Frontiers in Optics + Laser Science 2023 (FiO/LS) © Optica Publishing Group 2023

Digital Electronic and Analog Photonic Acceleration for Point Cloud Classifiers

James Garofolo,[1] Taichu Shi,[1] Thomas Ferreira de Lima,[2,3] Paul Prucnal,[3] and Ben Wu[1,*]

[1]*Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ 08028, USA*
[2]*NEC Laboratories America Inc. 4 Independence Way Suite 400 Princeton, NJ 08540*
[3]*Lightwave Communications Laboratory, Department of Electrical Engineering, Princeton, NJ 08544*
wub@rowan.edu

Abstract: We propose a method for training photonic neuromorphic accelerators using only software, and demonstrate its potential in autonomous driving. The method is modular and drop-in compatible with existing PyTorch training pipelines.

OCIS codes: 200.4260 Neural networks; 200.4700 Optical neural systems

1. Introduction

LiDAR is a promising technology for applications in machine vision and autonomous driving systems [1]. Because of this, various machine learning researchers have designed architectures which are able to classify objects using LiDAR point clouds. One such model is PointNet [2], which is depicted in Fig. 1a. While models like these are designed to be as computationally efficient as possible, inference throughput remains a key challenge in implementing these models in such real-time applications. This is due to the high volume of data being presented to the models by commercial sensors, which cannot be lowered without a severe loss in classification accuracy, meaning that the computations must be accelerated to meet the real-time requirements of autonomous vehicles.

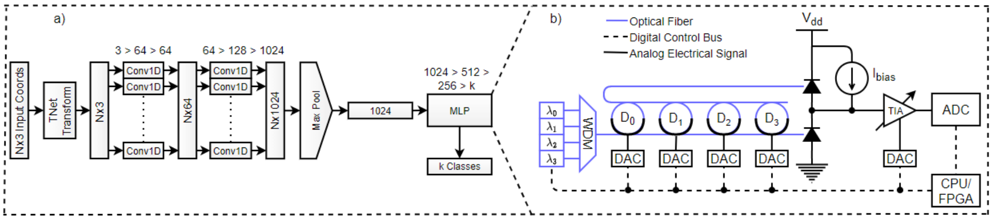

Fig. 1. a) Block diagram of PointNet [2] b) Schematic of the DEAP accelerator described in [3]

Photonic analog tensor processors have been an emerging technology in academia [4], and support from digital electronics has recently enabled them to be scalable enough to accelerate deeper neural networks in practice [3]. These scalable architectures, referred to as Digital Electronic and Analog Photonic (DEAP) architectures, perform partial Multiply-Accumulate (MAC) operations on vectors encoded as optical intensities. This is done using electrically tunable optical micro-ring resonators (MRRs), which can be combined with a balanced photodetector to independently multiply different wavelengths of light by coefficients on the interval [-1,1], sum the products together, and output the resulting value as an electrical current. A transimpedance amplifier (TIA) can then be used to transform the weights into any arbitrary range, along with converting the electrical current to a voltage which can be measured and digitized. This architecture, proposed by [3] and depicted in Fig. 1b, is capable of performing four multiplications and four additions at once, with a maximum operation rate of 5GHz being set by the sampling circuitry. These operations can be chained together with digital additions to implement full vector and matrix multiplications at much higher rates than conventional digital computers.

In this paper, we propose a method for training these DEAP neuromorphic accelerators, and demonstrate its effectiveness on PointNet. This is done through digital simulations implemented on the PyTorch machine learning library [5], so as to enable scalability of the model, flexibility of the setup, and ease of reproduction for machine learning engineers with no prior knowledge in analog computing. The model uses the control signals as trainable parameters, meaning that the training process requires no access to hardware, as well as that the effects of all MRRs are accounted for when calculating weights, as recommended by [6].

2. System Setup

The digital DEAP processor model is described in Eqs. 7, 10 and 11 of [3], and has been fitted to match the transmission spectrum of the device used in that paper's experiments. This simulated spectrum enables the implementation of four neural network weights at a time, which can be tuned via the power in mW applied to the four MRR heaters and the gain of the TIA as shown in Fig. 1b. For vectors of more than four elements, different tuning powers and gains can be applied, representing different passes through the circuit.After generating the weights, a simple batch-matrix multiplication can be used to apply the weights to the inputs, enabling any basic PyTorch module to be replaced by a simulated DEAP processor.

