JCBI
Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 04 Issue 02
2023

Research Article
https://doi.org/10.56979/402/2023

Hybrid Quantum Neural Network Approach for Rapid Response to Cyber Attacks

Arshad Iqbal[1], Muhammad Masoom Alam[1], Nadeem Javaid[1], Sadia Nishat Kazmi[2,*], Fraz Ahmad[1], Allay Hyder Urooj[3]

[1]Department of Computer Science, COMSATS University Islamabad, Islamabad 45550, Pakistan.
[2]Faculty of Automatic Control, Electronics, and Computer Science, Silesian University of Technology, Gliwice, 44-100, Poland.
[3]Directorate of Outreach, University of Agriculture, Faisalabad 38000, Pakistan.
[*]Corresponding Author: Sadia Nishat Kazmi. Email: contactkazmi4@gmail.com

Received: December 21, 2022 **Accepted:** February 20, 2023 **Published:** March 29, 2023.

Abstract: Cyber attacks on data servers and critical infrastructure of fifth-generation (5G) wireless communication networks are increasing day by day. Therefore, an intelligent, reliable and cost effictive security of sixth-generation (6G) wireless communication network is inevitable. Quantum technology is one of the enabling technologies of future computing, as well as of the 6G. In the event of a cyberattack, rapid response is critical to minimize the risk of data loss and denial-of-service (DoS) attacks. Quantum computers are a new generation of computers that are expected to be much faster than classical computers. In this study, we have proposed and implemented a hybrid quantum neural network (HQNN) model. The proposed HQNN model was trained and tested on a dataset from the Australian center for cybersecurity. Simulation results show that the proposed HQNN model is faster in training as well as in the testing phase compared to classical neural networks. Moreover, our proposed model helps to overcome the underutilization of centeral processing unit (CPU) resources. CPU utilization of our proposed HQNN model is 95-100%, while that of the classical model is only 35-75% during the training and testing of the dataset.

Keywords: Cyber security, Sixth generation (6G) wireless communication network, Quantum computing, Cyber-attack, Big data, Resource utilization.

1. Introduction

To ensure low latency and the highest reliability, fifth-generation (5G) wireless communication networks are being standardized. Massive data generation is an inherent feature of 5G communication networks, and 5G networks will not be able to meet all the requirements of future networks. Therefore, the need for a new communication network has arisen and the research community is working on 6G communication network. It is expected that the sixth generation (6G) communication network will provide improved cost efficiency, global coverage, improved energy efficiency, higher intelligence level, enhanced security, and better spectral efficiency. To deliver improved services, 6G will use new technologies and high-speed computing for fast data processing.

Along with benefits, the 6G wireless communication network is also prone to cyber-attacks. According to IBM [1, 2], the average cost of a single data breach in 2019 was approximately 3.92 million USD which indicates that cyber-attacks are extremely destructive. Consequently, cybersecurity becomes a top priority for businesses. In the event of a cyber-attack, incident response is an approach to cope with and manage information systems.

A quick response to any attack/incident is necessary to minimize the damage by taking appropriate countermeasures. Rapid incident response helps to mitigate the damage and reduce the recovery time and

cost. Machine learning using quantum could lead to a new area of research and the detection of innovative models that revolutionize the field of machine learning (ML). Machine learning and in particular quantum computing ML will gradually permeate all facets of quantum computing by revolutionizing the way we look at quantum computing. Google's quantum computer did a certain number-crunching in 200 seconds that would take the most advanced supercomputer 10,000 years to do [3, 4]. Incident response time depends on big data training, big data testing, and the use of central processing unit (CPU) also plays an important role.

Quantum technology is one of the most advanced computing technologies that has made many breakthroughs in the near past. Quantum technology is rapidly evolving from infancy to a mature field of science and technology. The parallel information processing capability of quantum technology suggests better performance in terms of security, memory requirements, computational speed, etc., than conventional methods.

Richard Feynman proposed in the early 1980s that quantum computers would be able to solve extremely difficult problems in physics and chemistry in the coming era [5]. Feynman's vision opened the doors to many practical and theoretical challenges for the research community. The first challenge was to develop a true quantum computer that could perform computations in a large (Hilbert) space at higher speeds. The second challenge was to formulate problems that can be solved easily by quantum computers, but with great difficulty by classical computers. The superconducting qubit processor [6], a milestone on the road to quantum computing, has proven its superiority by efficiently solving the above two challenges [7, 8].

