A Comparative Analysis of Quantum Neural Nets on Imbalance Dataset

Master of Science in Computer Science, Dec 2024

University of North Georgia

Approval Page

Abstract

Traditional approaches to training neural networks often emphasize the necessity of relatively balanced datasets to ensure accurate and reliable outcomes. Despite the importance of balanced data, real-world scenarios are filled with imbalanced datasets. Many domains, from medical diagnostics to fraud detection, naturally produce datasets where certain classes are underrepresented. This imbalance can lead to biased models that perform well in the majority class while failing to accurately predict the minority class. The conventional methods of addressing this imbalance, such as data resampling and synthetic data generation, have their limitations and often do not fully resolve the underlying issues. This research aims to explore an approach to overcoming the challenges posed by imbalanced data in Deep Neural Networks (DNNs) by integrating quantum mechanics. Quantum computing, with its potential for computational speed-ups and unique data handling capabilities, could offer new pathways to address the limitations of classical neural network training. The primary research question guiding this thesis is: Does Quantum Neural Network (QNN) improve accuracy when they are applied to imbalanced data? By investigating this question, this research will contribute to the potential benefits of integrating quantum computing to improve the performance of DNNs. . Overall, this research will examine the intersection of quantum mechanics and deep learning to overcome the limitations imposed by imbalanced data.

Copyrights page

Table of Contents

[Introduction 8](#_Toc736032307)

[Neural Network and its Dependency on Data Quality 8](#_Toc564684405)

[Limitation of DNNs motivate QNNs. 8](#_Toc1280778366)

[Potential Contributions to the Fields of AI and Quantum Computing 9](#_Toc530259353)

[Breakdown of the Analysis 9](#_Toc2073012313)

[Literature Reviews 10](#_Toc108434546)

[Importance of Balanced Datasets for Achieving Accurate and Reliable Outcomes in Neural Network Models 10](#_Toc1913807765)

[Impact of Underrepresented Classes on Model Performance and Accuracy 11](#_Toc220764383)

[Quantum Machine Learning on imbalanced Data 11](#_Toc1821194947)

[Quantum Neural Networks (QNNs) 11](#_Toc52948049)

[Current Challenges and Potential of Quantum Neural Networks (QNNs) 12](#_Toc1460817625)

[Methodology 12](#_Toc1103284558)

[Research Methods and Plan of Implementation 13](#_Toc924348578)

[Model Development and Theory 15](#_Toc1992678407)

[Hypothesis 15](#_Toc273336252)

[Theoretical Underpinnings 15](#_Toc357976457)

[Conceptual Research Model 15](#_Toc1681721443)

[IMPLEMENTATION 16](#_Toc821661005)

[Environment setup 17](#_Toc1846291462)

[Dataset Description 17](#_Toc818525804)

[Nomenclature 17](#_Toc127611936)

[Differences from Classical NNs 18](#_Toc788207596)

[Feature Map and Ansatz 19](#_Toc363100150)

[Conceptual Order 19](#_Toc2013127268)

[Results and Analysis 20](#_Toc1124199036)

[Text processing 21](#_Toc1752998367)

[Data preprocessing 21](#_Toc1425353572)

[Results 23](#_Toc316175214)

[Image Processing 24](#_Toc2027115519)

[Data Preprocessing 24](#_Toc474819499)

[Discussions 28](#_Toc877485973)

[Limited Qubits 29](#_Toc1868214131)

[Feature Map and Ansatz 29](#_Toc1781154153)

[Future Directions 29](#_Toc913864530)

[Hybrid Computing 29](#_Toc1796381091)

[Current Limitations 29](#_Toc1643553381)

[Conclusion 30](#_Toc294009851)

[References 30](#_Toc1537608540)

List of Diagrams

Figure 1 Research roadmap.....................................................................9

Figure 2 Implementation Map .................................................................13

Figure 3 Research Model Flow chart .........................................................15

Figure 4 QNN Structure ..........................................................................21

Figure 5 Feature Map .............................................................................22

Figure 6 Ansatz Circute ............................................................................23

Figure 7 Traning Grpahs text .....................................................................24

Figure 8 QCNN 8qbits Circute diagram.......................................................27

Figure 9 Image Traning graph ....................................................................27

List of Tables

Table 1 Implementation Map ......................................................................13

Table 2 QNN vs NN text processing .............................................................25

Table 3 QCNN vs CNN balanced data ..........................................................27

Table 4 QCNN on imbalanced data ..............................................................28

