Hashing for Secure Optical Information Compression in a Heterogeneous Convolutional Neural Network

Maria Solyanik-Gorgone¹, Behrouz Movahhed¹, ², and Volker J Sorger¹, ² *

¹Department of Electrical and Computer Engineering, George Washington University, Washington DC, 20052
²Optelligence LLC, 10703 Marlboro Pike, Upper Marlboro, MD, 20772, USA
*sorger@gwu.edu, @optelligence.co

ABSTRACT

In the recent years, heterogeneous machine learning accelerators have become of significant interest in science, engineering and industry. The major processing speed bottlenecks in these platforms come from (a) an electronic data interconnect; (b) an electro-optical interface update rate. In this light, information compression implemented in native to incoming data optical domain could mitigate both problems mentioned above by reducing the demand on data throughput at the camera side and beyond. In this paper we present an optical hashing and compression scheme that is based on SWIFFT - a post-quantum hashing family of algorithms. High degree optical hardware-to-algorithm homomorphism allows to optimally harvest well-understood potential of free-space processing: innate parallelism, low latency tensor by-element multiplication and Fourier transform. The algorithm can provide several orders of magnitude increase in processing speed by replacing slow high-resolution CMOS cameras with ultra fast and signal-triggered CMOS detector arrays. Additionally, the information acquired in this way will require much lower transmission throughput, less in silico processing power, storage, and will be pre-hashed facilitating cheap optical information security. This technology has a potential to allow heterogeneous convolutional 4f classifiers get closer in performance to their fully electronic counterparts.

1 Introduction

Optical compression methods have been extensively studied for decades inspired by bandwidth, storage and processing power hungry real-time image recognition and holography.1–3. Recently, the transformative advances in technology such as internet of things, edge-computing and 3D visualisation with light-field cameras posed new challenges, such as algorithms capable of real-time image pre-processing, compression and pattern recognition that resulted into the emergence of a new field of compressive sensing.4, 5 Large image size and strong temporal correlations that distinguish multimedia data from other formats pose a great computational challenge due to convolutions and dot-product multiplications being the main component of image processing ML,6, 7 and simultaneously the main computer power consumer.

In most currently deployed heterogeneous implementations, optical data processing, encryption and compression are accomplished electronically8–10, hence, after the electro-optical domain conversion. Often, image quality reduction is implemented first to alleviate corresponding processing load. This way, the benefits of having a high-resolution camera in free-space setting are not fully harvested due to the digital image processing in order to meet electronic streaming throughput and/or data storage limitations. Alongside with that, other venues for development recently have been gaining momentum such as pre-processing in photonic integrated circuits11, custom CMOS chips for optical encryption12, and extra-low formfactor metasurface-based optical systems13. These shows multi-directional and cross-disciplinary interest in novel hardware solutions for high-speed data processing and security.

Explored here, Fourier optics based heterogeneous data compression and encryption is by far not a new idea14–20. Free-space 4f-systems have been previously considered for optical encryption in phase-only modulating setups, oftentimes employing random phase filters14, 17. Exotic proposals include approximating digital fractional Fourier transform algorithms in Fourier optics18. However, to our knowledge, there has been no attempt to (1) directly adopt post-quantum encryption protocols in the field of optical processing; (3) simulate simultaneous amplitude-and-phase modulation filtering in Fourier domain to achieve confusion step in heterogeneous hashing; and (3) achieve simultaneous optical hashing and compression.