To overcome the processing and speed limitations of classical computing, quantum computing was proposed, and now quantum computers are ready to be used for high-end scientific computations. Powerful applications and algorithms, especially ML, have been developed to support quantum computing [9, 10]. Quantum computers with sufficient computing power can accelerate the most important ML algorithms, namely: neural networks, Boltzmann machines [12], Bayesian interface, data fitting, support vector machines [14], recommendation systems, Monte Carlo methods [20], and principle component analysis [21].

Deep learning, on the side of classical computation, has neural network-based machine learning techniques [11, 22, 23] and is powered by specialized hardware and new specialized software libraries. However, computational units in digital computers are bit registers. In deep learning, continuous vectors and tensors transformed into a higher dimensional space are used as computational units.

Considering the advantages of neural networks and quantum computing, we proposed and implemented a hybrid quantum neural network (HQNN) in this study. The proposed HQNN model was trained and tested on a dataset from the Australian Center for Cybersecurity (ACCS). Simulation results show that the proposed HQNN is faster in the training phase and the testing phase compared to classical neural networks. Moreover, our proposed model helps to overcome the under-utilization of the CPU resource.

The remainder of this paper is structured as follows. Section 2 provide the Preliminaries and background of neural network and quantum computing. Section 3 presents the proposed system model followed by implementation details of the proposed system model in section 4. The results and discussions of the study are presented in section 5. Eventually, the conclusion and expected future research directions are given in section 6.

2. Preliminaries and Background

2.1 Classical neural network (CNN)

The CNN is composed of a set of nodes arranged in a different number of layers. The actual processing of information in CNN takes place at the nodes. CNN has an input layer, one or more hidden layer(s), and an output layer. The basic architecture of the classical neural network is shown in figure 1. Input data i.e. ACCS dataset is received at the input layer and it reaches to output layer through one or more hidden layer(s). The number of iterations, number of hidden layers, and learning rate are the controlling parameters of CNN. Different forms of neural networks are discussed in [11].

2.2 Quantum technology

Classical computers execute tasks based on physical states i.e. 0 or 1, [15] whereas, in quantum computers, the computation is performed on quantum states known as qu-bits. Quantum states represent the unclear position and undefined properties of an object. Special computational techniques namely; superposition and entanglement are involved in quantum computations, which increase the performance of computers far beyond that of classical computers. Superposition is a concept that allows computers to have simultaneous states of one and zero, whereas classical computers have only one state, i.e. either one or zero [15]. Entanglement, on the other hand, is a phenomenon in which the quantum states of two or more objects must be described in terms of each other. The quantum machine takes input λ and the parameters α. The quantum processing unit (QPU) produces output β using function $f(\lambda, \alpha) = \beta.A$
Description of the Quantum model's parts is given in the following.

1) Initialization module: At this stage, quantum states are created against the description given at the input. Later on, these quantum states are saved into quantum memory.
2) Quantum memory module: This module stores the quantum states received from the initialization module and exchanges the qubits with quantum transistors.
3) Quantum transistors: Quantum transformation takes place in this module and there may be an array of quantum transistors working in parallel.
4) Measurement: This module provides a classical bit out of the qubit being measured through the implementation of a photon detection procedure.

2.3 Quantum neural network

A general quantum neural network (QNN) is constructed as a layered system, where each layer contains a gate from the collective set of gates. This layered approach is based on the analogy that as the number of layers increases, the mean square error (MSE) is lowered. This approach is useful up to a certain number of layers where there is no or very little improvement in MSE with further addition of layers.

2.4 Variational quantum circuit

We have used a variational quantum circuit (VQC), which transfers the concept of a linked layer model from the neural networks of classical computers to the world of quantum neural networks. the variational quantum circuit includes a set of actions based on some parameters common to continuous-variable quantum computations.

