Gradient-based Quantum Neural Network Using Quantum Computing Chips

Y.C. Zhan1*, H. Zhang1*, H. Cai1,2, D. P. Poenar1, L. C. Kwek1,3, and A. Q. Liu1†

1 Quantum Science and Engineering Centre (QSec), Nanyang Technological University, Singapore
*2 Institute of Microelectronics, A*STAR (Agency for Science, Technology and Research), Singapore*
3 Centre for Quantum Technologies, National University of Singapore, Singapore
*†Email Address: eaqliu@ntu.edu.sg (*equal contribution)*

Abstract: An analytic gradient algorithm for training neural networks in an 8-mode quantum photonic chip is proposed. The algorithm shows 4x faster-training speed and 3% accuracy rate enhancement than existing approaches in solving two classification tasks.

OCIS codes: (270.0270) Quantum Optics; (200.4260) Neural Networks; (250.5300) Photonic Integrated Circuits

1. Introduction

Quantum neural networks [1] develop classical models with the advantages of quantum computing, providing faster speed and power-efficient calculation mode. However, it is difficult to train the quantum neural network in the photonic chip with traditional gradient descent algorithms [2], as a physical system is not differentiable. The alternative training methods include gradient-free methods [3] and numerical gradient methods [4,5], which are less effective for their high computation and measurement complexity. Here, we propose an analytic gradient algorithm to train a photonic chip-based quantum neural network. The method calculates the gradient of the chip in a divide-and-conquer strategy and updates the chip parameters by the obtained gradient. The method is evaluated by classification tasks including Binary classification and Iris Flower classification on the chip. The results present more than 4 times faster speed and 3% accuracy rate enhancement.

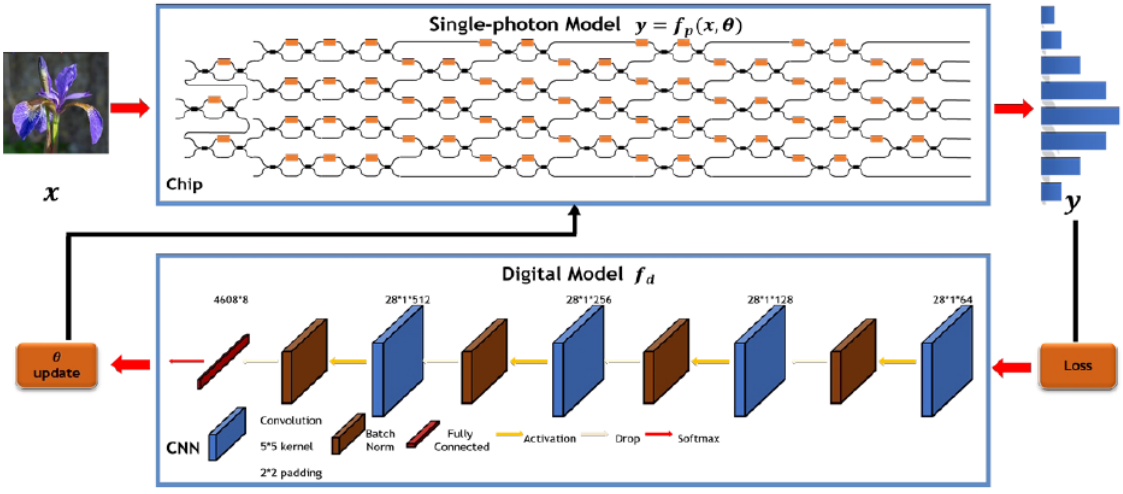

Figure 1. The schematic diagram of quantum classifier training. The quantum classifier model consists of a single-photon model fp and a pretrained digital model fd.The fp describes an 8-mode photonic chip and fd is a 5-layer CNN. The parameters θ are updated according to the analytic gradient in the digital model.

2. Design

A quantum classier for the tasks is designed in an 8-mode chip, as shown in Fig. 1. The algorithm is gradient-based, where analytic gradients are derived in software on the backward propagation, and parameters are updated in hardware by gradients on the forward propagation. In forward propagation, the model f_p categorizes input datasets x into different classes y depending on parameters θ, the process of which is given by $y=f_p(x,\theta)$. The f_p adopts Clement's network of Mach–Zehnder Interferometers (MZIs), where a 2×2 MZI has two free parameters defined by

$$T_{MZI} = ie^{i\frac{\theta}{2}}\begin{bmatrix} e^{i\phi}\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \\ e^{i\phi}\cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \end{bmatrix}, \quad (1)$$

and the arrangement of 28 MZIs realizes an 8-mode quantum computing chip. In backward propagation, a digital model fd is pretrained to model the property of fp with the same inputs and outputs, to calculate the analytic gradient for each θ in fp. It adopts a 5-layer convolutional neural network (CNN) with 1 million training parameters. The loss between y and target $y$target is defined by Kullback-Leibler divergence loss $L(y,y_{\text{target}})=y_{\text{target}}\log(y_{\text{target}}/y)$, and the gradients g_{out} on the output vectors are derived by partial differentiation, which is given by

$$g_{out} = \frac{\partial L}{\partial y}(y,y_{\text{target}}). \quad (2)$$

The gradients $g\theta$ of parameters θ are computed according to the Jacobian Matrix and the Chain Rule for the derivation of the digital model fd, which is given by

$$g_\theta = [\frac{\partial f_d}{\partial \theta}(x,\theta)]^{\mathrm{T}} g_{out}, \quad (3)$$

where T is the transpose matrix. Then the parameters can be updated according to the gradient descent algorithm, $\theta\rightarrow\theta—\eta g_\theta$, where η is the learning rate and g_θ is the average gradient in one batch.

