TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

A Parallel Quantum Feature Encoding Scheme for Effective Classical Data Classification in Quantum Convolutional Neural Networks

Raisa Mashtura[1], Jishnu Mahmud[1], Shaikh Anowarul Fattah[2] and Mohammad Saquib[3]

Abstract*—Quantum machine learning is one of the most exciting new avenues in the world of artificial intelligence, especially because of the enormous computational power of quantum computers and the promise of the development of near error-free quantum computers in the not-so-distant future. For quantum algorithms to be used in real-life applications, quantum computers must be able to work with classical data. One of the key steps in quantum algorithms dealing with classical data is the encoding of classical data points to quantum states, which can then be processed by quantum gates. It is known that the type of encoding technique that works best for a particular network is dependent on the dataset being used. In this paper, a new parallel structure is proposed utilizing two encoding techniques, namely amplitude encoding and angle encoding, for effective classical data classification via quantum neural network. The paper further proposes a maximally expressible and entangled ansatz used to design a simple Quantum Convolutional Neural Network (QCNN) with only 32 parameters, that is used in the latter stages of the network and is kept the same across all encoding instances so that a comparison between the different encoding methods is possible. Extensive experimentation is carried out on two publicly available image datasets, namely *MNIST* and *Fashion MNIST*. The results show that the proposed method achieves better results than any of the encoding techniques deployed alone for binary classification.

***Index Terms*—Amplitude, Angle, Encoding, Quantum, Qubit**

I. Introduction

This decade has seen enormous growth in the field of Machine Learning, and it is being realized that its application potential in various fields is infinite. As the computational power of modern computers increases, more complex and deeper networks with an enormous number of parameters can be trained to capture the most nuanced information from the data. Image classification is one of the most fundamental problems of computer vision and has enjoyed an exponential increase in its abilities using Convolutional Neural Networks (CNNs). However, with the increase in depth and the number of trainable weights, new problems have been introduced in modern networks. Such impairments include the barren plateau problem, where the training process is trapped in a local minimum of the cost function and therefore the network stops the optimization of the weights despite a substantial learning rate. In the post-modern world, where there is access to an enormous amount of data, the development of algorithms that can truly reap the benefits of the volume to learn more nuanced information is deemed imperative.

The field of quantum computing seems to be a formidable competitor when it comes to solving this problem. The enormous computational power of quantum computers coupled with the recent developments which indicate the creation of near error-free quantum computers in the not-so-distant future makes it a compelling choice for implementing machine learning algorithms in this field [1]. Quantum Computers can capture complex information from the data provided with only a fraction of trainable weights and shallower networks compared to their classical counterparts. This makes them more resistant to the problems arising from deeper networks with millions of trainable parameters and suitable for tasks where more complex information must be extracted from the data. The quantum advantage primarily arises due to their operability based on *qubits*, which are the superposition of the two fundamental bit states and the entanglement phenomenon that these qubits exhibit [2].

However, Quantum Machine Learning (QML) has its impairments, and its use is therefore still limited in applications in the real world. Although there are promises of near-error-free quantum computers, the reality of today is that most quantum computers available are very noisy and sensitive to external environmental factors [3]. In this era of (NISQ) computing, the limitation in the number of feasible qubits presents a great challenge in the design of Quantum Neural Networks, making it essential to reduce classical data dimension before it can be embedded into initiated qubits as well as design a quantum network with minimum cost. The importance of mapping this reduced classical data into the Hilbert Space effectively thus arises, to be classified with a variational quantum circuit whose parameters are to be optimized using classical back-propagation. The form of network used in this paper mimics a convolutional structure, as they are resistant to the barren plateaus issue [4]. The effective method of mapping classical data into quantum states for various purposes has been found to be dependent on the data to be processed, leaving the choice of quantum feature encoding an open problem. Many proposed methods accomplish this task, with Amplitude Encoding and Angle Encoding being one of the most widely used encoding methods.

