Frontiers in Optics / Laser Science © Optica Publishing Group 2021

High Throughput Multi-kernel Fourier Optic Classifier

Zibo Hu[1], Maria Solyanik-Gorgone[1], Shurui Li[2], Puneet Gupta[2], Volker J Sorger[1]

1.Department of Electrical and Computer Engineering, George Washington University, Washington DC, 20052, USA.

2.Department of Electrical and Computer Engineering, University of California, Los Angeles, California 90095, USA.

sorger@gwu.edu

Abstract: An optical Fourier Network is a system empowered by a passive Fourier transformation enabling convolutional neural networks. Here, we demonstrate high throughputs by parallelizing the system via multiple-input and multiple-kernel capability of this convolutional classifier.

1. Introduction

In recent years, neural networks have demonstrated to meet numerous demands in science and engineering. In particular, Convolutional Neural Networks (CNN) have shown higher image recognition accuracy than human brain. However, this extraordinary ability comes at a cost of computational complexity: $O(MNmn)$ for $M*N$ image, $m*n$ kernel dimensions. In a regular CNN, the convolution process often consumes most of the computational resources, see AlexNet [1]. Therefore, new task-specific hardware has been developed to elaborate heavy computational load, such as Graphics Processing Units (GPUs), Tensor Processing Units (TPUs) and Fiend Programmable Gate Arrays (FPGAs), to name a few. However, these electronic processors are rather power-hungry and have a relatively long initialization time (>20ms) to implement Neural Networks (NNs). Considering these, research on heterogeneous computing systems for NN and machine learning applications has been recently gaining momentum in the realm of computer science and optical engineering.

In our previous work, we have demonstrated an electronic-optical hybrid NN architecture with an optical 4f convolution layer [2]. We achieved a 10x faster (latency) optical convolution process compared to a state-of-art GPU for a given matrix size. Here, we further expand the parallelism by exhausting the nominal throughput with multi-inputs and multi-kernels processing paradigm.

2. Method

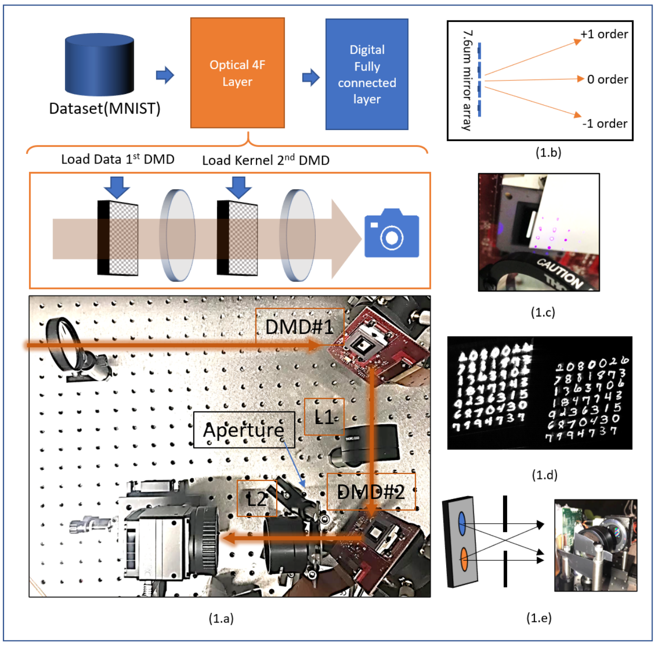
Our Optical Fourier Network is a CNN with an Optical Fourier layer instead of a digital convolutional layer, or so-called optical 4f system layer. Based on first principles, a thin lens performs a passive Fourier Transform (FT) in paraxial approximation. Hence, by placing a mask in the Fourier domain of a well-collimated coherent 4f laser system, one can acquire an ultra-fast and energy-efficient convolution processor. Alongside with FT coming at no energy cost, this approach allows to delay the domain crossing and perform one of the most computationally heavy tasks of image processing in optical domain.

