Research Statement

Sudhanshu Shrivastava

Department of Physics, Indian Institute of Technology (IIT-BHU), Varanasi 221005, India

In the past decade, Machine learning has been the reason behind the revolutions in many fields. Deep learning, especially, has been making real strides due to the advancements in computational technology. Deep learning has produced remarkable results on perceptual and automation problems such as seeing, hearing, and driving—problems involving skills that seem natural and intuitive to humans but have long been elusive for machines. Besides these popular applications, Deep learning has led to significant advancements in the field of computational imaging and digital image processing.

My research interests lie in the broader field of Machine Learning and specifically in the intersection of deep learning, physics, computer vision, computational imaging and digital image processing. I have been intrigued by the application of Deep Learning to physical problems, especially those involving the application of all optical systems, mostly diffractive layers, to implement deep learning algorithms for holographic reconstruction, phase recovery, computational imaging, and classification [1-3]. Currently, I am working on my UG Honours Project under the supervision of Dr. Rakesh Kumar Singh*, and co-supervision of Dr. Rajeev Singh* on "Physics informed Machine Learning".

As a first step, I attempted to reproduce the results of the paper "Computer-Free, All-Optical Reconstruction of Holograms Using Diffractive Networks." This paper proposes a compact set of diffractive layers network that can all optically process Gabor inline holograms to give the reconstructed image. The proposed conceptual framework achieves twin artifact free reconstruction of Gabor inline holograms[4] through light-matter interactions facilitating speed of light passive reconstruction as opposed to iterative phase retrieval techniques[5-11] that involve time and electrical power consuming digital processors.

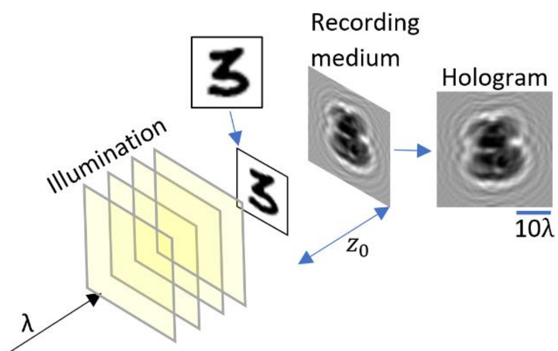

Figure 1 Inline Hologram Recording,
(contains graphical representation of numerical simulations that I performed)

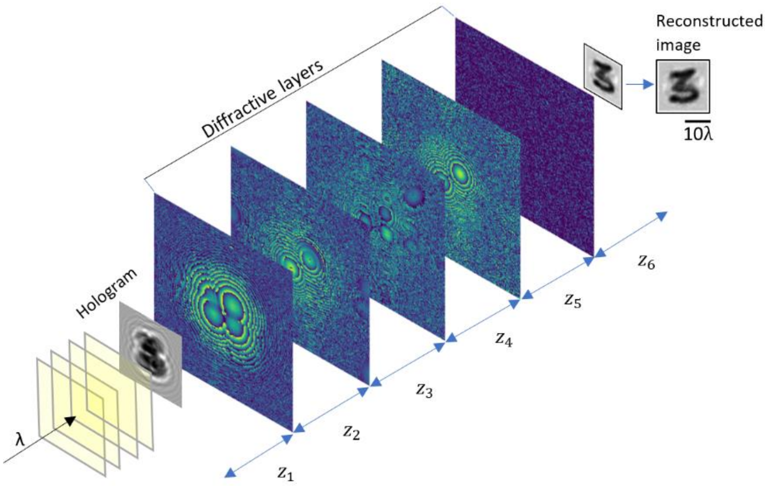

Figure 2: All-optical reconstruction process.
(contains graphical representation of numerical simulations that I performed)