Table 5 Filling accuracy ..............................................................................29

# Introduction

### Neural Network and its Dependency on Data Quality

Neural networks, particularly Deep Neural Networks (DNNs), have revolutionized machine learning by providing robust solutions for complex tasks such as image, text, and audio classification. The success of these models largely depends on the quality and quantity of the data used for training. A well-curated dataset that adequately represents all classes is crucial for the development of accurate and reliable models. However, real-world datasets often suffer from imbalances, where some classes are significantly underrepresented compared to others. This imbalance can severely affect the performance of neural networks, leading to biased predictions and reduced generalization capabilities (Goswami, S. 2020, January 20. Impact of Data Quality on Deep Neural Network Training)

### Limitation of DNNs motivate QNNs.

The limitations of DNNs in handling imbalanced datasets motivate the exploration of alternative models like Quantum Neural Networks (QNNs).( . Beer, K. 2022 May 17. Quantum neural networks)

QNNs combine the power of neural networks with the principles of quantum mechanics. ( M. A., & Chuang, I. L. (2000). Quantum computation and quantum information)

They utilize quantum bits (qubits) that can exist in multiple states simultaneously (superposition), potentially offering advantages in dealing with imbalanced data.

### Potential Contributions to the Fields of AI and Quantum Computing

This research has the potential to make significant contributions to both AI and quantum computing by demonstrating the effectiveness of QNNs in addressing the challenges of imbalanced datasets. Successful implementation and validation of QNNs could pave the way for their adoption in various applications, enhancing the accuracy and fairness of AI systems. Additionally, this study could provide valuable insights into the practical benefits of quantum computing in real-world machine learning tasks.

### Breakdown of the Analysis

Based on Figure 1, my research roadmap will follow these steps

Step 1 Literature Review - Identify existing work on handling imbalanced data in Deep Neural Networks (DNNs).

Step 2: Methodology - Research Methods and Plan of Implementation, Including Model Development Theory, Hypothesis, and Research Question

Step 3 Implementation - Define the DNN and Quantum Neural Network (QNN) models I will use, including data preprocessing techniques and training procedures, Implement the models using appropriate frameworks and data processing methods.

Step 4 Results and Analysis - Analyze the performance of both models on imbalanced datasets.

Step 5 Discussion - Interpret the results, highlighting the advantages and limitations of QNNs compared to DNNs.

Step 6 Conclusion - Summarize the findings and discuss the implications for future research.

**Fig 1 Research roadmap**

# Literature Reviews

### Importance of Balanced Datasets for Achieving Accurate and Reliable Outcomes in Neural Network Models

Balanced datasets are essential for neural network training as they ensure that the model learns to recognize and predict all classes accurately. When a dataset is balanced, each class contributes equally to the learning process, allowing the model to develop a comprehensive understanding of the problem domain. In contrast, imbalanced datasets can cause the model to become biased towards the majority class, ignoring or misclassifying the minority classes. This imbalance not only diminishes the accuracy of the model but also undermines its reliability and fairness, particularly in applications where accurate classification of minority classes is critical, such as medical diagnosis or fraud detection. (Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer.,2002, "SMOTE: synthetic minority over-sampling technique.")

### Impact of Underrepresented Classes on Model Performance and Accuracy

Underrepresented classes in imbalanced datasets pose a significant challenge for neural network models. The disparity in class representation can lead to several issues, including:

* **Biased Learning:** The model may become biased towards the majority class, resulting in poor performance on the minority classes.
* **Misclassification:** Underrepresented classes are more likely to be misclassified, as the model does not have sufficient examples to learn their characteristics effectively.
* **Overfitting:** The model may be overfit to the majority class, failing to generalize well to new, unseen data. ( Blohmke, S., & Fetahu, B. (2021, January). Fairness in Machine Learning: A Survey.)

### Quantum Machine Learning on imbalanced Data

According to Schetakis et al. (2023), research on the FH model, a Quantum Machine Learning (QML) architecture, suggests robustness to data noise. This noise tolerance could potentially contribute to effective classification even in scenarios with imbalanced data sets.

