Million-channel parallelism Fourier-optic convolutional filter and neural network processor

Mario Miscuglio1, Zibo Hu1, Shurui Li2, Jiaqi Gu3, Aydin Babakhani2, Puneet Gupta3, Chee-Wei Wong2, David Pan2, Seth Bank5, Hamed Dalir4, Volker J. Sorger 1,*

1Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA
2Department of Electrical and Computer Engineering, University of California Los Angeles, Los Angeles, CA, 90095, USA
3Department of Electrical and Computer Engineering, University of Texas, Austin, TX 78758, USA
4Omega Optics, Inc. 8500 Shoal Creek Blvd., Bldg. 4, Suite 200, Austin, Texas 78757, USA
*Correspondence: sorger@gwu.edu

Abstract: Here we report on a massively-parallel Fourier-optics convolutional processor accelerated 160x over spatial-light-modulators using digital-mirror-display technology as input and kernel. Testing the system on MNIST and CIFAR-10 datasets shows 96% and 54% accuracy, respectively. © 2020 The Author(s)

1. Introduction and Rationale
The rise of machine-learning has shown significant utility in i) deep-learning such as image and language classification, ii) nonlinear optimization such as in predictive control including target tracking, and iii) generative adversary networks. While electronic implementations provide utility, they often do not scale well in terms of inference delay or power consumption; for instance, a single game of AlphaGo Zero costs about $3,000 in electricity, and transistor scaling does not improve performance. Within the trend of hardware-specific accelerators driving a trend in computing heterogeneity, we envision future accelerators to include photonic process units (PPU), which utilize 1) one-shot (non-iterative) executions towards delivering \( O(1) \) process capability; 2) massive \( 10^9 \) parallelism such as form free-space digital light processing technology; 3) 'cheap' convolutions enabled by a ‘free’ Fourier transform performed by a lens; and 4) maturing foundry PDKs of high-performance photonic integrated circuits (PIC) [1-4]. Here we show a prototype of a Fourier-optics enabled 4f systems utilizing digital-mirror-display (DMD) technology for convolutional filtering and convolutional neural networks (CNN) (Fig. 1a).

2. Results and Discussion
The main idea is to perform massively parallelized vector matrix multiplications (VMM) in the Fourier domain as point-wise multiplications reducing the elsewise (e.g. GPU) \( O(n^2) \) multiplications to \( O(n) \); that is, any multiplication can be simplified by summing the digits of each \( n \)-digit factor and then multiplying. Furthermore, the in electronics costly Fourier transformation is here performed passively by a lens of this 4f Fourier processor (Fig. 1b), where in the Fourier-domain spatial frequency filtering is performed (kernel). However, unlike existing light modulators (SLM), which clock at 60 Hz rely on slow liquid crystal technology, here we deploy fast DMDs with the same spatial resolution but 10 kHz fast rates (~160x kernel speed-up, Fig. 1c). However, since DMDs are phase insensitive, phase information is not considered, but could be accounted for in a Vanderlugt lens system by sacrificing portion spatial resolution.

We built such a dual DMD system where the first DMD loads the image (or signal) and the second DMD performs the amplitude-based filter (Fig. 1c). As a test we loaded an image consisting of a differently rotated bar test-pattern into the processor. Then changing the kernel DMD, by rotating an exemplary (and ad-hoc selected) edge-detection kernel, we show that individual bar patterns can be selected. Such frequency processing can further be performed not only with images but also with bit-strings, RF-signals, or any other signal encoded in the optical domain.

In order to gain further insights into the accuracy resolution losses and performance of this 4f Fourier processor, we a) developed a physical accurate model of the 4f diffraction-based filtering of the 4f system, and b) perform offline kernel training and use the trained weights as Fourier kernels. Regarding the former, we develop a model that considers the phase-noise, limited numerical aperture, effects originating from aberration and diffraction of the lenses, and an accurate transfer function of the DMDs, which include dead-zones of the pixels of the DMD. Using this accurate description of our physical 4f system, we can now turn to the many offline training procedures in the field of machine-learning to train the kernels accurately (Fig. 1d). In brief a select an application-oriented open dataset (here MNIST, CIFAR-10) and perform a relatively standard gradient descent back-propagation-based training algorithm. For this, the network structure consists of 1x Fourier convolution layer, 1x fully connected (FC) layer (128-10) and use the ADAM optimizer. For the Fourier Conv. layer, the kernel is initialized with real numbers matching the
We acknowledge support from the Office of Navy Research under award number N00014-19-1-2595 of the Electronic Warfare program.