In Figure 1, an unknown object is illuminated with a coherent plane wave of wavelength λ and the pattern produced by the interference of the scattered wave and directly transmitted wave is recorded on a recording medium place at a distance z_0 from the object plane. For free space holographic reconstruction, the recording medium i.e., photographic emulsion or the optoelectrical sensor array (in numerical simulation) is illuminated with reference plane wave and the resulting wavefield is recorded on a screen at distance z_0. This reconstructed image contains the overlapping virtual image of the object, known as twin artifact. To eliminate this twin artifact, as per the all-optical method proposed in the paper, in Figure 2, the light emanated after illuminating the hologram recording medium with the reference wave is allowed to pass through a set of transmissive, spatially engineered set of diffractive layers and a twin artifact-free reconstructed image is obtained on the screen. The size of the object being recorded, diffractive layers, and hologram are assumed to be 25λ, 100λ, and 42λ, respectively with the axial separation between the transverse planes containing— the input hologram, diffractive layers and the output plane is taken to be z_i, $\mathrm{i} = 1, 2, \ldots N+1$. N is the number of diffractive layers. In the paper and my attempt, the operating wavelength(λ) is taken to be 600 nm and $\mathrm{N} = 5$ diffractive layers have been taken, each with 200×200 trainable parameters, each parameter corresponds to the phase value of a diffractive feature which acts as a neuron.

Skip to results

* *Assistant Professor, IIT BHU*

The Rayleigh Sommerfeld formulation can be written as follows for the forward pass of scattered light through the diffractive network: -

$$u_l^k = \sum_m \frac{z_l}{\left(r_{mk}^l\right)^2}\left(\frac{1}{2\pi r_{mk}^l} + \frac{1}{j\lambda}\right)\exp\left(j\frac{2\pi r_{mk}^l}{\lambda}\right)v_m^{l-1}$$

$$r_{mk}^l = \sqrt{(x_k - x_m)^2 + (y_k - y_m)^2 + z_l^2}$$

$$v_m^l = \left(\sum_m w_{km}^l v_m^{l-1}\right).t_k^l$$

$$w_{km}^l = \frac{z_l}{\left(r_{mk}^l\right)^2}\left(\frac{1}{2\pi r_{mk}^l} + \frac{1}{j\lambda}\right)\exp\left(j\frac{2\pi r_{mk}^l}{\lambda}\right)$$

Where u_m^l , v_m^l respectively represent the complex amplitude of the optical wave before and after optical modulation by the feature m of the diffractive layer l, $l = 1, 2, \dots, N$. $l = 0$ and $l = N + 1$ are the index for the input hologram plane and the reconstruction plane. Here λ is the wavelength of the optical wave; z_l is the axial distance between layers $l - 1$ and l; $(x_m, y_m)$and (x_k, y_k) are the transverse coordinates of feature m of layer l − 1 and feature k of layer l, respectively; v^0 and u^{L+1} are the optical fields at the input and the output fields-of-view of the diffractive network, respectively. Also, $v_k^l = t_k^l.u_k^l$ can be written by the definition of t, under the assumption the that thin diffractive layers have been used. t_k^l is the complex valued transmittance of pixel m of layer l.

Drawing parallels between the above equation and the equation for a feedforward neural network, $w_{km}^l$and t_k^l are analogous to the weights and multiplicative biases respectively, with t_k^l being trainable since for thin diffractive layers, phase φ_k^l of t_k^l is related to height of a feature by $\varphi_k^l = \frac{2\pi(n-1)h_k^l}{\lambda}$. Since in the proposed design, phase only modulation diffractive layers have been used, the amplitude of t_k^l is taken to be 1, hence, $t_k^l = \exp(j\,\varphi_k^l)$. For the numerical simulation of diffraction of wavefield by the diffractive layer, I defined **a custom layer (diffractive layer)** of shape (200,200) in TensorFlow 2.9.1[12] by subclassing tf.keras.layers, and defined its forward pass as $v_k^l = \exp(j\,\varphi_k^l).u_k^l$, u_k^l being the input to the layer l, obtained by propagating the output wavefield matrix v_{l-1} of the $l-1$ layer using the Angular Spectrum Method for the propagation of wavefields between traverse planes. The angular spectrum of the wavefield to be propagated at the traverse plane $i-1$ be given as $G(p, q, z_{i-1})$,

