can process the data further.

2.4 Decoherence

Decoherence is how a quantum system interacts with its environment and loses its quantum properties. This interaction causes the system's wave function to collapse, making it behave more predictably and classically.

2.5 Effective Dimension

The effective dimension measures a model's ability to generalize to new data. It includes both global and local effective dimensions, with the local one focusing on the trained model's expressiveness.

2.6 Fisher Information Matrix

The Fisher Information Matrix captures how sensitive a neural network's output is to network parameter changes. It is used in natural gradient optimization to navigate the parameter space efficiently.

2.7 Barren Plateaus

A barren plateau is a phenomenon where the cost function gradient with respect to the parameters is close to zero across a vast region of the parameter space. This makes it difficult for optimization algorithms to find a direction that leads to the minimum of the cost function. This is often due to quantum noise and imperfections in the quantum hardware.

2.8 Quantum Noise

Quantum noise includes quantum decoherence and fluctuations, which pose significant challenges to developing practical quantum computers. It refers to the random fluctuations that affect quantum systems arising from the inherent uncertainty and the probabilistic nature of quantum mechanics.

2.9 Parameterized Quantum Circuits

Parameterized Quantum Circuits (PQCs), also known as variational quantum circuits or quantum ansatzes, are a central component in quantum computing, particularly in quantum machine learning and variational quantum algorithms. PQCs are quantum circuits with tunable parameters, which can be adjusted to optimize the circuit's output for a specific task.

2.10 Over-parameterized Quantum Neural Networks

Over-parameterized Quantum Neural Networks (PQNs) have more parameters than necessary, giving them a high degree of freedom to learn complex patterns in data.

2.11 NISQ Devices

NISQ devices (Noisy Intermediate-Scale Quantum Computing) refer to the current generation of quantum computers. These devices are characterized by their relatively small size, typically ranging from a few dozen to a few hundred qubits, and their susceptibility to errors and noise.

2.12 Open Quantum Systems

Open quantum systems refer to quantum systems that interact with their environment instead of closed quantum systems, which are ideally isolated from any external influence.

3 Annotated Bibliography

3.1 The Quest for Quantum Neural Networks

In the last 20 years, many researchers have suggested different ideas for quantum neural networks (QNNs), but the confusion about what exactly a quantum neural network is remains. Maria et al. [6] were the first to propose technical guidelines for the design of the QNNs. Indeed, these guidelines were later adopted in designing the quantum neuron by Cao et al. [3].

3.1.1 Criteria for Quantum Neural Networks

Maria et al [6], outline three key requirements for a system to be considered as a QNN:

1. The initial state of the quantum system must be able to take any binary input string of length N, process it, and generate an output string that is the closest match to the input by some measure of distance.

2. The QNN should reflect one or more of the neural computing mechanisms, such as the structure of neural networks, training and updating weights, and the use of activation functions.

3. The quantum system should also utilize or exhibit effects of quantum computing like superposition, entanglement, and interference and be consistent with quantum theory.

3.1.2 Challenges in Quantum Neural Network Design

As mentioned in the study by Maria et al. [6], the major problem lies in using non-linear activation functions, which contradicts the basic principle of linearity in quantum theory. Using non-linear functions is crucial for the model to solve trivial tasks involving learning complex patterns in the data. Based on the minimal requirements, the review of QNN approaches presented by various authors concluded that none of the presented ideas could completely claim to be a QNN based on the proposed requirements. This outlines the problem that when the proposed models tend to follow quantum theory, they do not obtain the non-linear convergence dynamics of the neural networks, and vice versa. They highlight the concepts of open quantum systems and the dissipation models, including an environment, as important factors to consider to investigate the concept of QNNs further.

3.2 The Power of Quantum Neural Networks

3.2.1 Introduction

In contrast to classical models, understanding the capacity of quantum neural networks is not well explored. The classical neural networks are known to produce highly degenerate Fisher Information matrices, which have the potential to slow down training significantly. In the case of quantum neural networks, no Fisher Information matrices analysis has been done. Amira et al. [1] demonstrate that quantum neural networks offer an advantage over classical neural networks in terms of higher effective dimension (capacity of the model) and faster training ability (Fisher Information Matrix) over classical neural networks. Amira et al. [1] designed a quantum neural network (briefly described and the overview is provided in the paper), which is studied with two choices of feature maps termed as a special feature map (referenced from a study), easily simulable feature map (simplified feature map with less number of quantum gates). The use of different feature maps of the quantum neural network is to investigate how data encoding techniques impact capacity and trainability. The trainability of the model is further verified for the quantum neural network on the ibmq_montreal 27-qubit device available through the IBM Quantum Experience via Qiskit. In terms of classical neural networks, the networks with and without biases and different activation functions are explored. In particular, RELU, leaky RELU, tanh, and sigmoid activations are considered.