In this experimental setup, we load the images into 1st Digital Micromirror Device (DMD) with 7x7 parallel inputs as the NN input, and 2 kernels into 2nd DMD - first layer of the heterogeneous classifier, Fig. 1. By passing through the 4f system, one gets an amplitude-only convolution of an images projected by 1st DMD with the kernels loaded on 2nd DMD. The resulting optical signal is captured by a CCD camera (XIMEA CB019MG-LX-X8G3) simultaneously acting as a non-linear quadratic activation function and optical-to-electronic domain converter. The algorithm performs a digital fully connected hidden layer for the classification downstream, (Fig. 1). In this architecture, Fig. 1b,1c, we exploit multiple diffraction orders coming from 1st DMD to perform multi-kernel convolution. In the experimental setup, we demonstrate 49x input channels processed in-parallel on 2x diffraction kernels, to achieve 98x parallel channels.

Currently used experimental setup is shown in Figure 1a. To reduce the diffraction gap and fit multiple diffraction orders into the 2nd DMD, we had to shrink the focal length f of the lenses and laser wavelength λ while using the same DMDs:

$$gap = \lambda/\Delta * f$$

were Δ is the DMD pitch.



Caption.1, The schematics of the hybrid Fourier Network: (1.a) a collimated laser beam (Thorlabs CPS450 450nm laser) is projected onto first DMD (TI DLP6500 DMD) with 7X7 matrix image. Lens L1 (AC508-100-A) converts the signal to a Fourier domain, where it gets convoluted with two kernels on second DMD. The output passes lens L2 (AC508-100-A) and gets captured by the camera (XIMEA CB019MG-LX-X8G3); (1.b) the DMD naturally generates multiple diffraction orders due to micro-mirrors acting as a grating; (1.c) multiple different diffraction orders generated by 1st DMD and projected onto 2nd DMD; (1.d) output from the camera for 49x inputs, 2x kernels;(1.e) The custom aperture blocks unwanted diffraction contamination form ghost images.

3. Result

The CNN classifier has been performed both in electronics and in Optical Fourier Network. The structure of the electronic NN exactly maps onto the corresponding structure of the opto-electronic classifier. First hidden layer is the convolution which, in the opto-electronic classifier, is performed by a 4f setup. The electro-optical conversion by the camera has been mimicked in electronics by a square activation function. Second hidden layer, the fully-connected layer, contains 500 neurons with Relu activation function, and in both systems is performed electronically. The resulting accuracies have been compared for two standard validation data sets: MNIST and CIFAR-10, see Table 1. We additionally analysed the stand-alone performance of the convolution layer by implementing a fully-connected-only classifier, that resulted in $\sim$2 times accuracy drop for CIFAR-10. The accuracy in Table 1 shows that the optical 4f layer has the comparable contribution as the convolution layer.

In our optical Fourier Network, due to a binary modulation, each pixel has depth of one bit, which is a major accuracy drop in comparison to a digital CNN. Additionally, exploiting two diffraction orders in Fourier domain comes at a price of a slight accuracy-to-speed trade-of due to uneven optical power distribution and difference in optical path. We modified the pre-trained kernels with a high-pass filter in the Fourier Domain to increase the misalignment tolerance, see Table 1.

Table 1. Experimental and reference digital network accuracy

Neural Network Structure	MNIST dataset	CIFAR-10 dataset
Digital single layer CNN	99%	50%
Digital Fully Connected only	92%	23%
49x input channel optical Fourier Network	**98%**	**45%**

4. Conclusion

In this work, we have demonstrated a combined 98x faster optical Fourier Network than previous work [2]; this was achieved by simultaneously realizing 49x parallel data input channels and 2 parallel kernel channels. Henceforward, further increase in parallelism can be achieved by adapting shorter wavelength laser, larger size micromirrors and shorter focal length lenses. This idea can expand the channels for many optical neural networks without additional hardware cost.

References

1. Xiaqing Li, Guangyan Zhang, H Howie Huang, Zhufan Wang, and Weimin Zheng. Performance analysis of gpu-based convolutional neural networks. In *2016 45th International conference on parallel processing (ICPP)*, pages 67–76. IEEE, 2016.
2. Mario Miscuglio, Zibo Hu, Shurui Li, Jonathan K George, Roberto Capanna, Hamed Dalir, Philippe M Bardet, Puneet Gupta, and Volker J Sorger. Massively parallel amplitude-only fourier neural network. *Optica*, 7(12):1812–1819, 2020.