3. Proposed system model

The proposed hybrid neural network model is based on quantum computing. Some layers of the classical neural network are supplemented with layers of the quantum neural network to minimize the obstacles of technological limitations. The proposed system model of a hybrid quantum neural network is shown in figure 2, where three basic parts are shown which are; classical encoder, encoding layer, and quantum decoder. The dataset is fed at the classical layer as an input and then the encoding layer transforms the input data into quantum readable format. The next stage of the quantum decoder performs the quantum processing in a way that the output of the quantum state of the first layer is used as the input quantum state to the second layer. The output of the quantum state of the second layer is fed as the input quantum state to the third layer, and this process continues [16]. The different layers consist of different sizes by eliminating qumodes within the layers. The elimination of qumodes can be done by calculating the replacement qumodes. These qumodes are the basic information-carrying units of CV quantum computers.

To overcome the limitations of the classical neural network and to utilize the quantum computing power, we have proposed to transform CNN into a continuous variable quantum neural network by integration of CNN and quantum computer. In our proposed hybrid quantum neural network model, a scenario is developed where the gates in VQC and the variables of the neural network do not form any entanglement or superposition. By accumulating numerous modules of VQC, we can generate multi-layer neural

networks that are very sensitive. Since this neural network is equipped with a universal set of gates, therefore, the model provides a quantum advantage.

A quantum machine can be programmed and its parameters can be changed to alter its specifications. We can set some parameters as input data variables x and associate further constraints as learnable variables θ. The quantum machine eventually produces an output, shown in equation (1), which depends on the input variables. Devices using this method operate on the principle of supervised learning. This learning model is also called a variational classifier.

A general continuous variable QNN is built in a layered structure, where each layer contains a gate from the universal set of gates. Each layer l in continuous variable neural network (CV-QNN) consists of a successive sequence of gates.

$$l(x) := \emptyset_1 \hat{\mu}_2 \hat{D} \hat{S} \hat{\mu}_1 \quad (1)$$

The function computed by a quantum device may be unique to the architecture of its hardware. This means the way variables enter the computation and how one describes the inputs and outputs of the quantum model. If we do not know how to simulate the quantum model [16] with a classical computer, we cannot do machine learning that can be done with a quantum device. Mathematically, we could say that the layer $l: R^n \rightarrow R^m$ for each input $x \epsilon R^n$ constructs a circuit that performs the transformation.

$$l(x) = \varphi(Wx + b) \quad (2)$$

In equation (2), $W \epsilon R^m$ shows the weight matrix, φ is the sequential function, and $b \epsilon R^m$ represents a bias vector.

Evolutionary work on quantum circuits with continuous variables shows how such hardware-derived models can be trained with conventional computers, and researchers around the world are currently calling for the superiority and limitations of these quantum models to be investigated. Quantum computers can be universally implemented, as we have proposed with classical neural networks. To make CV-QNN universally implementable, we will introduce operator versions of the variable $\hat{x}$, as indicated by equation (3).

$$\hat{x} = \int_{-\infty}^{\infty} \langle x|x\rangle dx \quad (3)$$

Developing a new model for machine learning using quantum computers is akin to digging gold out of a mine. In the case of machine learning with quantum computers, we have discovered some promising gilt symbols through Google research. Therefore, we confidently continued to work and finally achieved our goal of better performance. In this work, we aim to develop a new machine-learning model that can be trained on cyber incident data to predict whether a certain event is an attack or not.

4. Implementation details

We used a multi-layer neural network where each layer has a linear transformation and then a variable activation function, as shown in figure 2. For every input vector in this system, there is an output vector. Training data is given as an input to the classical layers of our model, which is then passed to the quantum neural network layer to increase processing efficiency. The first section of the proposed network consists of three classical neural network layers. These classical layers collect features from the ACCS dataset (UNSW-NB15). These classical layers are followed by two invisible layers of the same size and the result is stored on another layer. In this hybrid model, the second part consists of five layers of quantum neural network (QNN). Finally, the system classified an incident as normal or an attack based on a set of variables in our data.