3.2.2 Experimental Setup and Evaluation

Amira et al [1] deemed two models comparable if they share the same number of trainable parameters (d), input size (s_{in}), and output size (s_{out}), and considered $d \leq 100$, $s_{in} \in \{4, 6, 8, 10\}$ with $s_{out} = 2$ for the experiments. The Fisher information matrix of each model is computed 100 times using initial parameters sampled uniformly at random. The number of trainable parameters $d = 40$, input size $s_{in} = 4$ and output size $s_{out} = 2$ is fixed. All three models were trained on the first two classes of the Iris dataset, using $d = 8$ trainable parameters with full batch size. The ADAM optimizer with an initial learning rate of 0.1 is selected. For a fixed number of training iterations = 100, all models were trained over 100 trials and the average training loss is plotted along with ± 1 standard deviation. The experiment on the IBM hardware was performed only once, due to limited hardware availability.

The quantum neural network achieved the highest effective dimension compared to the classical neural networks. The quantum neural network converged quickly and had lower training loss than the chosen classical neural network. The data for evaluating the models is small and constrained by hardware limitations. Although the experiments were performed several times on the simulator, the results include only a single run on the quantum hardware, highlighting the challenge of accessing quantum computing resources for testing. Amira et al. [1] did not provide any information on the model structure they used for classical neural networks to compare with quantum neural networks.

3.2.3 Conclusion

A quantum neural network with an easier data encoding strategy increases the likelihood of encountering a barren plateau, while a harder data encoding strategy shows resilience. The authors conclude that the choice of the feature map is important in designing a powerful quantum neural network and choosing a simpler feature map impairs these advantages. The quantum neural networks can possess a desirable Fisher information spectrum that enables them to train faster and express more functions than comparable classical models. Amira et al. [1] point out that using a particular higher-order feature map in this study needs to be investigated along with the possibility of noise-induced barren plateaus. They also mention that understanding generalization performance on multiple datasets and larger models will prove insightful.

3.3 Quantum Recurrent Neural Networks (QRNN)

3.3.1 Introduction

Recurrent neural networks are the foundation of many sequence-to-sequence models in machine learning, such as machine translation and speech synthesis. Johannes Bausch [2] is the first to introduce the concept of Quantum Recurrent Neural Network (QRNN). He proposes the implementation of the QRNN and explains how it achieved notable performance in tasks like sequence learning and integer digit classification.

3.3.2 Experimental Setup

Johannes Bausch [2] initially experimented with sequence memorization to reproduce two sequences $44444\ldots$, $123123\ldots3$, to benchmark the optimizer (RMSprop, Adam, L-BFGS, SGD) and learning rate hyperparameters. Adam is chosen as an optimizer based on the results of the experiment. The experiment to learn XOR sequences is performed to explore which parameter initialization (quantum neuron) converges quickly to a certain validation threshold.

The (train: validate: test) ratio of $(11 : 1 : 2)$ of the MNIST dataset (10×10 downscaled images) is used to assess the performance of the QRNN.

The long sequence test consists of gene sequences of letters G, A, T, and C. A letter U is added within the first half of the string sequence, and the network's task is to identify the letter following the letter 'U'.

3.3.3 Conclusion

The performance of the designed QRNN is assessed on the MNIST dataset using data augmentation techniques, resulting in a notable accuracy of 95%. However, it didn't surpass the state-of-the-art results. In the long sequence tests, QRNN can retain trainability while RNN and LSTM show a clear increase in convergence time or fail to converge within the set limit of 16000 steps.

Johannes Bausch [2] concludes that because of the inability to simulate many qubits on classical hardware, the proposed QRNN is outclassed even by simple RNNs. The work has valuable insights as the architecture presented in the paper can be run on the current hardware (classical hardware), and this is the first recurrent quantum neural network presented. It also demonstrated that more than a few bits of size data can be ingested, and models with large parameter counts can be evaluated and trained. Due to the intrinsic capability of a QRNN to keep track of a quantum state, it holds promise to better capture the exponentially-growing phase space dimension of the system to be modeled.

3.3.4 Evaluation

Johannes Bausch [2] compared QRNN with RNN on the Image classification dataset, which is an untypical task for a recurrent network. The paper had too many trivial tasks, with each having very little information on the experimental procedure, their validation, testing, and their results.

3.4 Quantum Convolutional Neural Networks (QCNN)

3.4.1 Introduction

Tak et al. [4] propose a quantum neural network inspired by Classical Convolutional Neural Networks (CNN) and investigate the performance of various QCNN models for binary classification. These models are differentiated by different parameters like structure and data encoding methods on MNIST and Fashion MNIST datasets, comparing their performance to classical CNNs.

3.4.2 Experimental Setup

The simulation for binary classification of QCNN on the datasets MNIST (0, 1) and Fashion MNIST (shirt/top, trouser) was conducted with respect to various parameters, encoding techniques, and data preprocessing techniques like PCA and auto-encoding. To make a fair comparison between QCNN and CNN, all hyperparameters of the two methods are the same, except the Adam optimizer was used for CNN since it performed better. The binary classification of the QCNN models is performed on the quantum simulator Pennylane.