### Quantum Neural Networks (QNNs)

In the paper "Randomness-Enhanced Expressivity of Quantum Neural Networks" by Liu et al. (2023), the authors explore how incorporating randomness into the parameters of Quantum Neural Networks (QNNs) can significantly enhance their expressivity. Expressivity, in this context, refers to the ability of a QNN to approximate a wide variety of functions, crucial for solving complex computational problems. The study demonstrates through theoretical analysis and numerical simulations that QNNs with random initial parameters can produce a greater diversity of quantum states compared to deterministic ones. This increased diversity allows the QNNs to capture more intricate patterns within the data, leading to more robust and efficient quantum algorithms. The findings suggest that leveraging randomness in QNNs can help overcome limitations faced by classical neural networks, such as avoiding local minima and reducing the need for extensive training data, thereby advancing the field of quantum computing.

### Current Challenges and Potential of Quantum Neural Networks (QNNs)

While DNNs have achieved remarkable success, they struggle with imbalanced datasets, often exhibiting bias towards the majority class. This highlights the need for alternative models that can handle imbalanced data more effectively.

However, a critical approach is necessary when evaluating claims about QNNs outperforming DNNs. A recent study ([Arute, Marc, et al. ,2023,"Better than classical?, The subtle art of benchmarking quantum machine learning models.") analyzing research papers on quantum machine learning found that:

* Many claims of outperformance are based on small-scale benchmarks and may not translate to real-world scenarios.
* Some reported improvements focus on specific aspects of QNNs (e.g., optimization algorithms) rather than overall model performance.
* Very few studies have demonstrably shown QNNs outperforming classical models, often citing limitations in current hardware as a reason for the gap.

**This highlights the importance of careful evaluation when comparing DNNs and QNNs.** While QNNs hold promise for overcoming challenges faced by DNNs with imbalanced data, further research and advancements in quantum hardware are needed to fully assess their potential in this domain.

This revised approach acknowledges the limitations of current QNN research and emphasizes the need for a more critical evaluation before asserting superiority over DNNs. It also paves the way for your research on exploring QNNs for imbalanced data classification.

# Methodology

### Research Methods and Plan of Implementation

To investigate the effectiveness of Quantum Neural Networks (QNNs) in handling imbalanced data, this research will employ text, image and audio classification tasks. Classification using quantum frameworks like PennyLane with TensorFlow Quantum (TFQ) or Qiskit. The research will involve data preprocessing tailored for quantum deep learning, followed by the development and testing of quantum neural network models. The primary research question guiding this investigation is whether Quantum Deep Neural Networks can effectively address the challenges of imbalanced data. The hypothesis posits that QNNs have the potential to perform well on imbalanced datasets, thereby overcoming the limitations of traditional neural network approaches.

According to Figure 2, I will first perform a classification task on balanced datasets for text, image, and audio data using both Deep Neural Networks (DNNs) and Quantum Neural Networks (QNNs). Since DNNs can struggle with imbalanced data, I will then evaluate the performance of QNNs on imbalanced datasets with two split ratios: 80% majority class and 20% minority class, and 90% majority class and 10% minority class. In all these experiments, the focus will be on classifying images, text, and audio data, with accuracy as the primary dependent variable.

The table 1 “Accuracy Table " summarizes the planned evaluation of classification accuracy for Deep Neural Networks (DNNs) and Quantum Neural Networks (QNNs) across text, image, and audio data. It compares performance on both balanced datasets (equal class representation) and imbalanced datasets (uneven class distribution) with two split ratios (80%/20% and 90%/10%). Accuracy, represented by question marks ("?"), will be measured for each data type and model combination to determine which performs better under balanced and imbalanced conditions.

A diagram of data processing

Description automatically generated

**Fig 2 : Implementation Map**

A black grid with white lines

Description automatically generated **Tab1: Accuracy Table**

### Model Development and Theory

In this section, a theoretical framework for the development and evaluation of Quantum Neural Networks (QNNs) for handling imbalanced data. The proposed framework will encompass hypotheses, theoretical underpinnings, and a conceptual research model to guide our investigation.

### Hypothesis

Our hypothesis posits that Quantum Deep Neural Networks (QNNs) can effectively address the challenges posed by imbalanced data, resulting in improved accuracy and performance compared to traditional neural network approaches.

### Theoretical Underpinnings

The theoretical foundation of our research lies in the principles of quantum mechanics and their potential application in deep learning. Quantum computing offers unique advantages, such as superposition, entanglement, and interference, which can potentially enhance the processing power and data handling capabilities of neural networks. By harnessing these quantum properties, QNNs may exhibit superior performance in handling imbalanced datasets, mitigating biases, and improving predictive accuracy.

### Conceptual Research Model

Our conceptual research model outlines the key components and relationships underlying the development and evaluation of QNNs for imbalanced data.