A dataset of UNSW-BN15 is trained on the proposed hybrid quantum neural network model, which is a combination of CNN and QNN layers. The HQNN is based on a continuous variable architecture (CV) that encodes quantum information, as shown in figure 2. Our proposed HQNN model consists of a layered array of uninterrupted parameterized gates, as is common for continuous quantum computing. Owing to the construction of the model with continuous variables, HQNN can deal with extremely discrete variables

while preserving unity. While preserving parallel relations in rotations and discrete activation functions, the main foundations of neural networks in QNN are ratified using non-Gaussian and Gaussian gates.

We used the Pennylane model, which works with the strawberry field and is based on the CV model, as described in [18]. We used this model in our security domain to detect anomalies within the dataset UNSW-NB15, compiled by the Australian center for cyber security (ACCS). Types of attacks in this dataset are given in the following.

1) Analysis of data attack
2) Backdoors attack
3) Denial of service (DoS) attack
4) The exploitation of critical data attack
5) Fuzzers attack
6) Generic attack
7) Reconnaissance attack
8) Shellcode attack
9) Worms attack

ACCS has used Argus, Bro- IDS tools [19]. The implementation of the HQNN model using Pennylane has proved that the proposed HQNN model can be trained faster than the CNN model on the dataset of cyber incidents. We have worked on the proposed model as described in [17] to reduce the training time and improve the response time.

The hybrid neural networks in our research are replicated mathematically using the quantum circuit simulator, Python version 3.7, TensorFlow with Strawberry Fields and Pennylane, and automatic differentiation [24] and numerical algorithms are used to train these networks. Automatic differentiation techniques allow researchers to directly use accepted optimization algorithms based on stochastic gradient descent.

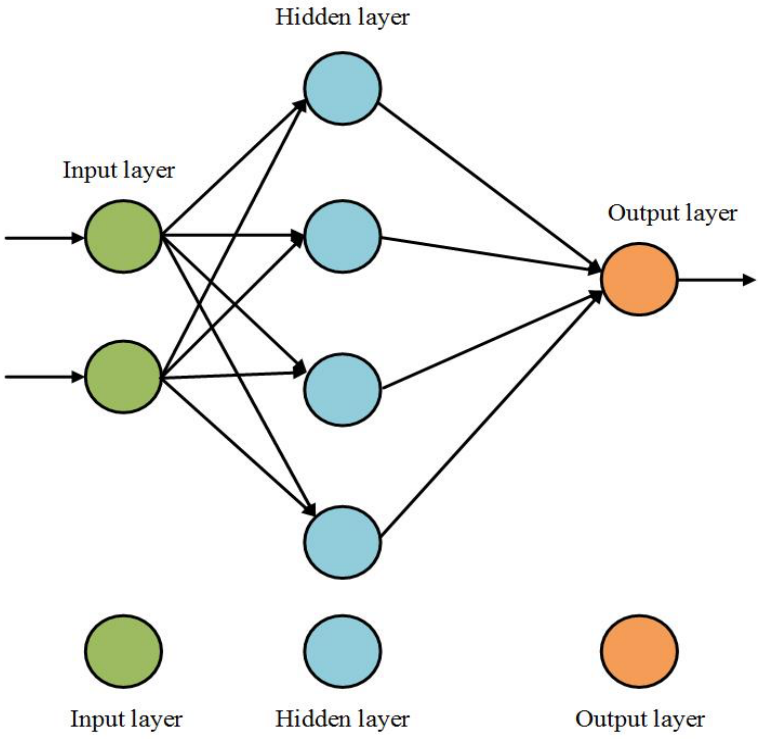

Figure 1. The basic architecture of classical neural network.