In previous works, the best quantum encoding technique has been shown to vary with the dataset to be classified. The performance of a network had to be evaluated individually for both mapping techniques for each dataset [5], [6]. While

[1]Raisa Mashtura and Jishnu Mahmud are with the Faculty of CSE, BRAC University, 66 Mohakhali, Dhaka 1212, Bangladesh. They are pursuing their MSc. at the Dept. of EEE at the Bangladesh University of Engineering and Technology (BUET), West Palashi, Dhaka 1000, Bangladesh. raisa.mashtura@bracu.ac.bd, jishnu.mahmud@bracu.ac.bd

[2]Shaikh Anowarul Fattah is a professor at the Dept. of EEE at the Bangladesh University of Engineering and Technology (BUET), West Palashi, Dhaka 1000, Bangladesh. fattah@eee.buet.ac.bd

[3]Mohammad Saquib is a professor at the Dept. of EE at The University of Texas at Dallas 2601 N. Floyd Road Richardson, Texas 75083-0688, USA. saquib@utdallas.edu

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

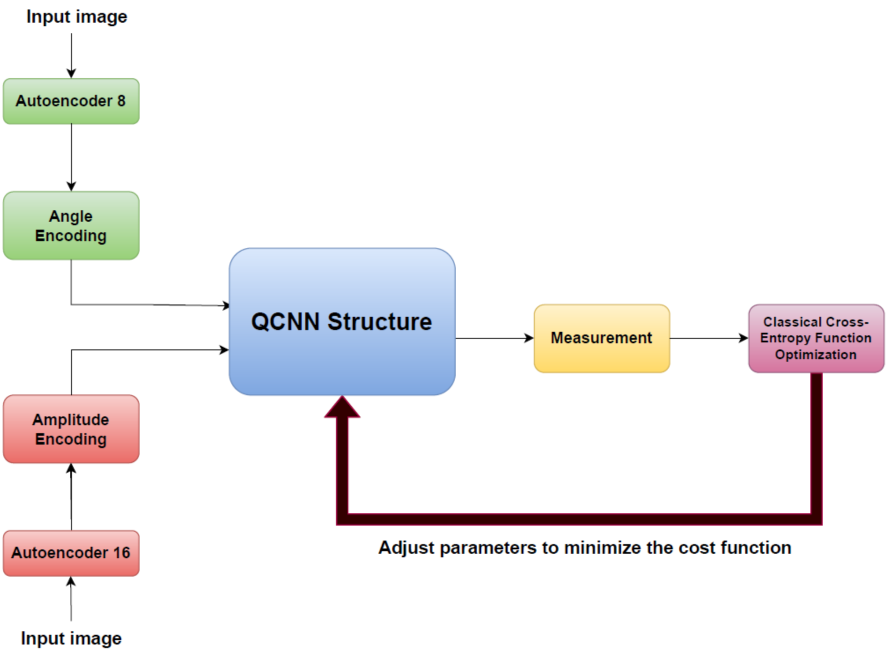

Fig. 1: A simplified block diagram illustrating the flowchart of the proposed methodology

this issue has motivated the invention of hybrid encoding techniques, classification using a weighted result by carrying out both encoding techniques simultaneously, especially for a Quantum Convolutional Neural Network (QCNN), is a first to the best of our knowledge.

Furthermore, an ansatz has also been devised, which is a building block of a quantum convolutional layer, that displays a maximum expressibility and entanglement while maintaining an optimum number of parameters and depth for proper trainability of the network.

II. Proposed Method

A. Quantum Feature Encoding

Due to the limitation in the number of employable qubits, the features of images are extracted using a classical auto-encoder. Auto-encoders help reduce the unwanted content in the data by reducing dimensions while preserving only the regions of interest. This is an effective way to reduce data especially when data of smaller dimension is needed to be fed into a quantum network. In order to map the classical output values of the auto-encoder from the classical domain to quantum states in the Hilbert space, two commonly used quantum encoding techniques namely Amplitude Encoding and Angle Encoding are conducted in parallel in our proposed network. The two encoding techniques are discussed in the subsequent section,

1) Amplitude Encoding: In Amplitude Encoding, the classical data points of the features are normalized and then represented as amplitudes of the basis states formed by n-qubits. As a result, features of size 2^n can be represented by n-qubits [7] as follows:

$$\left|\psi_s\right\rangle = \sum_{i=1}^{2^n} c_i |i\rangle \tag{1}$$

Here, $|\psi_s\rangle$ is the quantum state prepared after mapping the 2^n-dimensional classical datapoint C after reduction, $|i\rangle$ is the i-th computational basis state and c_i is the i-th element of the datapoint C.