This flowchart explains using Quantum Neural Networks (QNNs) for real-world imbalanced data classification. We design and train QNNs on the data, then evaluate and compare them to traditional models. Finally, we assess their generalizability across datasets.

**Fig 3: Research Model Flow Chart**

By elucidating these components and their interrelationships, our conceptual research model serves as a guiding framework for the empirical investigation into the efficacy of QNNs for imbalanced data classification. Through empirical validation and analysis, we aim to contribute novel insights into the intersection of quantum computing and deep learning, advancing the frontier of AI research.

# IMPLEMENTATION

## Environment setup

Establishing a robust quantum machine learning environment proved challenging. Initial attempts with TensorFlow Quantum (TFQ) were unsuccessful in multiclass image classification due to complexities with qubits and quantum circuits. This necessitated a more flexible framework. Consequently, we opted for a combination of Qiskit and PyTorch. While configuring Qiskit 1.1.1 and its dependencies was intricate, the detailed setup process is documented in our GitHub repository for reference. Given the substantial complexities encountered, a decision was made to limit the scope of this research to image and text processing using Quantum Neural Networks (QNNs), excluding audio processing.

All implementation details and environment setup are present in my GitHub repo:https://github.com/Ironman20121/Quantum-on-imbalanced-data-.git

## Dataset Description

Datasets used in this research include:

* **Image Data:** Multi-class fruit dataset from Kaggle (<https://www.kaggle.com/code/aravindanr22052001/tutorial-3-faster-rcnn/input>).
* **Text Data:** SMS spam vs. ham classification dataset from Kaggle (<https://www.kaggle.com/code/mykeysid10/sms-spam-ham-classification-using-nlp/input>).

## Nomenclature

#### **Qubits**

**Definition:** A qubit is the basic unit of quantum information, analogous to a classical bit. However, unlike a classical bit which can be in a state 0 or 1, a qubit can be in a superposition of both states.

**Superposition:** This property allows a qubit to be in a state *∣ψ⟩=α∣0⟩+β∣1⟩|\psi\rangle = \alpha|0\rangle + \beta|1\rangle*∣ψ⟩=α∣0⟩+β∣1⟩, where *α\alpha*α and *β\beta*β are complex numbers such that *∣α∣2+∣β∣2=1|\alpha|^2 + |\beta|^2 = 1*∣α∣2+∣β∣2=1.

**Entanglement:** Qubits can also be entangled, meaning the state of one qubit can depend on the state of another, no matter the distance between them.

#### **Quantum Circuits**

**Definition:** A quantum circuit is a computational routine consisting of coherent quantum operations on quantum data, such as qubits, performed using quantum gates.

**Quantum Gates:** Analogous to classical logic gates, quantum gates manipulate qubits. Examples include the Hadamard gate, Pauli-X gate, and CNOT gate. These gates can create superposition and entanglement.

**Circuit Design:** A typical quantum circuit starts with qubits in a known state, applies a sequence of quantum gates, and ends with a measurement of the qubits.

#### **Quantum Neural Networks (QNNs)**

### Differences from Classical NNs

* **Data Representation:** Classical NNs operate on classical data, while QNNs operate on quantum data (qubits).
* **Parallelism:** QNNs leverage quantum superposition and entanglement, potentially allowing them to process multiple states simultaneously, which could offer computational advantages over classical NNs.
* **Training and Inference:** QNNs require quantum measurements for output, and training involves adjusting quantum gates (parameterized gates) to minimize a loss function, like classical NNs but often more complex due to the nature of quantum mechanics.

#### Sample-based QNN and Estimator-based QNN

* **Sample-based QNN:** Utilizes samples from a quantum circuit to estimate the output. This approach involves running the quantum circuit multiple times and averaging the results, which can introduce statistical noise.
* **Estimator-based QNN:** Uses estimators to predict the output without needing as many samples, reducing noise and potentially increasing efficiency.

### Feature Map and Ansatz

* **Feature Map:** A quantum feature map encodes classical data into a quantum state. It uses a set of quantum gates parameterized by the input data to transform the initial state of the qubits. This step is crucial for leveraging the power of quantum computing for machine learning tasks.