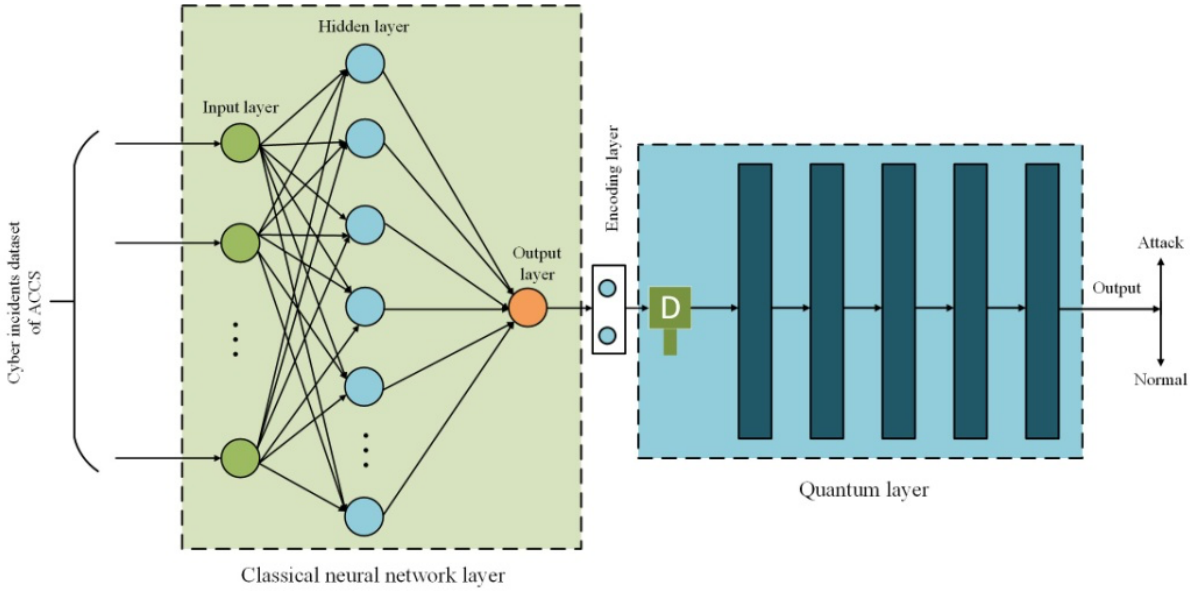

Figure 2. Proposed system model of a hybrid quantum neural network.