Fourier 4f-system-based21 simultaneous optical input hashing and compression system presented here elegantly addresses multiple drawbacks of optical processing: (a) feature extraction; (b) feature hashing; (c) data compression; (d) data security. As opposed to compact and fast on-chip solutions11, throughput of the featured optical compression system is limited only by the resolution of the light modulating device and diffraction properties of light, while phase control can be achieved via
Figure 1. The concept of the SWIFFT-based heterogeneous (optical) hashing/compression system and its integration into a generic edge system data interconnect. The advantages of (1) harvesting the features of the input in analogue domain, hence, before introducing analogue-to-digital conversion related noise; (2) utilising low-latency and high-speed convolutions and matrix multiplication in optical domain; (3) conventional free-space alignment methods. The theoretical core of the design is based on one of the lattice-based post-quantum hashing algorithms and uses it as a prototype for the featured here implementation of heterogeneous 4f-based optical data hashing. SWIFFT family of FFT-based hashing functions considered here are not lossless, but asymptotically secure, highly parallelizable, collision resistant and algorithmically homomorphic to optical processing. The underlying class of linear algebra operations for lattice-based cryptography, matrix-vector multiplication, has shown the potential to be performed optically or in a photonic chip with projected efficiency exceeding currently existing fully-electronic accelerators, supplied high-speed domain conversion and electronic interconnect. With a thin lens being a passive Fourier transform engine, the algorithm has a potential to elegantly integrate into modern optical interconnect potentially providing an unprecedentedly fast additional layer of security.

The proposed optical setup can be realized on a set of spatial light modulators (SLMs) with employed full amplitude-phase convolution in Fourier domain of a 4f system. This ASIC optical co-processor is capable of hashing 2D arrays of binary data at kHz speed and a fundamental latency in optical domain of $1.49 \cdot 10^8$ s, limited by the speed of light in free space. Key practical limitations come from the update rate of (1) modulating devices; (2) image acquisition and processing hardware; (3) speed of the supporting electronic interconnect. First limitation, high-speed SLMs with kHz update rate that could be an optimal choice for this project, remain to be a major holdup in the field of optical processing currently limiting the achievable speed of heterogeneous co-processors by 120Hz (off-the-shelf). However, if compression speed is a priority as opposed to hash security, the requirement on SLMs’ update rate can be substantially relaxed. In the limiting case, SLMs can even be replaced with static masks not compromising compression quality. Second and third limitations can be mitigated by an order of magnitude in speed due to the one-dimensional optical output of the proposed scheme. That allows one to use ultra-high-speed sensor arrays (e.g. ~0.5MHz Linea ML detector arrays) as opposed to many earlier proposed architectures that require order of magnitude slower CCD/CMOS cameras at the domain crossing and ultra-high bandwidth electronic interconnect. The back-end processing can be accomplished via comparing to a hash-table or using ML algorithms depending on the optical system’s and information channel’s stability, and a particular application.
Figure 2. CNN architecture and training results for MNIST compression/hashing algorithm: (a) the used in the experiment CNN architecture comprising one convolution, one pooling and one fully-connected layer with softmax activation function; (b) the convolution layer at training and inference stages are proposed to be replaced by an optical 4f-convolution layer; (c) training metric (accuracy and loss) for the original MNIST dataset of 60,000 images training part and the corresponding hashed/compressed images; the 60,000 images are separated into 40,000 for training and 20,000 for validation; (d) study of lossiness of the hashing/compression function, Algorithm 2, where information loss exhibits power-low behavior as a function of the compression coefficient.

2 Results

Here we propose an optical compression/hashing system inspired by the SWIFFT family of post-quantum lattice-based hashing and encryption algorithms. This 4f-based free-space optical system accomplishes a two-step hashing and compression via matrix-matrix multiplication and convolution in Fourier domain. The samples of input versus output data from the algorithm, and the schematics of the proposed experimental setup is shown in Fig. 1. The envisioned optical setup consists of (a) a source, e.g. input delivered via fiber optics interconnect or an otherwise spatially modulated collimated laser beam; (b) SLM1 with a projected weight matrix modulating in amplitude-only regime; (c) complex modulation unit comprising two SLMs embedded in a Michelson-type setup. The output of this system will have a 1D-like appearance and spread longitudinally. The cylindrical lens positioned in front of the detector array DA accomplishes the final step of the compression. The sensor’s readout can subsequently be stored in a hash table and/or used as an input for a heterogeneous CNN classifier. Further technical details and the theoretical background are presented in Sec.4.