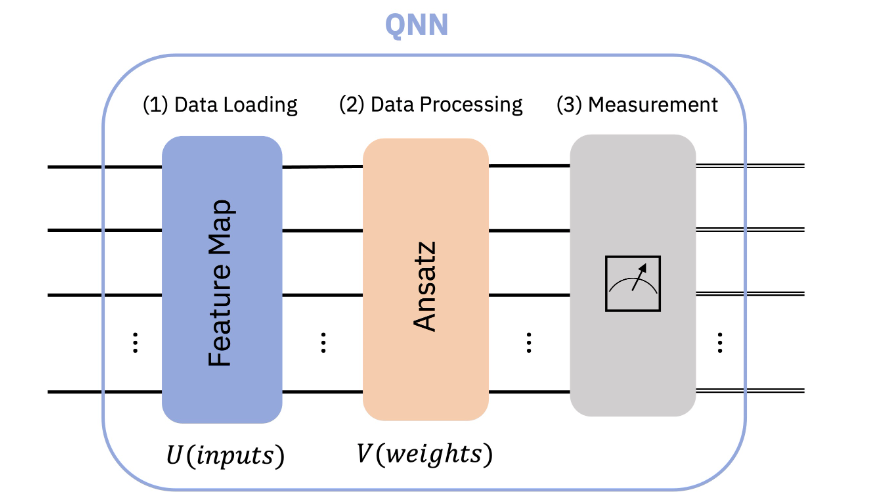
**Example:** For input *xx*x, a feature map might encode *xx*x into a quantum state using a series of rotations and entangling gates.

* **Ansatz:** An ansatz is a parameterized quantum circuit used to approximate the desired output state. The parameters (trainable weights) are adjusted during the training process to minimize the loss function.

**Example:** An ansatz might consist of several layers of quantum gates with parameters that are updated using classical optimization techniques.

### Conceptual Order

1. **Qubits Initialization:** Start with qubits in a known state (usually *∣0⟩|0\rangle*∣0⟩).
2. **Feature Map:** Encode the input data into the quantum state using a parameterized feature map.
3. **Ansatz Application:** Apply the ansatz to the quantum state, which consists of trainable quantum gates.
4. **Measurement:** Measure the qubits to obtain the output.
5. **Loss Calculation:** Calculate the loss based on the difference between the predicted and actual outputs.
6. **Parameter Update:** Update the parameters of the ansatz using classical optimization algorithms.
7. **Iteration:** Repeat the process until the loss converges to an acceptable value. (Qiskit Community. (n.d.). Qiskit machine learning tutorial: Neural networks. Retrieved from <https://qiskit-community.github.io/qiskit-machine-learning/tutorials/01_neural_networks.html#1.2.-Implementation-in-qiskit-machine-learning)>



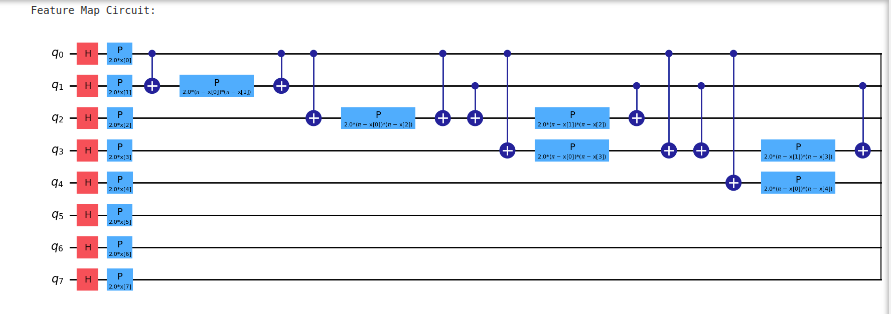
**Fig 4: QNN structure.**

# Results and Analysis

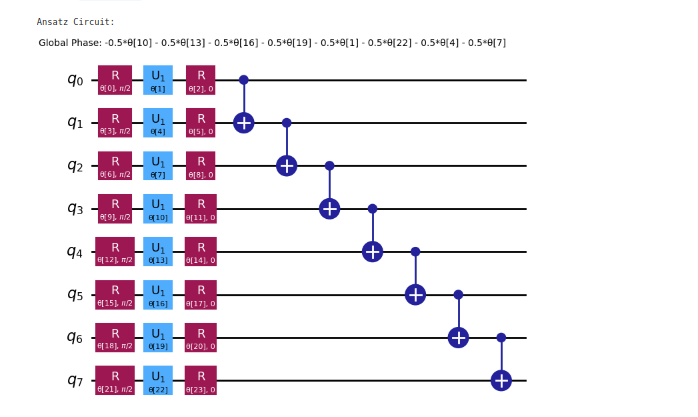
## Text processing

### Data preprocessing

The CSV dataset was initially converted to UTF-8 encoding to ensure compatibility and proper handling of text characters. Subsequently, Label Encoding was applied to transform categorical labels into numerical representations, a necessary step for most machine learning algorithms. The dataset comprised 747 instances labeled as 'spam' and 4825 instances labeled as 'ham'.