In the block for Amplitude encoding, as shown in Fig.2, 4 qubits have been assigned so that features of size 16 could be mapped using them.

2) Angle Encoding: In Angle Encoding [8], the features are embedded as phases into the parameters of quantum gates, hence determining the angles of rotation gates. The embedded feature vector of length n can be represented by n qubits as:

$$\left|\psi_s\right\rangle = \otimes_{i=1}^{n} R(x_i)\left|0^n\right\rangle \tag{2}$$

$R(.)$ can denote any of the rotation gates R_x, R_y, or R_z.

In this case, as illustrated in Fig.2, 8 qubits are fed into the encoder block to perform Angle Encoding on a feature vector of dimension 8.

It is anticipated that broadening the domain of data being represented as both amplitude and angle of quantum states independently will help aid the training of a QCNN and lead to enhanced classification performance.

B. Proposed Quantum Convolutional Neural Network

Analogous to a classical CNN, a QCNN consists of convolutional and pooling layers, which are stacks of convolutional and pooling filters, respectively. These filters or ansatzes are quantum circuits made of various parameterized quantum gates. In order to constrain a quantum circuit's cost, all gates employed in the proposed network are kept limited to one-qubit or two-qubit interactive gates as shown in Fig.3.

The ansatz used to constitute the convolutional layer of the proposed architecture is aimed to provide maximum expressibility as well as entanglement with a minimal number of layers of ansatz. Among various types of circuits in [9], a four-qubit ansatz shown in Fig.2 is chosen that satisfies the aforementioned demand. However, the number of parameters in this ansatz only being two, affected the trainability of the