To verify that this model works as intended, we replaced all of the trainable layers in PointNet with simulated DEAP processor layers, and trained the new model as one would the original. The model was made to minimize negative-log-likelihood loss using the ADAM optimizer with a learn rate of 0.001 and β values of 0.9 and 0.999. The model was trained on point clouds of 2500 points sampled from one of 16 types of ShapeNet object [7], with 12137 objects per epoch separated into batches of 32 and evaluation stops every 10 batches. A test benchmark was made to be used after training on model snapshots from after each epoch. To ensure a quality local minimum is reached, the model was set to anneal its learn rate by a γ of 0.9 in the event of 5 evaluation stops not yielding improved results over the observed peak performance.

3. Results and Analysis

The classification accuracy over epoch for the DEAP model is shown in Fig. 2b. This figure shows that the control parameters of a DEAP accelerator can be trained directly to learn a pattern, though the training process is slower and less stable than its digital counterpart. It is important to note that the final accuracy displayed in this model is lower than that of a software network due to the high sensitivity of the model's parameters, and not the ability of the architecture to learn patterns. This accuracy could be improved with longer training, hyper-parameter tuning, or by using least-squares minimization to fit the DEAP model to a pre-trained software model with satisfactory accuracy.

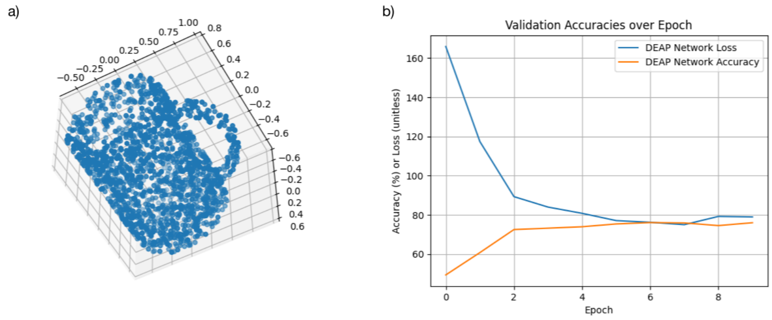

Fig. 2. a) Point cloud of a coffee mug sampled from the ShapeNet dataset [7]. b) Evaluation accuracy and loss over epoch for DEAP model.

697609280 passes through the DEAP core are required to compute one inference of PointNet. This means that, at the current bandwidth limit of 5GHz, a single instance of the DEAP architecture is capable of performing an inference on PointNet in as little as 139.522ms. This could be reduced further by adding more DEAP units, or by adding more MRRs to the units under test, allowing for the process to be further parallelized.

4. Conclusion

We propose a method for training digital electronic and analog photonic neuromorphic accelerators using only software. This method uses the modeled spectral response from existing hardware to simulate the circuits that generate weights in photonic neural networks. We then illustrate that these models can be made drop-in compatible with existing software modules and trained in the same pipeline. The model was able to reach accuracies above 75% without specialized hyperparameter tuning, and is capable of classifying point cloud object data in as little as 139.522ms with only one core operating at a time.

References

1. Ying Li, Lingfei Ma, Zilong Zhong, Fei Liu, Michael A. Chapman, Dongpu Cao, and Jonathan Li. Deep learning for lidar point clouds in autonomous driving: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8):3412–3432, 2021.
2. R. Qi Charles, Hao Su, Mo Kaichun, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 77–85, 2017.
3. Viraj Bangari, Bicky A. Marquez, Heidi Miller, Alexander N. Tait, Mitchell A. Nahmias, Thomas Ferreira de Lima, Hsuan-Tung Peng, Paul R. Prucnal, and Bhavin J. Shastri. Digital electronics and analog photonics for convolutional neural networks (deap-cnns). *IEEE Journal of Selected Topics in Quantum Electronics*, 26(1):1–13, 2020.
4. Paul R. Prucnal, Bhavin J. Shastri, and Malvin Carl Teich. *Neuromorphic Photonics*. CRC Press, Taylor amp; Francis Group, 2017.
5. Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
6. Weipeng Zhang, Chaoran Huang, Hsuan-Tung Peng, Simon Bilodeau, Aashu Jha, Eric Blow, Thomas Ferreira de Lima, Bhavin J. Shastri, and Paul Prucnal. Silicon microring synapses enable photonic deep learning beyond 9-bit precision. *Optica*, 9(5):579–584, May 2022.
7. Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, and Fisher Yu. Shapenet: An information-rich 3d model repository, 2015. cite arxiv:1512.03012.