**Fig 5: Feature Map**



**Fig 6: Ansatz Circuit**

Builds a NeuralNetworkClassifier using the QNN, COBYLA optimizer, and a callback function for visualization of training.

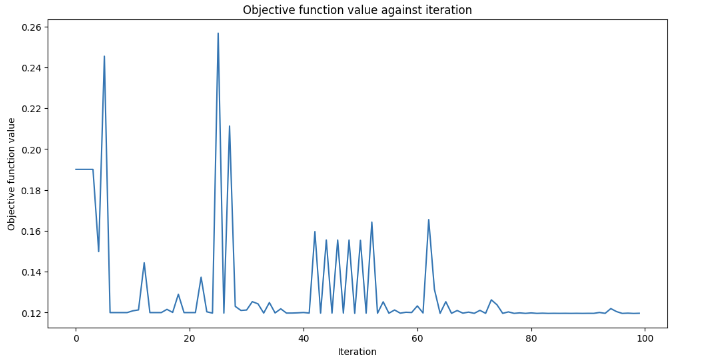


Fig 7 : Training visualization

### Results

The classical deep learning model achieves a high accuracy of 97% and completes its tasks in just 5 minutes, indicating both effective learning and time efficiency. In stark contrast, the quantum deep learning model attains a significantly lower accuracy of 12% and requires 3 hours to complete the same tasks. This disparity highlights the current limitations of quantum deep learning models. The quantum model's lower accuracy could be due to various factors such as the nascent stage of quantum algorithms, the complexity of encoding classical data into quantum format, and current hardware limitations. Additionally, the longer time required for the quantum model might be attributed to the overhead of simulating quantum computations on classical hardware or the inherent complexity of quantum operations. While quantum computing holds promise for the future, this comparison shows that classical deep learning models currently outperform quantum models in both accuracy and efficiency for this specific task, emphasizing the need for further advancements in quantum computing technology and algorithms.

|  |  |  |
| --- | --- | --- |
| Models | Accuracy | Time |
| Classical Deep Neural Nets | 97% | 5 min |
| Quantum Neural Nets 8 qbit | 12% | 3hr |

Table 2 NN vs QNN text processing

## Image Processing

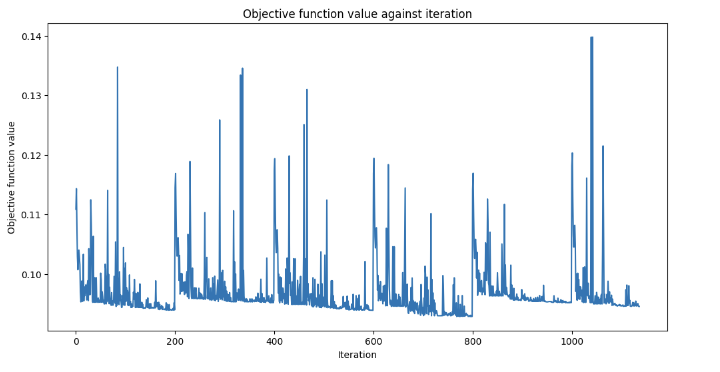
### Data Preprocessing

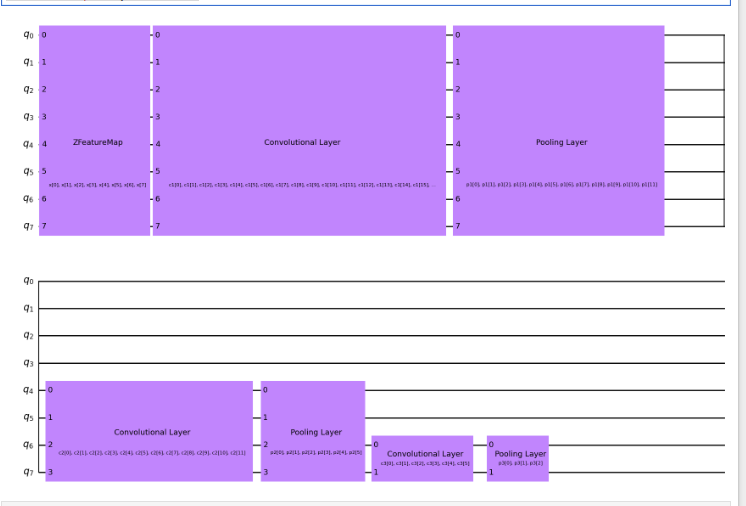
The dataset employed in this study is derived from the Kaggle Multi-Class Fruits dataset, encompassing ten distinct fruit categories: apple, banana, avocado, cherry, kiwi, mango, orange, pineapple, strawberries, and watermelon. The training partition consists of 254 images per fruit category, totaling 2,540 images. Conversely, the testing partition comprises 10 images per category, amounting to 100 images overall. Both traditional Neural Networks (NNs) and Quantum Neural Networks (QNNs) will be evaluated on this balanced dataset.

