ClassificationCam:
Toward an all-optical convolutional neural network

Julie Chang
Stanford University
jchang10@stanford.edu

Gordon Wetzstein
Stanford University
gordon.wetzstein@stanford.edu

ABSTRACT
Abstract here.

1 INTRODUCTION

Deep neural networks have found success in a wide variety of applications, ranging from computer vision to natural language processing to game playing [4]. Since the explosion of interest following the achievements of convolutional neural networks on ImageNet classification, deep learning has transformed algorithms in both academic and commercial environments. While accuracy has improved to a remarkable level, the number of parameters and connections has grown dramatically, and the power requirements to build and use these networks have increased correspondingly.

While the training phase of learning parameter weights is often considered the slow stage, large models also demand significant energy during inference due to millions of repeated memory references. For example, AlphaGo has a power consumption of approximately 300 W per game. To this end, there has been a large effort to develop new software methods and specialty hardware for improved efficiency. Algorithms for improved efficiency include pruning, quantization, low rank, parallelization, mixed precision, and model distillation. Compressed models have demonstrated preserved accuracy with much fewer parameters [2, 3]. On the hardware front, there are now specialized processing units for deep learning, such as TrueNorth, Movidius's USB-based neural compute stick (NCS), and Google's tensor processing unit (TPU). Despite all these efforts, it remains difficult for embedded systems such as mobile vision, autonomous vehicles/robots, and wireless smart sensors to deploy CNNs due to stringent constraints on power and bandwidth.

Optical computing has been tantalizing for its high bandwidth and inherently parallel processing. If we can come up with scalable optical configurations that together act as a framework for an optical CNN, this would be of interest to computer vision, robotics, machine learning, and optics communities. Optical neural networks (ONNs) could also potentially exploit wave optics for complex-valued neural networks and new types of non-linearities that are currently unavailable to digital computation. Linear transformations can be performed with lens systems or interferometer meshes [7]. Optical nonlinearities include saturable absorption and bistability. Liquid crystal light valves have been used as optical thresholding devices [6]. Convolutions are commonly performed with PSF engineering. We use these to design a model for cascaded parallel convolutions with sandwiched nonlinearity layers.

In this paper, we follow the vein of computational photography to create a zero-power, all-optical convolutional neural network for image classification. We choose a simple classification task, e.g. classify handwritten digits, and build a prototype that performs inference on projected images. We precompute the weights by training with the MNIST dataset on a computer, then fabricate the optical elements accordingly. We compare performance with the same inference performed on the computer. Here we demonstrate proof-of-concept with bulk optics and free-space propagation, but we recommend photonic integrated circuits for scalability. Photonic circuits with up to 4,096 optical devices have been demonstrated [8], and there have also been new three-dimensional photonic integrations that could enable larger networks [5]. Combination of these next-generation large-scale photonic circuits with compressed deep learning models could provide a potential route for high performance ONNs.

Limitations. We are limited to non-negative values since we are working with light intensities. This may be avoided with coherent processing. We also hope to proceed without normalization. We focus here on inference and do training on the computer. In our current implementation without active elements, we lose flexibility to update the network. However, in fixed applications, this is not a problem. Photonic circuits remain expensive to fabricate as they are not used in common consumer applications.

To summarize, we make the following contributions:

- We propose an optical toolbox of building blocks for convolutional neural networks.
- We build a zero-power, all-optical two-layer network with precomputed weights for image classification.
- We evaluate against the computer implementation of the same network and show that we achieve similar accuracy.

[Handwritten annotations: "W (not per game)"; margin list: Coherent, Fathom computing, Emphasize camera aspect, Related works, CNNs, Hybrid approaches, Photonics]

2 RELATED WORKS

Angle sensitive pixels to approximate the kernels of the first layer of a typical convolutional neural network [1]

Recently, two-layer fully connected NN demonstrated with intermediate simulated nonlinearity units on 1D data [7]. However, this required photodetection and reinjection, and it did not involve convolutional layers.

REFERENCES

[1] Huaijin G Chen, Suren Jayasuriya, Jiyue Yang, Judy Stephen, Sriram Sivaramakrishnan, Ashok Veeraraghavan, and Alyosha Molnar. 2016. ASP vision: Optically computing the first layer of convolutional neural networks using angle sensitive pixels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 903–912.

[2] Song Han, Huizi Mao, and William J Dally. 2015. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *arXiv preprint arXiv:1510.00149* (2015).

[3] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and¡ 0.5 MB model size. *arXiv preprint arXiv:1602.07360* (2016).

[4] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature* 521, 7553 (2015), 436–444.

[5] Mikael C Rechtsman, Julia M Zeuner, Yonatan Plotnik, Yaakov Lumer, M Segev, and A Szameit. 2013. Photonic Floquet topological insulators. In *Lasers and Electro-Optics (CLEO), 2013 Conference on*. IEEE, 1–2.

[6] Indu Saxena and Emile Fiesler. 1995. Adaptive multilayer optical neural network with optical thresholding. *Optical Engineering* 34, 8 (1995), 2435–2440.

[7] Yichen Shen, Nicholas C Harris, Scott Skirlo, Dirk Englund, and Marin Soljačić. 2017. Deep learning with coherent nanophotonic circuits. In *Photonics Society Summer Topical Meeting Series (SUM), 2017 IEEE*. IEEE, 189–190.

[8] Jie Sun, Erman Timurdogan, Ami Yaacobi, Ehsan Shah Hosseini, and Michael R Watts. 2013. Large-scale nanophotonic phased array. *Nature* 493, 7431 (2013), 195–199.