Then,

$$G(p,\ q;\ z_{i-1}) = \mathcal{F}\{v(x, y; z_{i-1})\}$$

$$G\ (p, q; z_i) = \begin{cases} G(p, q; z_{i-1})\exp(w(p,\ q)d), & \lambda^{-2} - p^2 - q^2 \geq 0 \\ 0 & ,otherwise \end{cases}$$

$$w(p,\ q) = \sqrt{\lambda^{-2} - p^2 - q^2},$$

w, p, q represent Fourier frequencies of a wavefield

$$u(x, y; z_i) = \mathcal{F}^{-1}\{G\ (x, y; z_i)\}$$

I defined a **custom AngularSpectrumPropagator layer** and used fft2d, ifft2d, fftshift functions from tf.signal module in its call for the numerical implementation of the above equations. For the hologram recording process, I built a **custom data pipeline** in which I first resized the 28×28 EMNIST images to 50×50 using bilinear interpolation and then used adequate normalizing and padding to get matrix(y) of shape (200,200) which contain the object of (50,50), with values between 0 and 1, embedded in (200, 200) matrix of values 1. Then the wavefield resulting from the illumination of the object can be given by the above (200,200) matrix since a plane wave of unit amplitude is assumed to be used for illumination. I calculated the recorded hologram intensity matrix(x) by taking the square of amplitude of the wavefield resulting from the propagation of object wave field by distance 30λ using the Angular Spectrum Propagator layer. Similarly, in the all-optical reconstruction model, upon illumination of the hologram, the wavefield is given by x matrix, and I defined the output wavefields for 1st layer to be the result of the application of angular spectrum propagator and diffractive layer on the x matrix and applied a similar treatment to the subsequent layers, hence obtaining the wavefield on the output plane. The square of the amplitude of this complex wavefield obtained gave the reconstructed image matrix ($\hat{y}$). The defined **custom loss function** to implement the loss as per the paper, $\mathcal{F}\{.\}$ represents Fourier transform operation $L = L_{pixel} + 1000L_{fourier} + \eta L_{efficiency}$

$$L_{pixel} = \frac{1}{N_p}\sum_{p=1}^{N_p}|y_p - \hat{y}_p|$$

$$L_{fourier} = \frac{1}{N_p}\sum_{p=1}^{N_p}\left|\mathcal{F}\{y\}_p - \mathcal{F}\{\hat{y}\}_p\right|^2$$

$$L_{efficiency} = 1 - \frac{P_{out}}{P_{illum}}$$

I defined P_{out} to be the reduced mean of $\hat{y}$ and $P_{illum} = 1$, y_p and $\hat{y}_p$ are the elements of y and $\hat{y}$ respectively. I trained the model using **distributed training** on 2 GPUs of the **HPC cluster PARAM Shivay** taking penalty term(η) be 0.5, δ_{tr}(*Figue 4(c)*) *to be* 0.2 and mini-batch size 8.