To further investigate the impact of class imbalance, additional experiments were conducted with varying ratios of apple images to other fruit categories. Specifically, two imbalance ratios were tested: 90:10 and 80:20. In both cases, the apple class was overrepresented, comprising 90% and 80% of the dataset, respectively, while the remaining nine fruit categories were equally distributed with 10% and 20% of the total images. These imbalanced datasets were processed using both QNN and NN models to assess the algorithms' performance under skewed class distributions.

Quantum circuits for convolutional and pooling layers, combining them into an ansatz circuit, and transforming classical image data into a quantum format using the ZFeatureMap. The conv\_circuit function defines a two-qubit unitary operation, forming the basic building block for convolutional layers. The conv\_layer function arranges these operations across multiple qubits, applying them in a staggered pattern for effective quantum convolution. The pool\_circuit function similarly defines a two-qubit unitary operation for pooling, while the pool\_layer function applies these operations to reduce the dimensionality of the qubit space, akin to classical pooling layers.

To transform classical image data into a quantum format, the ZFeatureMap is used. This quantum circuit encodes classical data into the amplitudes of a quantum state by applying parameterized gates based on the input features. In this code, the ZFeatureMap(8) maps classical data into an 8-qubit quantum state, preparing it for further quantum processing. The ansatz circuit, which combines the convolutional and pooling layers, is applied to the quantum state produced by the ZFeatureMap. This ansatz circuit is designed to extract and process the features encoded in the quantum state, similar to how classical neural networks process data. Finally, the entire quantum circuit, including the feature map and ansatz, is used to create an EstimatorQNN, which serves as the quantum neural network for the classifier. The observable is defined to measure the output, allowing the QNN to be trained and evaluated.

Fig 8: training graphs for balanced image classification



**Fig 89: QCNN 8qbits circuit diagram**

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy** | **Time** |
| **CNN** | **87%** | **3hr** |
| **QCNN** | **40%** | **More than 1 day** |

**Table 3: QCNN vs CNN balanced data**

In this comparison, the CNN model achieves an accuracy of 87%, completing its tasks in 3 hours. This indicates that CNNs are highly effective and efficient for balanced datasets, providing a good balance between accuracy and computational time. On the other hand, the QCNN model achieves a significantly lower accuracy of 40% and requires more than one day for training and evaluation. This substantial difference highlights the current challenges in applying quantum convolutional models to practical machine learning tasks. The longer time required for QCNN is due to the complexity of quantum computations and the overhead of simulating quantum operations on classical hardware. Furthermore, the development of QCNNs is in its early stages, and there are no existing multi-classification codes or papers using QCNNs, making this research both pioneering and challenging.

|  |  |  |
| --- | --- | --- |
| **Imbalanced ratio** | **Accuracy** | **Time** |
| **80 :20** | **12%** | **More than 1 Day** |
| **90:10** | **10%** | **More than 1 Day** |

**Table 4: QCNN Imbalanced data**

When applied to imbalanced datasets, the QCNN model performs poorly, as shown in Table 4. For an imbalance ratio of 80/20, the model achieves an accuracy of only 12%, and for a more extreme imbalance ratio of 90/10, the accuracy further drops to 10%. In both cases, the training and evaluation time remains more than one day. These results underscore the difficulty of training QCNNs on imbalanced data, which further exacerbates the already challenging task of quantum neural network training. The inherent inefficiency and lower accuracy on imbalanced datasets suggest that QCNNs require significant advancements in both algorithmic development and quantum hardware to become viable alternatives to classical models.