The optical free-space propagation part of the featured compression algorithm was simulated electronically with the account of diffraction optics effects and within the paraxial approximation. The simulated parameters of the optical setup in Fig. 3(b) are $\lambda = 633$nm is the light wavelength; the system’s focal distance of the lenses $L_1$ and $L_2$ depends on the dimensions of the input mask as $f = (mask \hspace{1mm} size, \hspace{1mm} nm) \times (pixel \hspace{1mm} size)/(wavelength, \hspace{1mm} nm)$. SLMs are assumed to have $2\pi$ pixel-to-pixel phase modulation range, 1910X1080 pixel resolution, and 8 bit pixel depth (e.g. PLUTO-2 from HOLOEYE).

The 60,000 hashed images generated from the training portion of MNIST dataset have been split and 40,000 have been used for training, and 20,000 classification in a 3-layer CNN: 2D convolution layer with 10 filters of size 3X3, maxpooling layer with $p = 2$ pooling parameter, and the dense layer to output 10 categories for numbers from 0 to 9, see Figure 2(a). Training has been performed using Keras library in Python environment. A new dataset of 60,000 hashed images has been generated using our compression algorithm and utilizing the training set of MNIST is the input. Both original and hash datasets were trained on the same model with the same set of parameters: 60 epochs of batch size 1. It is worth to mention, that the choice of these parameters as well as the CNN architecture did not matter for the experiment. The goal is not to show the performance of a particular ML model on the given hashed data set in absolute, but to juxtapose performance of any given CNN on original MNIST data versus the hashed MNIST data.
Figure 3. The concept of optical hashing; (a) an example of a 1D compressed sequence, read of the simulated camera reading that corresponds to MNIST input “5”, “0” and “4”; (b) the scheme of a proposed optical setup where source can be a single-mode fiber array collimated into free space or coupled in a photonic chip; diffusion can be achieved by an optical amplitude-only modulation by an SLM; and confusion must be done with complex amplitude-and-phase convolution in the Fourier domain; the optional final step of image classification with a heterogeneous convolutional 4f classifier can be alternatively replaces by a CMOS diode array DA.  

The training and validation curves are shown in Figure 2(c), where the performance of the CNN model in Figure 2(a) on the original MNIST data is juxtaposed with its performance on generated in optical simulation dataset of hash/compressed MNIST images using the Algorithm 2. The results show that the proposed compression method successfully picks up features of the original images showing ~8% performance trade-off for compression 61 for 60k hash dataset generated in the optical simulation. We would like to emphasize, that this gap could be mitigated by increasing the depth of the NN in optical or electronic domain by increasing the number of parameters (layers) in the network architecture since systematic optical errors can be incorporated into the weights of the model by regular NN training techniques.

In the next step, we study classification accuracy as a function of the compression coefficient: (input resolution)/(output resolution). We start from compression of 10 and go up to 60 while keeping the aspect ratio ~ 9:88. The compression is accomplished by averaging over the adjacent pixels to emulate the performance of cylindrical lens $L_3$ in Figure 3(b). The MNIST and Fashion-MNIST datasets have been considered. For Fashion-MNIST, we increased the depth of the NN model adding one more dense layer with 70 categories and relu activation to accommodate the higher complexity of the dataset with respect to MNIST. The performance trade-off is about 15% for the compression of 60. The accuracy roll-off shows a similar tendency for both data set hashing/compression, see Figure 2(d). Based on these results, we project that compression loss may increase as a function of input data complexity due to the diffractive losses in an optical setup.