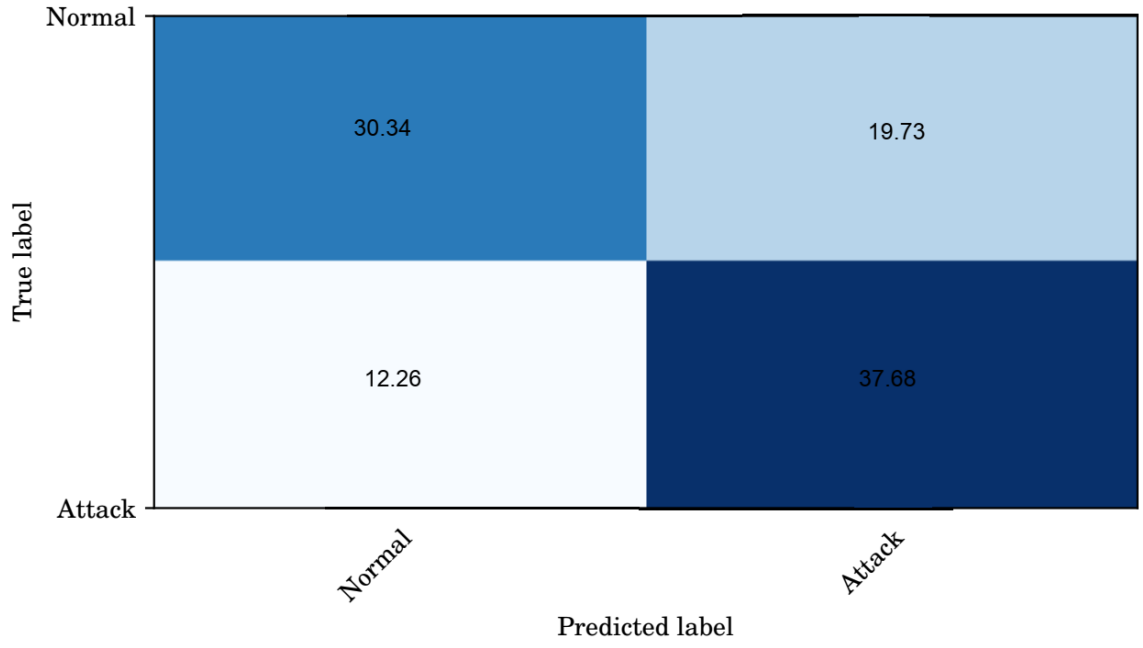

Figure 3: Confusion table of quantum neural networks

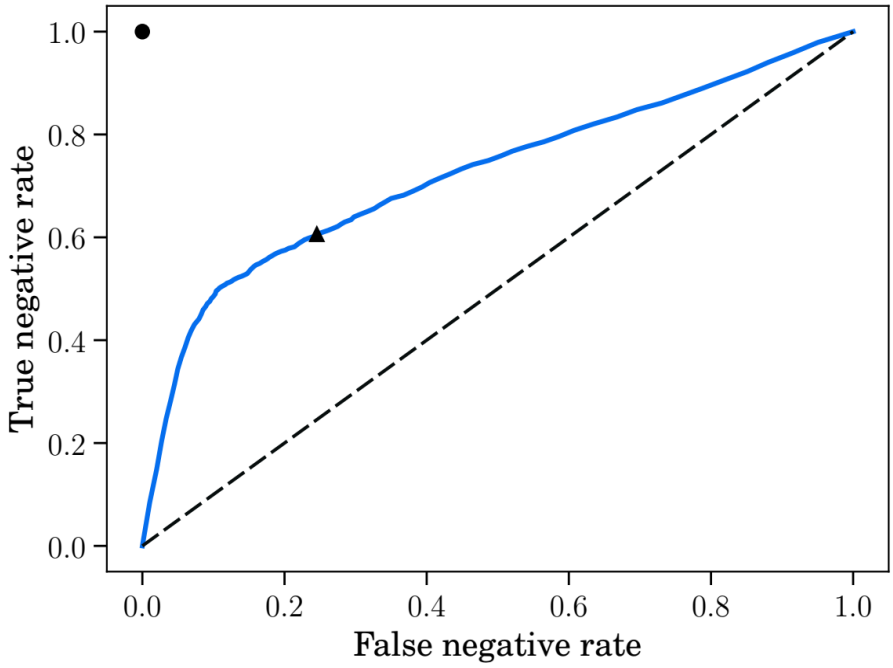

Figure 4: ROC curve of proposed hybrid quantum neural networks

Figure 5: CPU resource utilization graph of (a) classical neural network model and (b) proposed HQNN model

Table 1. Speed and accuracy comparison of CNN model and proposed HQNN model.

Entity of interest	CNN mode	HQNN model
Number of iterations	1000	1000
Size of training data set	115000	115000
Size of testing data set	40500	40500
Batch size	25	25
Testing accuracy	0.90	0.68
Training time	8742.48 s	1965.35 s
Testing time	11.8 s	3.7 s

5. Results and discussion

In the case of big data, predictive accuracy, and efficiency become a challenge and the complexity of training and testing increases manifold. In this study, we have created two classes where 0 represents normal and 1 represents attacks. A list of all attacks can be found in the UNSW-NB15 dataset [19].

Our proposed HQNN model successfully trained a neural network on sample data to predict whether a cyber incident was an attack or not. The proposed hybrid neural network model took only 1965.35 seconds while the classical neural network model took 8742.48 seconds, as shown in Table 1. The training and testing time of our proposed HQNN model is 4 times and 3 times faster than the classical model's training and testing times, respectively which indicates a great breakthrough in the field of threat hunting and incident response. In this research work, we have trained our proposed model on the dataset of ACCS with 115000 training examples and 40500 test examples. For training the data, we used the stochastic gradient descent (SGD) method with 1000 iterations and a batch size of 25.

The testing accuracy of the CNN model was observed 0.90 while it was 0.68 for our proposed HQNN model. The training time of the CNN model was 8742.48 seconds whereas, it was 1965.35 seconds for the proposed HQNN model using a similar dataset of ACCS. The testing time of the CNN model was recorded as 11.8 seconds and for the HQNN model, the testing time was recorded as 3.7 seconds. When testing the proposed hybrid neural network model, the security system responded faster than the classical model, enabling cybersecurity analysts to respond to malicious activities within seconds. The results prove that the proposed HQNN model is three times faster than CNN. Table 1 summarizes the statistics of the two models. There is a significant difference in the time consumption of both models for the same dataset.