## Discussions

The performance of QNNs in both text and image classification tasks did not outperform classical DNNs. The accuracy achieved was relatively low, suggesting that further optimization and research are required to fully leverage the potential of QNNs.

Due to the limitations of processing audio data with only 8 qubits, this modality was excluded from the experiments.

Focusing on text classification, specifically spam detection (ham vs. spam), the quantum model achieved a significantly lower accuracy of 12% compared to the classical model's performance. This imbalanced dataset, with a substantial disparity between ham and spam instances, likely contributed to the quantum model's difficulties.

Surprisingly, when applied to image classification with a dataset heavily biased towards apple images, the quantum model unexpectedly demonstrated high accuracy in predicting apple instances but struggled to recognize other fruit types. This behavior deviates from the typical bias observed in classical machine learning models, where an imbalanced dataset often leads to a classifier predominantly predicting the majority class.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data type | DNN | QNN Balanced | QNN 80/20 | QNN 90/10 |
| Text | 97% | NA | 12% | |
| Image | 87% | 40% | 12% | 10% |

**Tab5: Filled Accuracy Table**

### Limited Qubits

The use of only 8 qubits seems insufficient for the complexity of the problem at hand. Increasing the number of qubits might improve performance by allowing for more complex quantum states and computations.

### Feature Map and Ansatz

The current feature map and ansatz might not be effectively processing and converting classical data to quantum data. Further optimization of these components is essential to improve the quantum data representation and enhance the QNN's performance.

### Future Directions

Exploring additional methods for optimizing the quantum data conversion process is crucial. This could involve refining the feature map and ansatz or developing new techniques for better quantum data encoding.

### Hybrid Computing

Direct implementation on quantum hardware requires more expertise and may not be feasible with the current state of quantum technology. Leveraging hybrid computing, where classical and quantum computations are combined, could be a more practical approach. This involves using classical computers for some parts of the computation and quantum processors for others, potentially leading to better performance.

### Current Limitations

At present, the limitations of quantum hardware and algorithms present significant challenges. While a fully quantum approach might be a dead-end for now, hybrid computing represents a promising direction for future research and development.

## Conclusion

This research explored the application of Quantum Neural Networks (QNNs) to handle imbalanced datasets. While QNNs have the potential to address some limitations of classical DNNs, the experimental results showed that their performance on the chosen tasks was not superior.

# References

1. Goswami, S. (2020, January 20). Impact of Data Quality on Deep Neural Network Training. https://arxiv.org/abs/2002.03732
2. Beer, K. (2022, May 17). Quantum Neural Networks. https://arxiv.org/abs/2205.08154
3. Nielsen, M. A., & Chuang, I. L. (2000). Quantum computation and quantum information. Cambridge, UK: Cambridge University Press.
4. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357. https://arxiv.org/pdf/1106.1813
5. Blohmke, S., & Fetahu, B. (2021, January). Fairness in machine learning: A survey. https://arxiv.org/abs/2010.04053
6. Rehák, M., Semančíková, J., Fiedler, J., & Železný, J. (2020). Impact of Data Quality on Machine Learning Algorithms. https://arxiv.org/abs/2002.03732
7. Beer, K., Bondarenko, D., Farrelly, T., Osborne, T. J., Salzmann, R., & Wolf, R. (2022). Training deep quantum neural networks. https://arxiv.org/abs/2205.08154
8. Schetakis, N., Boguslavsky, M., & Griffin, P. (2023). Quantum Machine Learning for Imbalanced Data Classification. https://arxiv.org/abs/2308.10847
9. Liu, Y., Zhang, X., & Zhang, J. (2023). Randomness-Enhanced Expressivity of Quantum Neural Networks. https://arxiv.org/abs/2308.04740
10. Arute, M., Babbush, R., Bardin, J. C., Barends, R., Boixo, S., Broughton, M., ... & Roushan, P. (2024). Better than classical? The subtle art of benchmarking quantum machine learning models. https://arxiv.org/abs/2403.07059
11. Schetakis, N., Aghamalyan, D., Boguslavsky, M., Rees, A., Raktomalala, M., & Griffin, P. (2023, August 7). Quantum Machine Learning for Credit Card Score. <https://arxiv.org/abs/2308.03575>
12. Qiskit Community. (n.d.). Qiskit machine learning tutorial: Neural networks. Retrieved from <https://qiskit-community.github.io/qiskit-machine-learning/tutorials/01_neural_networks.html#1.2.-Implementation-in-qiskit-machine-learning>