One last test has been performed on GTSRB dataset of German traffic signs with 50,000 RGB images in total and 43 classes. The images have been first converted into single channel grayscale, then the dataset has been augmented with additional images to insure equal number of images per class. The CNN has been expanded to accommodate more complex data to two convolution and two fully-connected layers with about 8,000,000 parameters. The training was terminated by early stopping under the condition of loss not changing for 50 epochs. With compression coefficient of 17 the validation accuracy loss is 54.4%. This is a cumulative loss as the images have been both compressed and converted to grayscale from RGB.

Here, the in silico CNN that is used for image classification, 2(a). However, it can be replaced by a heterogeneous 4f CNN (HCNN) classifier,24,29 as shown in Figure 2(b) and in the scheme Figure 3(b). Delayed domain crossing can offer a benefit of harvesting maximum of the incoming analogue signal information and high-speed and high energy efficiency optical convolution layer, while not adding the optical IO related latency since the data is pre-existent in optical domain as it is.
| Dataset    | Compression coefficient | Drop, training (%) | Drop, validation (%) |
|------------|-------------------------|--------------------|----------------------|
| MNIST      | 60                      | 10.4               | 11.9                 |
| Fashion-MNIST | 60                 | 18.2               | 19.2                 |
| GTSRB      | 17                      | 47.6               | 54.4                 |

Table 1. Table of accuracy trade-off for three datasets: MNIST (binary), Fashion-MNIST (binary) and GTSRB (RGB) binarized.

3 Discussion

We demonstrate a new opto-electronic hashing/compression algorithm capable of analogue 2D signal processing in optical domain with minimal domain conversions. The Algorithm 2 is based on post-quantum hashing scheme SWIFFT\(^2\), see Algorithm 1, that takes advantage of information scrambling properties of Fourier transform and security of lattice-based algorithms. The proposed algorithm shows optical hardware-to-algorithm homomorphism as the associated tensor algebra operations, such as Fourier transform and by-element tensor multiplication, is native to Fresnel’s optics and have shown the potential to be low-bandwidth and latency. Theoretical limits of this compression algorithm suggest a possibility of turning a input 2D-array into a 1D-array on the output.

As compared to other compression algorithms used in compressive sensing and digital compression, where information gets compressed before transmission or processing, and subsequently decompressed to recover the original image, we propose to use this algorithm in combination with heterogeneous NN training techniques. This way, hashed information is inputted into the heterogeneous ML accelerator where training and classification happens on already augmented data. The output of this this system, Figure 3(b), can be used as a trigger for larger (and slower) edge data acquisition systems when the original high-resolution data of an event of interest is needed. Alternatively, for systems with small input data variability, look-up tables can be created to accomplish de-hashing back-end. The appeal is to reduce the requirement on throughput and latency of heterogeneous systems by pre-hashing and pre-compressing the data optically.

Our simulation results show, for compression coefficient 60 the MNIST accuracy drop is about 10.8% and for Fashion-MNIST 18.5%, see Table 1. For GTSRB 8 bit RGB dataset converted into 1-bit grayscale beforehand, the loss is \(\sim 50\%). This is expected since the original digital version of SWIFFT algorithm is not lossless. We would like to mention, that when designing the corresponding ML networks we used a minimalistic approach keeping the architecture of the classifier simple. We believe, that performance of an ML algorithm on the compressed datasets can be boosted with more complex and tuned NN architectures. Our main results and conclusions are: (1) the proposed algorithm is highly-parallelisable and capable of accomplishing fast and efficient 2D data compression with quadratically increasing loss as a function of compression coefficient information; (2) the algorithm is robust with respect to the image data of various nature (MNIST, Fashion-MNIST, GTSRB); (3) as a hash function, the resulting dataset fundamentally has features necessary for applications in cryptography, such as asymptotic security and collision resistance\(^2\,3\). In addition, the projected possibility of outputting a 1D-array back end suggest on-chip implementations for the further ultra-fast photonic data processing, as opposed to dealing with the known in silico interconnect bottlenecks in optical pre-processors.