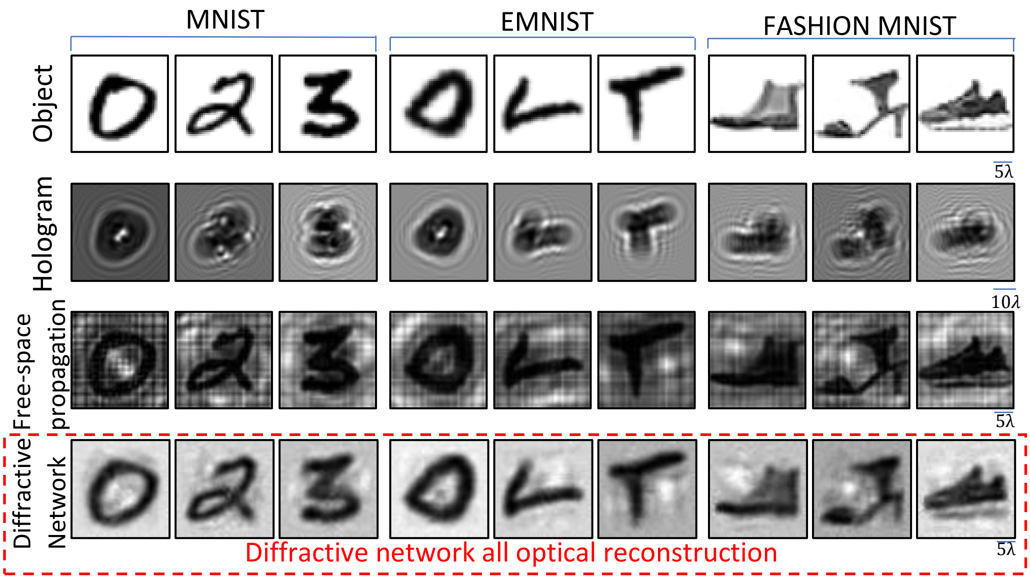

Figure 3. A graphical representation of results of numerical simulations that I performed, the first row shows the objects, second row shows the corresponding amplitude encoded holograms of objects in the first row that I obtained by propagating the wavefield 30λ using the Angular Spectrum method. The fourth row shows the free space propagation reconstruction of holograms by propagating the wavefield at hologram plane, resulting from illumination by reference wave, by 30λ. The last row shows the result of all optical reconstruction process using diffractive layers. I implemented the Angular Spectrum method of wave propagation by defining a custom keras layer, with non-trainable weights, called AngularspectrumPropagator. I defined another custom layer, with 200,000 trainable parameters, called DiffractiveLayer that numerically mimics the action of diffractive layer on a wavefield. I defined diffractive network as the sequential model consisting of 5 AngularspectrumPropagator and DiffractiveLayer layers.