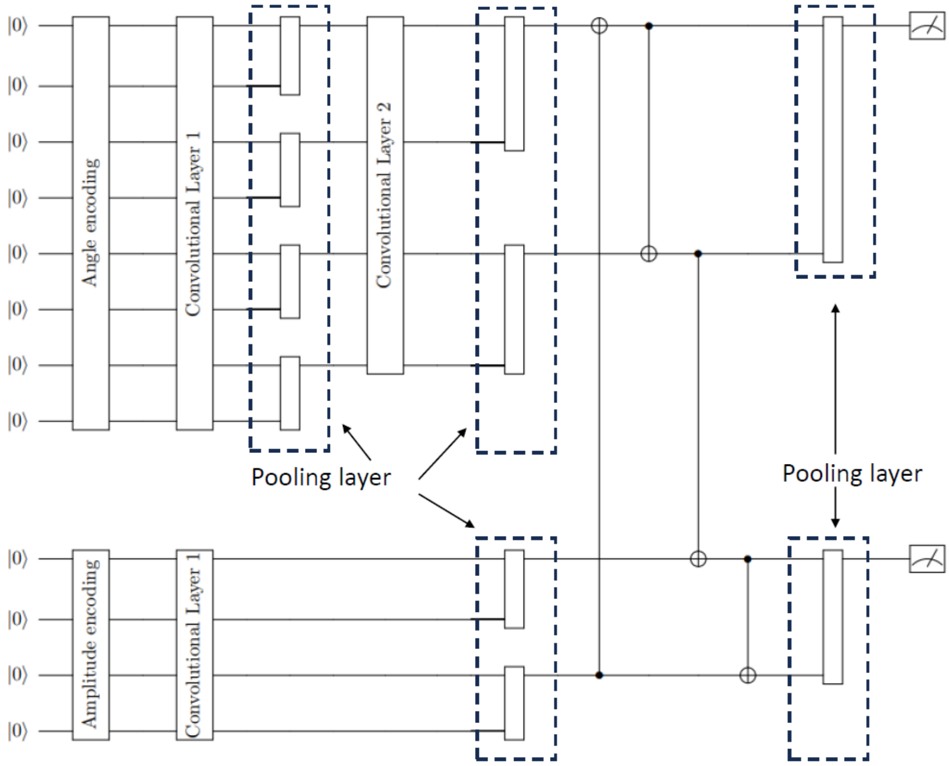

Fig. 2: The two different encoding schemes used in a parallel QCNN structure

network and consequently, this ansatz is added with a U_3 gate at each qubit after being reduced to a 2-qubit interactive structure as shown in Fig.3. The $Hadamard$ gate along with the $Controlled-Z$ gate establishes entanglement among the two qubits. The R_y gates cause rotation of the qubit about the y-axis and improve the expressibility of a quantum circuit [9]. The U_3 gate is a combination of rotation and phase shift gates added with the intention of increasing the flexibility of the network.

The pooling ansatz establishes a weighted interdependency with the help of *Controlled-Rotation* and $Pauli$ gates. Following that, one qubit is discarded and fed into the next part of the network as illustrated in Fig.2.

In the overall structure, two convolutional and three pooling layers are employed at each parallel path of the network. After the second pooling layer, a quantum analogous of a fully connected layer is executed in order to establish interaction among the two paths with the help of *CNOT* gates.

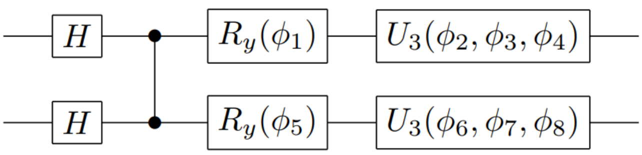

Fig. 3: The ansatz used to build the convolutional layers

C. Cost Function optimization

The quantum states prepared by the network are then measured, which means that they are made to collapse from a probabilistic quantum state to a deterministic classical value. The expectation values of the qubits are then measured and subsequently subjected to the softmax function. The outputs are then used to calculate the cross-entropy loss function, which is carried out in the classical domain. Mathematically the cost function can be expressed as the following.

$$loss = \sum_{i=1}^{output\ size} y_i.log(\bar{y}_i) \tag{3}$$

where y_i is the true-label and $\bar{y}_i$ is the predicted probability of the corresponding class.

The minimization of this objective function is carried out to optimize the parameters of the ansatz making up the QCNN layers.

III. Results and Analysis

A. Datasets

The first two classes of the two standard and widely employed datasets, *MNIST* [10] and *Fashion MNIST* [11], were chosen for classification. For both the datasets, the first two classes were chosen in order to perform binary classification, with each class containing 6000 training and 1000 test images, each of size 28×28. The images were then subjected to classical feature extraction using an auto-encoder.

TABLE I: Results for Binary Classification

Dataset	Quantum Encoding Techniques' Accuracy (%)		
	Proposed Amplitude & Angle parallel structure	**Amplitude**	**Angle**
Fashion MNIST	90.9	86.7	84.1
MNIST	93.3	86.4	90.8

B. Simulation and Results

The proposed Parallel QCNN is simulated in Pennylane [12] where the variational parameters of the quantum gates are optimized using the Nesterov moment optimization algorithm [13]. Small batches of 75 images are selected at random and fed into the network where it is duplicated and consequently encoded to their respective quantum states via Amplitude and Angle Encoding methods in parallel branches. Then they are passed through their respective QCNN structure where the parameters of the gates of both the convolutional as well as the pooling layers are initialized randomly. Finally, when the expectation values of the two measurements are taken, they are fed as input to the classical cost entropy function where the gradients are calculated. The use of small batches helps the training process to be quicker and avoids the cost function to be trapped in a local minima.

Table I shows the classification accuracies for the binary classification of classes 0 and 1 for *MNIST* and *Fashion MNIST* datasets for various encoding techniques. It can be observed that for different datasets the type of encoding that provides the best results are different; Amplitude Encoding results in an accuracy of 86.7% for *Fashion MNIST* and Angle Encoding results in a classification accuracy of 90.8% for *MNIST*.