4 Methods

**SWIFFT Hashing Algorithm**

In this paper, we focus on one of the FFT-based compression/hashing algorithms that belong to the family of lattice hashing algorithms. Since optical crypto is an interdisciplinary class of problems, lets begin with revisiting some fundamental concepts of number theory, cryptography and optical processing, that are essential for understanding of the SWIFFT algorithm\(^2\)

First, most deployed crypto algorithms heavily employs modular arithmetic of rings. In number theory a ring \(\mathbb{Z}\) can be broadly defined as a set of elements with defined operation of addition and multiplication. Addition is always commutative, while multiplication might be not. Hence, a commutative ring is any ring with commutative multiplication. By number of elements, a ring can be infinite (e.g. a ring of real numbers), and finite (e.g. a ring of Pauli matrices). The SWIFFT is requiring modular rings \(\mathbb{Z}_p\), also called in group theory "quotient groups", of a modulo \(p\) - a prime number.

Other two concepts that arise in mathematical theory of cryptography are "confusion" and "diffusion". Confusion is an algorithm that puts each value of an input message in dependence on several (maximum possible) pieces of a key. The goal of confusion is to scramble the input to achieve "one-way" functionality of a given hash algorithm. "Diffusion" is an algorithm that reduces collision probability by ensuring that small alteration of a plain text results in an avalanche of change in chipertext.

With these concepts in hand, let’s proceed to the description of the original SWIFFT algorithm. Given a binary message to encode \(M\), choose a large integer \(k\). Set \(n = 2^k\), a prime number \(p\) such that \(2n = p - 1\), and choose an integer \(m \geq \log p\). Split the binary input message \(M\) into a set of \(m\) \(n\)-bit vectors such that a matrix \(x_{i,j}\) of dimensions \(\{m,n\}\) is formed. Generate a
weight matrix $\omega_{mn}$ out of $m$ replicas of a sequence of length $n$, and a key $k_{mn}$ such that it consists of $n$ replicas of a sequence of length $m$. The elements of these sequences can be randomly drawn from a set of elements on a commutative ring $\mathbb{Z}_p$ or use a more sophisticated generation mechanism. SWIFFT hashing algorithm is laid out below.

**Algorithm 1:** Electronic SWIFFT hashing algorithm

| Step | Description |
|------|-------------|
| 1 | Multiply an input matrix $x_{mn}$ to an integer matrix of weights $\omega_{mn}$ to form a lattice $\mathcal{L}(\omega_n) = \omega_{mn}x_{mn}$; |
| 2 | DFT the earlier introduced key $\mathcal{A}_{mn} = \mathcal{F}[\omega_{mn}]$; |
| 3 | Accomplish diffusion of an input message by taking a DFT of the result of the previous step $y_{mn} = \mathcal{F}[\mathcal{L}(\omega_n)]$; |
| 4 | Accomplish confusion by multiplying the result of step 2 by the result of step 3: $\mathcal{A}_{mn}y_{mn}$; |
| 5 | Inverse Fourier transform the result $z_{mn} = \mathcal{F}[\mathcal{A}_{mn}y_{mn}]$; |
| 6 | To accomplish the data compression, sum the resulting matrix over $m$: $\sum_n z_{mn} = \tilde{z}_n$; |

Hence, SWIFFT hash function can be defined via the operation of convolution

$$h(x_{i,j}) = \mathcal{L}(x_{mn}) * \mathcal{A}_{mn}$$  \hspace{1cm} (1)$$

where $*$ denotes a convolution defined for 2D as:

$$\mathcal{L}_{ij} * \mathcal{A}_{ij} = \sum_{t=-m}^{m} \sum_{s=-n}^{n} \mathcal{L}_{i-t,j-s} \mathcal{A}_{t,s}$$