The resulting confusion matrix with P = 0.9 is shown in figure 3, where we can see that in the case of 12.26% of incidents when the true label was "Attack" our proposed HQNN model falsely labeled it as "Normal". In case of 30.34% of incidents the true label was "Normal", and the HQNN model truly predicted the label as "Normal". In 37.68% of incidents the true label was "Attack", and the HQNN model also truly predicted the label as "Attack". Whereas, in 19.73% of incidents when the true label was "Normal", our proposed HQNN model falsely labeled it as "Attack". Consequently, it is concluded that our proposed HQNN model truly predicted 68% of the total incidents.

For the analysis of two different classifiers i.e. CNN classifier and HQNN classifier, we have used the receiver operating characteristic (ROC) curve. The ROC curve is a method of graphical representation to illustrate the analysis performance of a binary classifier. The ROC method was developed in 1941 for radar operators at military installations, which is why it was also called the receiver operating characteristic curve [20].

An ideal classifier has a true-negative rate of 1 and a false-negative rate of 0 [17], as shown by a circle in figure 4. The classifier in the case of the ACCS dataset has an area under the ROC curve of 0.723, compared to the best value of 1. The results shown in figure 4 represent a significant difference in the efficiency of the proposed HQNN model compared to the classical neural network model.

Figure 4 is the ROC curve of the hybrid quantum neural network model. This ROC curve shows that the false negative rate is 0.2454, the positive predictive value is 0.712, the sensitivity is 0.606, and the specificity is 0.7545, which proves that our proposed HQNN model is a practical approach for training any dataset.

Rapid response is of paramount value in cybersecurity. Experiments have shown that the accuracy of our proposed hybrid quantum neural network is lower as compared to the accuracy of CNN. The accuracy of HQNN was 68%, while the accuracy of the CNN model was 90%. A 22% decrease in accuracy and an increase in false positive rate are negligible if the cybersecurity analyst can receive faster threat warnings and consequently respond quickly and remediate the cybersecurity threats.

The proposed hybrid quantum neural network model is better for CPU resource utilization. Results show that during training and testing of the dataset, the classical neural network model utilizes 35% to 75% of the total available CPU resource, as shown in figure 5 (a), while our proposed HQNN model utilizes 95-100% of CPU resource, as shown in figure 5 (b). The underutilization of CPU resources in classical models is a waste of CPU resources. Therefore, it is hereby proved that the proposed HQNN model makes comparatively better usage of CPU resources.

6. Conclusion and future work

Hybrid quantum neural networks (HQNNs) are a new research area that can be efficiently used in cybersecurity for threat hunting and rapid attack detection of 6G wireless communication networks. In this study, we have used a dataset of ACCS (UNSW-BN15) for training and testing the proposed HQNN model. During the study, it was found that a classical neural network took a long time to train and test, while our proposed hybrid quantum neural network performed the same tasks with the same dataset in a shorter time. The results proved that our proposed HQNN model is four times faster in training and three times faster in testing the dataset compared to the classical neural network. However, further research contributions from the research community will improve the hybrid quantum neural network and make it a strong candidate for threat hunting, attack detection, and malware analysis of 6G wireless communication networks. Continuing our current research, we intend to develop better and more efficient algorithms for hybrid quantum neural networks to make them the best choice for the international scientific community.

Supplementary Materials: There is no supplementary material available, except the data set.

Funding: This research received no external funding.

Data Availability Statement: The dataset is available online at https://www.kaggle.com/mrwellsdavid/unswnb15.

Acknowledgments: I acknowledge the technical support and continuous guidance provided by Dr. M. Masoom Alam, and Dr. Nadeem Javed for this research work. I am thankful to Fraz Ahmad for participating in implementation and providing the data from ACCS. I am thankful to Sadia Nishat Kazmi and Allay Hyder Urooj for reviewing the paper and removing grammatical errors from the manuscript.

Conflicts of Interest: The authors declare no known conflict of interest.