The use of the parallel hybrid encoding technique solves that problem where it manages to achieve state-of-the-art accuracies of 90.9% and 93.93% for *Fashion MNIST* and *MNIST* datasets respectively, with a few parameters and a minimum number of layers; displaying high expressibility of the network. The results show a clear superiority of the parallel structure to those where Amplitude and Angle Encoding has been deployed alone. The proposed parallel structure along with the modifications in the QCNN network enables the network to learn more nuanced information for a wider range of data and automatically applies more weight to the encoding technique which outperforms its counterpart. The fact that the parallel encoding technique outperforms the best encoding technique for each of the two datasets goes on to prove that even if a particular encoding technique performs better than the other for a dataset, there is still some information that is extracted by the lesser-performing encoding technique which goes on to be unnoticed by its better counterpart. It is this information that the proposed parallel encoding technique can extract from a dataset, which enables it to outperform even the best encoding techniques in each of the datasets.

IV. Conclusion

In this paper, a new parallel encoding technique coupled with a QCNN structure with a very small number of trainable parameters is proposed. It is shown that the network is extremely effective in dealing with different types of datasets. The paper aims to demonstrate that the quantum encoding methods that are not the best in terms of classification accuracy, still manage to extract some information that is overlooked by the encoding method bearing the best results for a particular dataset. This means that even the best quantum encoding techniques suited for a particular dataset are unable to extract and interpret all the classical information content when encoding them to quantum states and it is for that reason a parallel structure with different encoding techniques can reap more accurate results across multiple datasets. This means that the development of more generalized and robust quantum encoding techniques is immensely important for the development of classical data processing using quantum machine learning algorithms. Due to computational and hardware limitations, only a small network with a minimal number of qubits has been simulated in this paper. Future works may include the development of similar parallel structures using combinations of various existing encoding techniques and carrying out in-depth investigations regarding the correlations between the performance, type of data, and encoding techniques. In this way, a more robust, and generalized encoding method can be devised which can then be used to encode a wide array of classical data.

Acknowledgment

The authors would like to express their sincere gratitude towards the authorities of the Department of Electrical and Electronic Engineering at Bangladesh University of Engineering and Technology (BUET) and the authorities at the Department of Computer Science and Engineering at BRAC University for providing their constant support throughout this work.

References

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[2] N. Yanofsky and M. Mannucci, *Quantum Computing for Computer Scientists*. Cambridge University Press, 2008.

[3] B. Cheng, X.-H. Deng, X. Gu, Y. He, G. Hu, P. Huang, J. Li, B.-C. Lin, D. Lu, Y. Lu *et al.*, "Noisy intermediate-scale quantum computers," *Frontiers of Physics*, vol. 18, no. 2, p. 21308, 2023.

[4] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, "Absence of barren plateaus in quantum convolutional neural networks," *Physical Review X*, vol. 11, no. 4, p. 041011, 2021.

[5] T. Hur, L. Kim, and D. K. Park, "Quantum convolutional neural network for classical data classification," *Quantum Machine Intelligence*, vol. 4, no. 1, pp. 1–18, 2022.

[6] D. Sierra-Sosa, M. Telahun, and A. Elmaghraby, "Tensorflow quantum: Impacts of quantum state preparation on quantum machine learning performance," *IEEE Access*, vol. 8, pp. 215 246–215 255, 2020.

[7] M. Schuld and F. Petruccione, *Supervised learning with quantum computers*. Springer, 2018, vol. 17.

[8] M. Schuld, "Supervised quantum machine learning models are kernel methods," *arXiv preprint arXiv:2101.11020*, 2021.

[9] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1900070, 2019.

[10] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE signal processing magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[11] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[12] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi *et al.*, “Pennylane: Automatic differentiation of hybrid quantum-classical computations,” *arXiv preprint arXiv:1811.04968*, 2018.
[13] Y. E. Nesterov, “A method for solving the convex programming problem with convergence rate,” in *Dokl. Akad. Nauk SSSR,*, vol. 269, 1983, pp. 543–547.