**SWIFFT-based Optical Compression**

Let's first review several fundamental concepts that come from optical processing. The one to start with is the two-dimensional Fourier transform of an optical signal that is defined in, e.g., eqn. (2-1):

$$\mathcal{F}\{g(x,y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dxdy g(x,y) \exp[-2\pi i(f_x x + f_y y)] \hspace{1cm} (2)$$

where $f_x$ and $f_y$ are the $x$ and $y$ components of the corresponding spatial frequency spectrum. Interestingly, an ideal thin lens is a Fourier transform engine supplied by nature: when placing the input transparency centered at a front focal plane of an ideal thin lens one receives a Fourier-transformed signal at a rear focal plane. Despite the benefits, one needs to remember, that an ideal thin lens is, of course, a theoretical model of a real thin lens performance, that only works approximately in practice. Also, optical systems, free-space and on-chip, are prone to noise that has a potential to irreversibly corrupt a message. Keeping this noise near the level of common SNR of modern telecommunication links is one of the performance targets of optical hashing.

One more missing piece is the convolution theorem, e.g.

$$[g * h](x) = \int_{-\infty}^{\infty} d\psi g(\psi)h(x-\psi) = \mathcal{F}\{g(x)\} \cdot \mathcal{F}\{h(x)\} \hspace{1cm} (3)$$

where "*" denotes the convolution of signal $g(x)$ with a mask $h(x)$. One can see that applying the Fourier transform to the original functions $g(x)$ and $h(x)$ turns an elaborate multi-step matrix multiplication into a simple by-element product.

In the original SWIFFT implementation, convolution is achieved by applying the convolution theorem that equates convolution with dot-product multiplication in the Fourier domain (3). Fourier transform $\mathcal{F}\{\bigcirc\}$ is done electronically, typically with a fast Fourier transform or a modular Fourier transform algorithm. The computational complexity of such implementation is $n \log_p$, which is an achievement as opposed to mn-complexity of the brute force convolution. However, the Fourier transform remains the main consumer of computational resources in this algorithm, slowing down both key generation and encryption/decryption. These result in lower key update rate and, consequently, compromises the security.

In this paper we propose, simulate and analyze an optical SWIFFT inspired data compression algorithm applied to benchmark ML data sets such as MNIST. In the envisioned scheme, see Fig. 3(a), the computational engine consists of an optical Fourier transform accomplished with lenses L1 and L2, and a dot-product multiplication done by SLMs. The whole 4f-systems forms a complex amplitude-and-phase convolution in the Fourier domain.

One starts from acquiring a two-dimensional incoming message $\mathcal{M}$ into a matrix $x_{mn}$. One also generates $n$ elements on a ring $R_m$ to form the set of coefficients $\omega_n$. In case with using SLMs, one naturally works on a ring $\mathbb{Z}/256\mathbb{Z}$, in other words, mod 256. The dot-product multiplication in step 2 of an algorithm can be done electronically or optically, depending on the
original domain of the incoming message \( \mathcal{M} \). It is worth mentioning that key \( a_{mn} \) often also built of a polynomial on a ring, similar to the sequence \( u_n \), forming matrix \( \omega_{mn} \) in step 2 of the SWIFFT.

Steps 3-6, convolution (1) in Fourier domain, can be achieved optically using a 4f-system such as\(^{26} \). A thin lens, being a natural spatial-to-Fourier domain converter, provides low latency (limited by speed of light), and high parallelism passive FT-engine. Dot-product multiplication is pseudo-passive as multiplication occurs on the surface of a light-modulating device that does not require processing power. The matrices, generated in step 2 and step 4 loaded onto light modulators requiring electric input and initial data processing. The Fourier transform of the key, needed in step 4, can as well be accomplished optically. Last step of the algorithm can be accomplished by either summing over the 2D reading of the camera CCD or by replacing a 2D CCD with a 1D diode array with controllable number of activated diodes, allowing parsing through the output 2D signal during the image collection time.