3.4.3 Conclusion

The results show that the QCNN models performed better than their corresponding CNN models and had considerably smaller standard deviations in accuracy than the CNN models. The authors also speculate that the advantage of QCNN lies in the ability to exploit entanglement, which is a global effect, whereas CNN can only capture local correlations. Future research could focus on constructing a multi-class classifier and optimizing the data encoding.

3.4.4 Evaluation

The structure of the QCNN proposed in [4] is different compared to the existing structures of QCNN, as Tak et al. [4] implements the entire network as a parameterized circuit. While the experiments were conducted thoroughly, the presentation of the results could have been better. The authors claim that the proposed QCNN model is expected to be suitable for NISQ devices (prone to noise), but also inform that this needs to be verified via real-world experiments or performing noisy simulations.

3.5 The Dilemma of Quantum Neural Networks

3.5.1 Introduction

Yang et al. [5] compare the performance of Quantum Neural Networks (QNN) with classical Deep Neural Networks (DNN) on a larger scale than previous research to measure the trainability and generalization ability of QNNs under both a Noisy Intermediate Scale Quantum scenario (NISQ) and a Noiseless (N) scenario. They demonstrate how the noise of QNNs negatively impacts their trainability and generalization ability, discuss solutions to it, and finally include a benchmark that evaluates the performance of QNNs and DNNs. The paper introduces three types of QNNs, which differ in terms of the structure of their circuits:

1. Quantum Naive Neural Network (QNNN)
2. Quantum Embedding Neural Network (QENN)
3. Quantum Convolutional Neural Network (QCNN)

3.5.2 Experimental Setup

Yang et al. [5] compare the generalization ability of different types of QNN specified with Convolutional Neural Networks (CNN) and Multi-Layer Perceptron (MLP), with the number of trainable parameters equal to that of a QNN on two datasets: MNIST and WINE. MNIST, Wine, and Quantum synthetic datasets were used to assess the performance and trainability of the models. The effect of regularization by different optimization algorithms such as gradient descent optimizer, stochastic gradient descent optimizer, stochastic quantum natural gradient descent optimizer, and weight decay on the quantum models is studied. Additional experiments were conducted to study how the batch size affects the learning performance of QNNs and QENNs. The experiments were conducted ten times to suppress randomness, and the respective results were reported.

3.5.3 Conclusion

The results show that the current performance of QNNs is not better than current DNNs on real-world datasets, the effective model complexity of QNNs is much simpler than DNNs, failing to cover complicated target concepts, and the imperfection of NISQ machines further worsens the performance of QNNs. Quantum system noise suppresses the learnability of QNNs. Yang et al. [5] conclude that by addressing issues like the barren plateaus phenomenon, and quantum noise and introducing error mitigation techniques, QNNs can be improved further.

3.5.4 Evaluation

The authors introduced a benchmark to assess the learnability of various QNNs fairly and comprehensively. Ablation experiments were conducted to study the effect of batch size and the effect of regularization by different optimization algorithms. The experiments were conducted on the simulator, and noise simulation was also performed via a noisy module extracted from a quantum computer. The results are performed over ten times to suppress the effect of randomness.

4 Discussion

Almost all the papers that compared the Quantum Neural Networks (QNN) with their classical counterparts provided the architectures of the respective quantum networks used. However, most did not explain why specific classical neural networks were chosen. Some implied that they have chosen classical neural networks with the same number of trainable parameters as the quantum neural networks to enable fair comparison.

Most of the comparisons in terms of quantum neural networks were done on the simulators and not on real quantum computers due to hardware limitations of access to quantum computers. Almost all the experiments performed on the simulators were conducted several times, and the averages of the results are reported. QNN and QCNN models showed less standard deviation in terms of training loss and accuracy than their classical counterparts.

Ablation experiments are conducted to study the effect of the batch size and optimization algorithms on the training accuracy of different QNNs proposed. The comparison was made mostly with respect to binary classification. Access to quantum computers can help researchers understand how QNNs work in real-time and also help them study effects like quantum noise and barren plateaus.

The validation of performance on multiple and larger datasets and designing over-parameterized QNNs without the effect of barren plateau could be future research areas. Furthermore, more complicated QNNs can be constructed to toggle more specific datasets. Finally, we conclude that the development of the quantum neural networks is in their early stages.

References

[1] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, 2021.

[2] J. Bausch. Recurrent quantum neural networks. In *Advances in Neural Information Processing Systems*, volume 33, pages 1368–1379, 2020.

[3] Yudong Cao, Gian Giacomo Guerreschi, and Alán Aspuru-Guzik. Quantum neuron: an elementary building block for machine learning on quantum computers. 2017.

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

[5] Y. Qian, X. Wang, Y. Du, X. Wu, and D. Tao. The dilemma of quantum neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[6] M. Schuld, I. Sinayskiy, and F. Petruccione. The quest for a quantum neural network. *Quantum Information Processing*, 13:2567–2586, 2014.

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