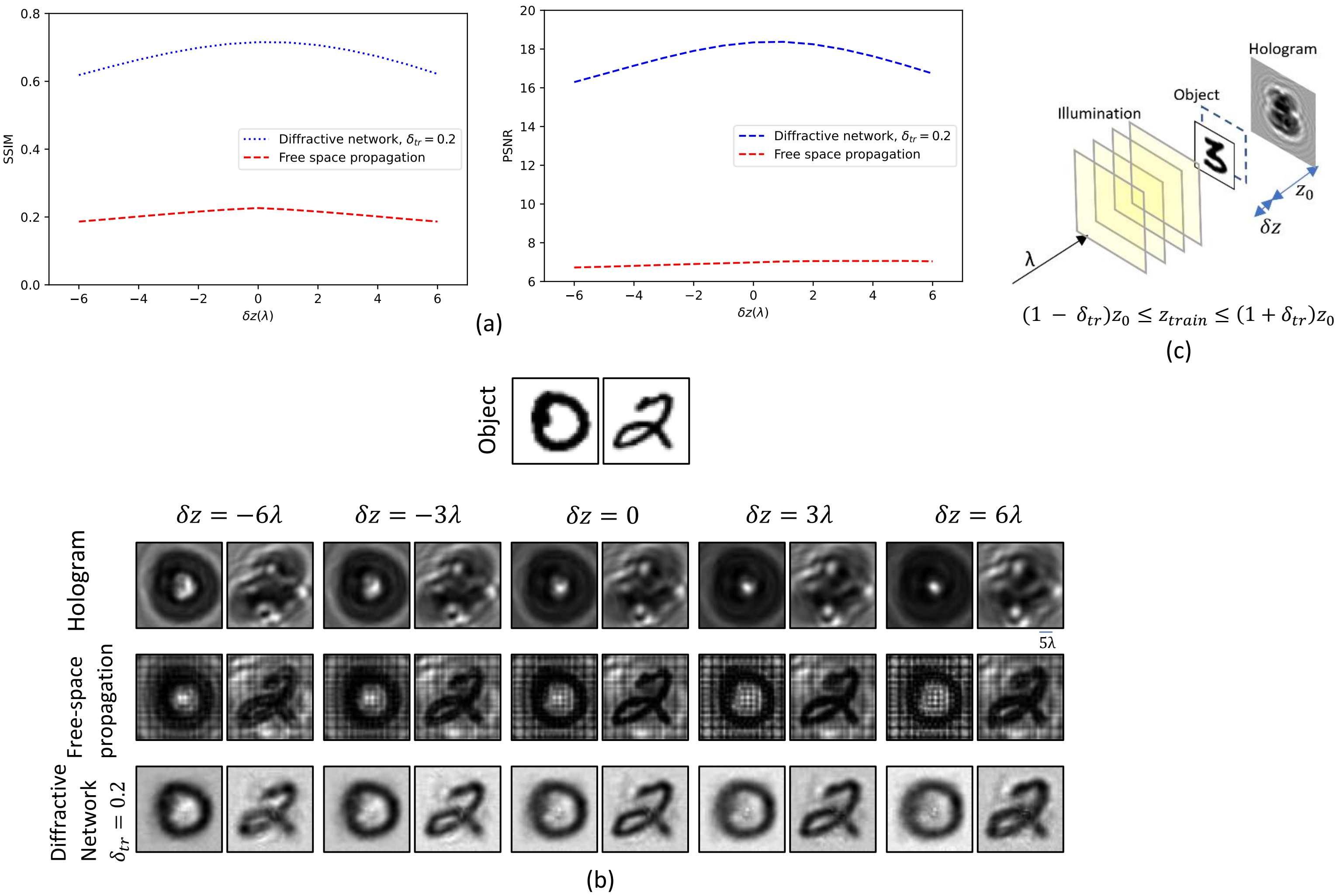

Figure 4. The trained diffractive network is robust to unforeseen changes to hologram recording distance($z_{test} = z_0 + \delta z$)(c). I trained the network with a hologram recording distance ($z_{train} = z_0 + \delta z$), δz is a discrete random variable distributed that takes values integral multiple of λ from -6λ to 6λ with equal probability. This treatment of z_{train} as a random variable vaccinates the network to variations in the value of hologram recording distance due to unavoidable alignment errors during the actual physical realization. I evaluated the average values of SSIM and PSNR over 10,000 test images from MNIST dataset for different values of δz and plotted this variation for free-space propagation and all- optical diffractive network reconstruction in (a). I have given a graphical representation of this robustness to changes in hologram recording distance in (b).