Algorithm 2: SWIFFT-like optical data compression algorithm

\[
\begin{align*}
\text{Input :} & \text{An incoming signal } \mathcal{M} \text{ of dimensions } m \times n \\
\text{Output :} & \text{CCD reading of hash matrix } |z_{mn}|^2 \\
1 \text{ Optically multiply } x_{mn} \text{ to an integer matrix of weights } \omega_{mn} \text{ by using modulating devices M1 and M2 in Fig. 3;} \\
2 \text{ Optically Fourier transform the resulting signal to achieve diffusion } y_{mn} = \mathcal{F}[\mathcal{D}(\omega_{n})]; \\
3 \text{ Optically generate or electronically pre-calculate and project onto a modulating device M3 in Fig. 3 the matrix } \\
A_{mn} = \mathcal{F}[a_{mn}] \text{ where it will get by-pixel multiplied by the output of step 2;} \\
4 \text{ Fourier transform with L2: } z_{mn} = \mathcal{F}[A_{mn} y_{mn}]; \\
5 \text{ Compress with cylindrical lens L3;} \\
6 \text{ Get the CCOS diode array reading } |z_{mn}|^2; \\
\end{align*}
\]

For completeness, one could alternatively accomplish data compression at the domain crossing stage in step 5 by the use of reconfigurable diode arrays or electronically.

We have taken MNIST, Fashion-MNIST and GTSRB datasets as input data. The resulting hashed images for MNIST are shown in Fig. 3(b). The cipher images have been generated for the same diffusion matrix \( \omega_{mn} \), and the same key \( a_{mn} \). For additional details see Sec. 2.

It is worth re-emphasizing that an accurate optical confusion can only be achieved if the dot-product multiplication in the Fourier domain captures both phase and amplitude components. In particular, \( A_{mn} \) is a complex valued matrix with, generally speaking, non-trivial real and imaginary parts. The signal \( x_{mn} \) will also carry a phase component after passing lens L1. Such complex convolution can be done interferometrically or with the help of new-generation modulating devices that are yet to be achieved experimentally.

Analysis

An electronic simulation of the optical performance of the proposed hashing system has been done using Fresnel’s transfer function as a propagator, and focusing function to simulate the lensing effect, see\(^{12} \). It is important to satisfy the applicability criteria to obtain adequate results. In our simulation, the focal distance is related to the size of an incoming image, the size of the spatial modulator pixel, and the wavelength as \( f = L \cdot \Delta / \lambda \). The input versus output hash are shown in Fig. 3(a). To correctly model such an optical setup in a simulation, just the same as in DFT, optical FT simulation needs padding. These results into an estimated input field of an image of about 250X250 pixel as opposed to its original size (e.g. 28X28 for MNIST and Fashion-MNIST). If the latter dimensions are taken as an input, then the compression rate of this system is 60. We would like to remind the readers, that comparing the nominal MNIST resolution to the system output resolution to assess compression efficiency in this case is not justified. For many potential applications, the input is analogue and the analogue-to-digital conversion is expected to happen only after the pre-processing with our custom system. Hence, direct comparison of the featured here information compression to the digital analogues is not straightforward. The presented system is not intended to replace any of the existing electronic systems. Instead, it is meant for use, e.g., in an edge processing setup for optical pre-processing. Hence, the input is envisioned to be, e.g., an optical fiber array or a real-life scene that is usually captured by a several megapixels CCD/CMOS camera. Taking these into consideration, the cited compression coefficient is significant and can get even higher depending on the nature of the data and the demand on classification accuracy.