3.Results and Discussions

The analytic gradient algorithm is experimentally applied in the quantum classifier to classify the Binary dataset and Iris Flower dataset. The results are shown in Fig. 2 and 3, respectively. Single photons are generated on the chip and supplied as the input to the quantum classifier. The output photons are detected by an 8-ports single-photon detector with an efficiency of ~88%.

The Binary dataset contains 2 nonlinear classes with 40000 points for training, 1000 points for validation, and 10000 points for testing. The quantum classifier is a 4-mode structure in the chip with 6 tunable weights. Fig. 2(a) shows the training and validation accuracy rate per measurement and Fig. 2(b) depicts the classification result. The

average accuracy in the testing dataset is 97.92%, presenting an accurate separation between data clusters than other training methods.

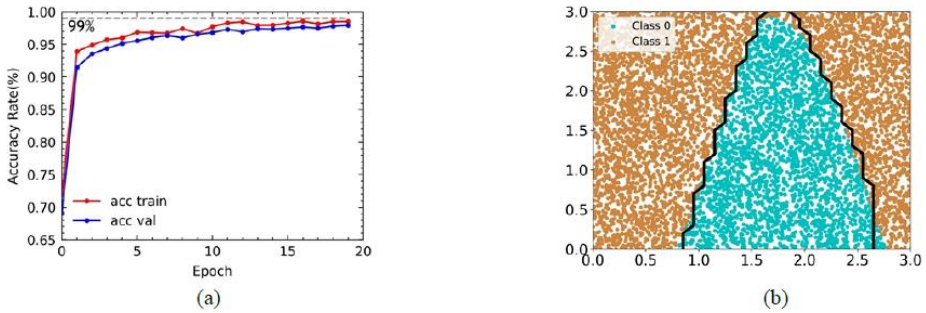

Figure2. Binary classification results. (a) Accuracy rate in training and validation dataset. (b) Decision boundary (black line) on the testing dataset.

The Iris Flower dataset comprises 3 flower species with 4-dimensional pattern information as shown in Fig. 3(a). The dataset is divided into 120 images for training and 30 images for testing. The quantum classifier uses the 8-mode chip with 28 tunable weights. Fig. 3(b) and Fig. 3(c) present the training and testing result. The results reveal analytic gradient training converges within 25 epochs ([3] is 105 epochs and [4] is 210 epochs) and has a 96.7% accuracy rate ([3] is 93.3% and [4] is 90.0%) for a real image dataset.

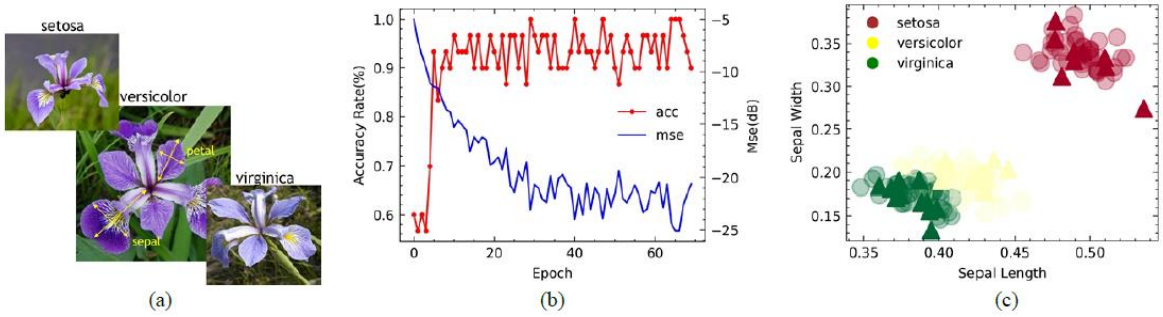

Figure 3. Iris Flower classification result. (a) Three flower species include setosa, versicolor and virginica. (b) Accuracy rate and Mean squared error (Mse) in training. (c) Visualization of the classification results in 2-dimension, the training dataset (round), and the testing dataset (triangle).

4. Summary

We propose an analytic gradient algorithm for training neural networks in the quantum photonic chip. In machine learning tasks of classification, the algorithm utilizes the photonic chip model for fast-speed forward computation and the digital model for high-accuracy backward computation. The gradient-based quantum chip training has broad application potential in variational quantum circuits and quantum machine learning.

5. Acknowledgement

This work was supported by the Singapore Ministry of Education (MOE) Tier 3 grant (MOE2017-T3-1-001), and PUB Singapore's National Water Agency grant (PUB-1804-0082).

5. References

[1] Zhang H, *et al*. An optical neural chip for implementing complex-valued neural network[J]. Nature communications, 2021, 12(1): 1-11.
[2] Bottou L. Stochastic gradient descent tricks[M]. Neural networks: Tricks of the trade, 2012: 421-436.
[3] Zhang H, *et al*. Efficient on-chip training of optical neural networks using genetic algorithm[J]. Acs Photonics, 2021, 8(6): 1662-1672.
[4] Mitarai K, *et al*. Quantum circuit learning[J]. Physical Review A, 2018, 98(3): 032309.
[5] Xu X, *et al*. Self-calibrating programmable photonic integrated circuits[J]. Nature Photonics, 2022, 16(8): 595-602.