References

1. Lu, J. (2019). Assessing the cost, legal fallout of Capital One data breach. *Law360 Expert Analysis*.
2. Jiang, B., Seif, M., Tandon, R., & Li, M. (2021). Context-aware local information privacy. *IEEE Transactions on Information Forensics and Security*, *16*, 3694-3708.
3. Guan, W., Perdue, G., Pesah, A., Schuld, M., Terashi, K., Vallecorsa, S., & Vlimant, J. R. (2021). Quantum machine learning in high energy physics. *Machine Learning: Science and Technology*, *2*(1), 011003.
4. Klco, N., Roggero, A., & Savage, M. J. (2022). Standard model physics and the digital quantum revolution: thoughts about the interface. *Reports on Progress in Physics*.
5. Feynman, R. P. (2018). Simulating physics with computers. In *Feynman and computation* (pp. 133-153). CRC Press.
6. Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., ... & Martinis, J. M. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, *574*(7779), 505-510.
7. Bouland, A., Fefferman, B., Nirkhe, C., & Vazirani, U. (2019). On the complexity and verification of quantum random circuit sampling. *Nature Physics*, *15*(2), 159-163.
8. Elben, A., Flammia, S. T., Huang, H. Y., Kueng, R., Preskill, J., Vermersch, B., & Zoller, P. (2023). The randomized measurement toolbox. *Nature Reviews Physics*, *5*(1), 9-24.
9. Schuld, M., & Petruccione, F. (2018). *Supervised learning with quantum computers* (Vol. 17). Berlin: Springer.
10. Bharti, K., Cervera-Lierta, A., Kyaw, T. H., Haug, T., Alperin-Lea, S., Anand, A., ... & Aspuru-Guzik, A. (2022). Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, *94*(1), 015004.
11. Aslam, S., Herodotou, H., Mohsin, S. M., Javaid, N., Ashraf, N., & Aslam, S. (2021). A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renewable and Sustainable Energy Reviews*, *144*, 110992.
12. Amin, M. H., Andriyash, E., Rolfe, J., Kulchytskyy, B., & Melko, R. (2018). Quantum boltzmann machine. *Physical Review X*, *8*(2), 021050.
13. Schuld, M., & Killoran, N. (2019). Quantum machine learning in feature hilbert spaces. *Physical review letters*, *122*(4), 040504.
14. Rengaraj, G., Prathwiraj, U., Sahoo, S. N., Somashekhar, R., & Sinha, U. (2018). Measuring the deviation from the superposition principle in interference experiments. *New Journal of Physics*, *20*(6), 063049.
15. Fujii, K., Kobayashi, H., Morimae, T., Nishimura, H., Tamate, S., & Tani, S. (2018). Impossibility of classically simulating one-clean-qubit model with multiplicative error. *Physical review letters*, *120*(20), 200502.
16. Killoran, N., Bromley, T. R., Arrazola, J. M., Schuld, M., Quesada, N., & Lloyd, S. (2019). Continuous-variable quantum neural networks. *Physical Review Research*, *1*(3), 033063.
17. Schuld, M., Bocharov, A., Svore, K. M., & Wiebe, N. (2020). Circuit-centric quantum classifiers. *Physical Review A*, *101*(3), 032308.
18. Australian Centre for Cyber Security (ACCS). (2018). UNSW-NB15 by IXIA Perfect Storm tool. https://www.kaggle.com/mrwellsdavid/unswnb15
19. Nevin, J. A. (1969). Signal detection theory and operant behavior: A review of David M. Green and John A. Swets' Signal detection theory and psychophysics1. *Journal of the Experimental Analysis of Behavior*, *12*(3), 475.
20. Hammersley, J. (2013). *Monte carlo methods*. Springer Science & Business Media.
21. Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, *2*(4), 433-459.
22. Ahmed, S., Khan, Z. A., Mohsin, S. M., Latif, S., Aslam, S., Mujlid, H., adil, M., & Najam, Z. (2023). Effective and Efficient DDoS Attack Detection Using Deep Learning Algorithm, Multi-Layer Perceptron. *Future Internet*, *15*(2), 76.
23. Akber, S. M. A., Kazmi, S. N., Mohsin, S. M., & Szczęsna, A. (2023). Deep Learning-Based Motion Style Transfer Tools, Techniques and Future Challenges. *Sensors*, *23*(5), 2597.
24. Yang, C., Deng, Y., Yao, J., Tu, Y., Li, H., & Zhang, L. (2023). Fuzzing automatic differentiation in deep-learning libraries. arXiv preprint arXiv:2302.04351.