References

1. Naughton, T. J., McDonald, J. B. & Javidi, B. Efficient compression of fresnel fields for internet transmission of three-dimensional images. Appl. Opt. 42, 4758–4764 (2003).
2. Dufaux, F., Xing, Y., Pesquet-Popescu, B. & Schelkens, P. Compression of digital holographic data: an overview. Appl. Digit. Image Process. XXXVIII 9599, 163–173 (2015).
3. Zhang, L., Xiong, R., Chen, J. & Zhang, D. Optical image compression and encryption transmission-based on deep learning and ghost imaging. Appl. Phys. B 126, 1–10 (2020).
4. Donoho, D. L. Compressed sensing. IEEE Transactions on information theory 52, 1289–1306 (2006).
5. Qaisar, S., Bilal, R. M., Iqbal, W., Naureen, M. & Lee, S. Compressive sensing: From theory to applications, a survey. J. Commun. networks 15, 443–456 (2013).
6. Li, X., Zhang, G., Huang, H. H., Wang, Z. & Zheng, W. Performance analysis of GPU-based convolutional neural networks. In 2016 45th International conference on parallel processing (ICPP), 67–76 (IEEE, 2016).
7. Jordà, M., Valero-Lara, P. & Pena, A. J. Performance evaluation of cudnn convolution algorithms on nvidia volta gpus. IEEE Access 7, 70461–70473 (2019).
8. Mills, G. A. & Yamaguchi, I. Effects of quantization in phase-shifting digital holography. Appl. Opt. 44, 1216–1225 (2005).
9. Singh, N. & Sinha, A. Chaos-based secure communication system using logistic map. Opt. Lasers Eng. 48, 398–404 (2010).
10. Juleang, P. & Mitatha, S. Optical hash function for high speed and high security algorithm using ring resonator system. In 2021 7th International Conference on Engineering, Applied Sciences and Technology (ICEAST), 160–163 (IEEE, 2021).
11. Zhou, H. et al. Photonic matrix multiplication lights up photonic accelerator and beyond. Light. Sci. & Appl. 11, 1–21 (2022).
12. Zhou, J. et al. Optical edge detection based on high-efficiency dielectric metasurface. Proc. Natl. Acad. Sci. 116, 11137–11140 (2019).
13. Situ, G. & Zhang, J. A cascaded iterative fourier transform algorithm for optical security applications. Optik 114, 473–477 (2003).
14. Hennelly, B. & Sheridan, J. T. Optical image encryption by random shifting in fractional fourier domains. Opt. letters 28, 269–271 (2003).
15. Situ, G. & Zhang, J. Double random-phase encoding in the Fresnel domain. Opt. Lett. 29, 1584–1586 (2004).
16. Jin, W. & Yan, C. Optical image encryption based on multichannel fractional fourier transform and double random phase encoding technique. Optik 118, 38–41 (2007).
17. Lang, J., Tao, R. & Wang, Y. Image encryption based on the multiple-parameter discrete fractional fourier transform and chaos function. Opt. Commun. 283, 2092–2096 (2010).
18. Zhou, N., Wang, Y. & Gong, L. Novel optical image encryption scheme based on fractional Mellin transform. Opt. communications 284, 3234–3242 (2011).
19. Zhang, L., Zhou, Y., Huo, D., Li, J. & Zhou, X. Multiple-image encryption based on double random phase encoding and compressive sensing by using a measurement array preprocessed with orthogonal-basis matrices. Opt. & Laser Technol. 105, 162–170 (2018).
20. Goodman, J. W. Introduction to fourier optics. Introd. to Fourier optics, 3rd ed., by JW Goodman. Englewood, CO: Roberts & Co. Publ. 2005 1 (2005).
21. Ajtai, M. Generating hard instances of lattice problems. In Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, 99–108 (1996).
22. Lyubashevsky, V., Micciancio, D., Peikert, C. & Rosen, A. SWIFFT: A modest proposal for fft hashing. In International workshop on fast software encryption, 54–72 (Springer, 2008).
23. Miscuglio, M. et al. Massively-parallel amplitude-only fourier optical convolutional neural network. In 2021 Conference on Lasers and Electro-Optics (CLEO), 1–2 (IEEE, 2021).
24. Miscuglio, M. & Sorger, V. J. Photonic tensor cores for machine learning. Appl. Phys. Rev. 7, 031404 (2020).
25. Miscuglio, M. et al. