Learning in Holographic Convolutional Neural Nets

Kelvin Wagner
Optoelectronic Computing Systems Center, Dept of ECEE, University of Colorado, Boulder, CO, USA

Abstract: We show how the array of convolutional adaptive interconnections needed for deep learning can be physically implemented and learned in an all-optical multistage dynamic holographically-interconnected architecture using thick Fourier-plane dynamic holographic photorefractive crystals. This optical architecture is self-aligned, phase-calibrated, and aberration compensated by using photorefractive phase-conjugate mirrors to record the dynamic-holographic Fourier-plane interconnections in each layer.

Van Heerden introduced the metaphorical connection between neural computation and volume holography nearly 60 years ago. [1] A variety of holographic interconnection approaches to building optical neural networks were investigated in the 80s and 90s. [2–4] Since these planar technologies have the same component density scaling as 2-D electronic neural networks numerous optical researchers instead investigated the use of 3-D photorefractive (PR) crystals for neural-network weighted interconnections. PR crystals were used both as nonlinear dynamical systems with emergent neural computational capabilities [5–7] as well as the weighted adaptive holographic interconnections in optical neural networks. [8, 9] These 3-D volume-holographic interconnections have a potential for a dramatically larger interconnection weights storage capacity ($10^{11}/\text{cm}^3$) than planar integrated-electronic or integrated-optical approaches. These adaptive holographic systems can be made to self align by using phase-conjugate mirrors. [10–12] Multi-layer optical networks using cascaded holograms and trained by back propagation were also investigated. [13–17] The critical missing ingredient in developing an all-optical multi-layer adaptive neural network is a practical nonlinear neuron. Recent quantum optic and integrated-photonic programmable activation function demonstrations have shown capability for a variety of flexible nonlinearities and indicated capability for digitally controlled or simulated learning. [18–21] But none of these approaches have addressed the critical requirement for appropriate bidirectional gradient operation through the nonlinearity needed for backwards error propagation, so we introduce in Fig. 1 a novel bidirectional optical ReLU (Rectifying Linear Unit), [22] the modern nonlinear neuron activation function that enables deep learning by avoiding the sigmoidb derivative blocking of back propagation errors from penetrating deep into the trainable multi-layer network.

Various approaches to a front-end convolutional layer and optical convolutional neural networks (CNNs) have been proposed and demonstrated as single layers or in simulations, [21, 23, 24] but have not addressed the key components of modern deep learning CNNs including multiple feature planes, variable template sizes, pooling and resolution decrease, and back propagation learning of the shared weights that we incorporate. [22] Adjoint method training using concepts of numerical phase conjugation has been developed for neural network and integrated optics optimization but without an appropriate differentiable nonlinearity multiple layers and convolutional operation has not been achieved. [25] An integrated-photonic approach utilizing a thin-film photorefractive adaptive holographically-interconnected vector-matrix-multiplier is being developed [26], but so far only single layers without back propagation, with a limit of about a million weights, and without the capability for multiple convolutional feature planes. We instead exploit the massive parallelism of bulk photorefractive holographic optical systems with the potential for billions of adaptive weights per layer (corresponding to 10^{11} convolutional multiply-and-adds per layer with scores of input and output feature planes). This approach is based on deep learning in a multi-layer architecture consisting of spatially-multiplexed arrays of optical rectifying neurons convolutionally interconnected by arrays of adaptive Fourier-holographic weights as in Figs. 2 and 3. By using photorefractive phase-conjugate mirrors for the adaptive holographic recording, this system avoids problems with alignment, aberrations, and device imperfections. The optical ReLU device enables practical, low-power, multi-layer

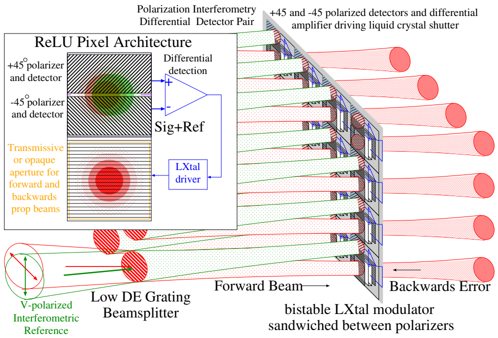

Figure 1: Bidirection optical ReLUs using polarization interferometry in an LCoS smart-pixel array geometry.

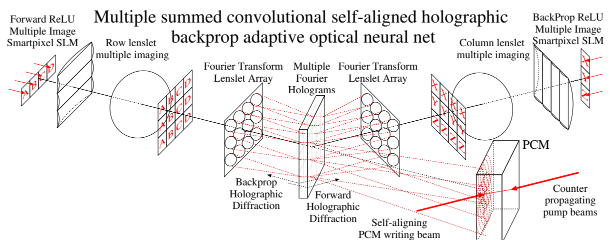

Figure 2: Multiple feature plane convolutions using lenslets and self-aligning Fourier volume hologram.

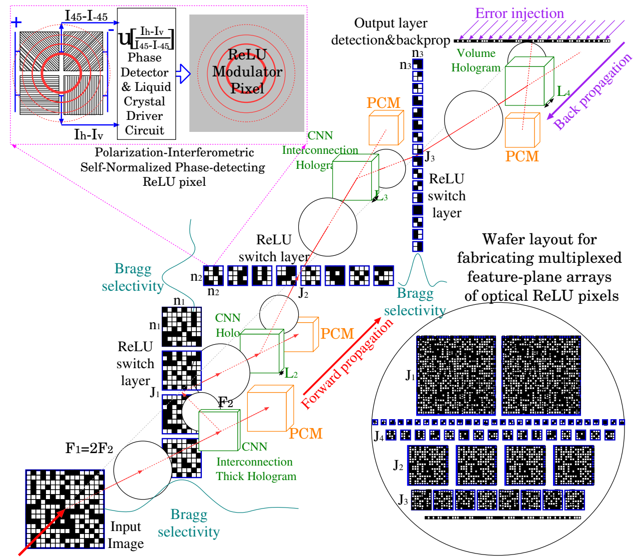

Figure 3: Multiple layers of self-aligning, angular-multiplexed, convolutional Fourier holograms using optical ReLU neuron layers and PR crystal interconnects.

neural-network cascades while enabling fully-reciprocal bidirectional operation for back propagation based adaptive holographic learning. When built, this system will be capable of achieving a competitive computational throughput to special purpose deep learning supercomputers with only moderate speed neurons and slowly evolving photorefractive interconnection convolutional feature plane weights.

References

[1] P. J. van Heerden, "Optical information storage in solids," *App. Opt.*, vol. 2, p. 393, 1963.
[2] E. G. Paek et al, "Optical associative memory using FT holograms," *Opt. Eng.*, vol. 26, p. 428, 1987.
[3] D. Psaltis et al, "Holography in artificial neural networks," *Nature*, vol. 343, p. 325, 1990.
[4] A.V. Pavlov et al, "Holographic correlators and optical neural networks," *J Opt Tech*, vol. 61, p. 42, 1994.
[5] D.Z. Anderson et al, "Resonator memories and optical novelty filters," *Opt. Eng.*, vol. 26(5), p. 434, 1987.
[6] Y. Owechko, "Optoelectronic resonator neural networks," *App. Opt.*, vol. 26(23), p. 5104, 1987.
[7] D. Anderson, "Information dynamics of photorefractiove two-beam coupling," *PRL*, vol. 82, p. 1418, 1999.
[8] D.Psaltis, D.Brady, K.Wagner, "Adaptive optical networks using PR Xtals," *AO*, vol. 27, p. 1752, 1988.
[9] K. Wagner, "Optical Competitive Learning With LXtal WTA Modulators," *AO*, vol. 32, p. 1408, 1993.
[10] B.H. Soffer et al, "Associative holographic memory with feedback using PCM," *Opt. Lett.*, vol. 11(2), 1986.
[11] Y. Owechko, "Optical interconnection for neural networks using SP-PCM," *Opt. Let.*, vol. 16, p. 675, 1991.
[12] I.M. Beldyugin et al, "Optical neurocomputers based on PR Xtals," *Kvan. Elek.*, vol. 19, p. 424, 1992.
[13] K. Wagner and D. Psaltis, "Multilayer optical learning networks," *Appl. Opt.*, vol. 26, p. 5061, 1987.
[14] I. Saxena et al, "Adaptive multilayer optical neural-network," *Opt. Eng.*, vol. 34, p. 2435, 1995.
[15] D. Psaltis and Y. Qiao, "Adaptive multilayer optical networks," *Prog. Opt.*, vol. 31, p. 227, 1993.
[16] Y. Owechko, "Cascaded-grating holography for artificial neural nets," *App. Opt.*, vol. 32, p. 1380, 1993.
[17] G. C. Petrisor, "Backward-error-propagation learning in PR crystals," *App. Opt.*, vol. 35, p. 1328, 1996.
[18] A. Ryou et al, "Free-space optical NN based on atomic nonlinearity," *Phot Res*, vol. 9(4), p. B128, 2021.
[19] I. Williamson et al, "Reprogrammable EO Nonlinear Activation Functions," *IEEE JSTQE*, vol. 26, 2020.
[20] M. Frad, "Experimental realization of arbitrary activation functions ," *Opt Exp*, vol. 28, p. 12138, 2020.
[21] N. Passalis et al, "Training Photonic CNN With Sinusoidal Act.," *IEEE T Comp Intel*, vol. 5, p. 384, 2021.
[22] K. Wagner, "Optical ReLUs for Back-Prop Learning in Deep Holographic CNN," *IEEE JSTQE v26*, .2020
[23] S. Colburn, "Optical frontend for a convolutions neural network," *App. Opt.*, vol. 58(12), p. 3179, 2019.
[24] B. Muminov et al, "Fourier optical preprocessing in lieu of deep learning," *Optica*, vol. 7(9), p. 1079, 2020.
[25] T. Hughes, "Training of photonic NN through in situ backpropagation ," *Optica*, vol. 5(7), p. 864, 2018.
[26] R. Dangel et al, "Integrated photonic synaptic interconnect processor," in *APC*, p. IM2A.1, OSA, 2020.