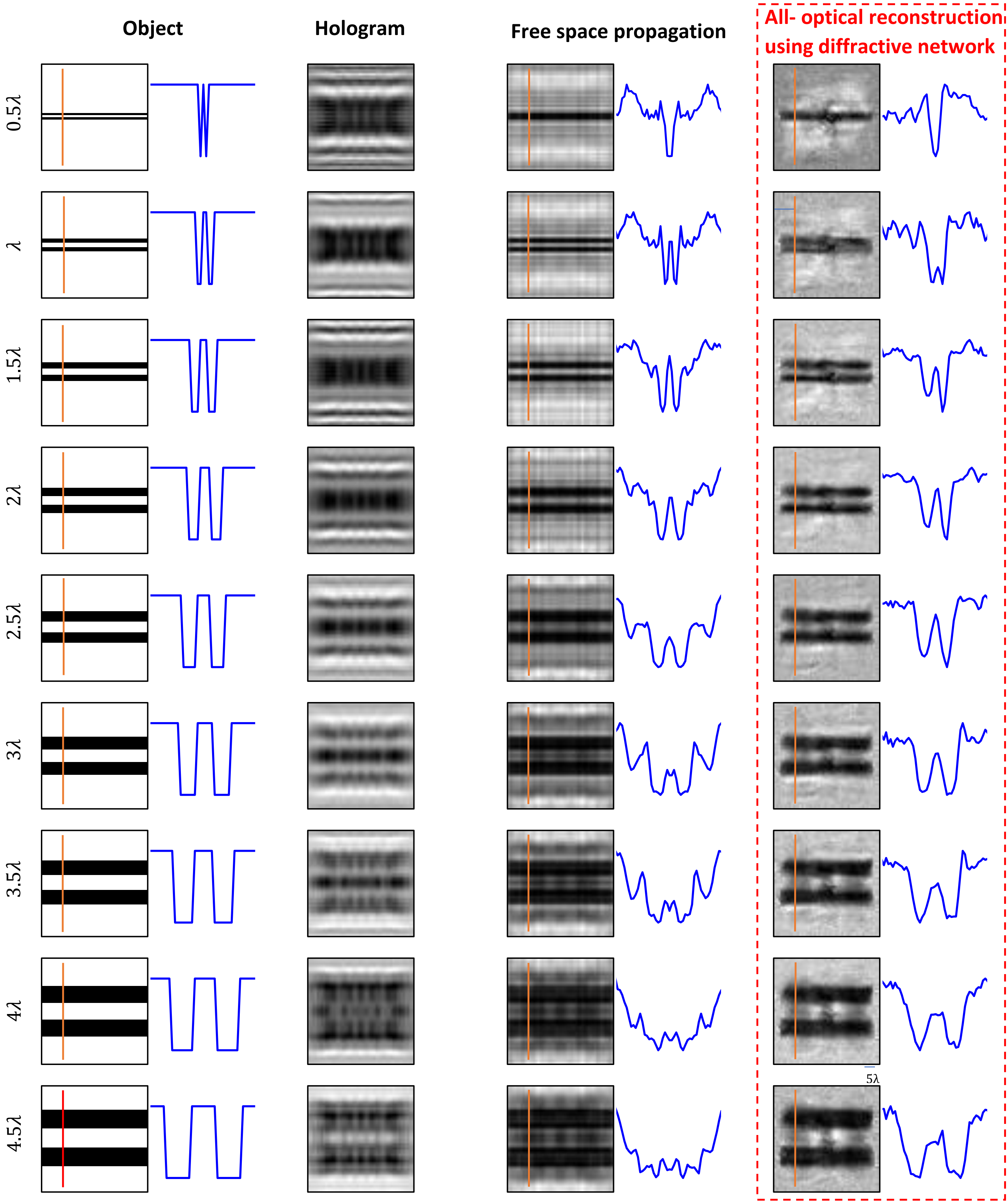

Figure 5 depicts the ability of the diffractive network to resolvable reconstructed images of inline holograms of objects with separations of order of magnitude similar to λ, in comparison to that of free-space propagation reconstruction. The variation of intensity value of an image along the orange vertical line is plotted alongside it. The diffractive network is able to resolve two lines separated by a distance greater than λ. The performance of the diffractive network in comparison to free-space reconstruction is especially appreciable when the lines separation is larger than 3.5λ.

The hallmark of my studies in the field of Deep Learning has been the application of D^2NN, to implement different DNN all-optically. I intend to use D^2NN to be able to optically implement computationally modelled solutions to complex deep learning problems to pioneer advancements in medical imaging, image processing, pattern recognition, image classification and deep learning in general, to exploit the light-speed, passive, highly scalable and easily reconfigurable nature of D^2NN. I would like to work towards the eventual realization of backpropagation algorithms through physical processes like light-matter interactions, and hence toward the generalization of D^2NN as a complete Deep Learning solution.

References

[1] *Md Sadman Sakib Rahman and Aydogan Ozcan, Computer-Free, All-Optical Reconstruction of Holograms Using Diffractive Networks,* ACS Photonics ***2021*** 8 *(11), 3375-3384, DOI: 10.1021/acsphotonics.1c01365*

[2] *Mengu, D., & Ozcan, A. (2022). Diffractive all-optical computing for quantitative phase imaging. arXiv preprint arXiv:2201.08964.*

[3] *Mengu, Deniz, et al. "Analysis of diffractive optical neural networks and their integration with electronic neural networks." IEEE Journal of Selected Topics in Quantum Electronics 26.1 (2019): 1-14.*

[4] *Gabor, Dennis. "A new microscopic principle." nature 161 (1948): 777-778.*

[5] *Teague, M. R. Deterministic Phase Retrieval: A Green's Function Solution. J. Opt. Soc. Am. 1983, 73 (11), 1434−1441.*

[6] *Streibl, N. Phase Imaging by the Transport Equation of Intensity. Opt. Commun. 1984, 49 (1), 6−10.*

[7] *Latychevskaia, T.; Fink, H.-W. Solution to the Twin Image Problem in Holography. Phys. Rev. Lett. 2007, 98 (23), 233901.*

[8] *Waller, L.; Tian, L.; Barbastathis, G. Transport of Intensity Phase-Amplitude Imaging with Higher Order Intensity Derivatives. Opt. Express 2010, 18 (12), 12552−12561.*

[9] *Greenbaum, A.; Ozcan, A. Maskless Imaging of Dense Samples Using Pixel Super-Resolution Based Multi-Height Lensfree on-Chip Microscopy. Opt. Express 2012, 20 (3), 3129−3143.*

[10] *Latychevskaia, T.; Fink, H.-W. Practical Algorithms for Simulation and Reconstruction of Digital In-Line Holograms. Appl. Opt. 2015, 54 (9), 2424−2434*

[11] *Rivenson, Y.; Wu, Y.; Wang, H.; Zhang, Y.; Feizi, A.; Ozcan, A. Sparsity-Based Multi-Height Phase Recovery in Holographic Microscopy. Sci. Rep. 2016, 6 (1), 37862*

[12] *Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; Kudlur, M.; Levenberg, J.; Monga, R.; Moore, S.; Murray, D. G.; Steiner, B.; Tucker, P.; Vasudevan, V.; Warden, P.; Wicke, M.; Yu, Y.; Zheng, X. TensorFlow: A System for Large-Scale Machine Learning. Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation; OSDI'16; USENIX Association: U.S.A., 2016; pp 265−283*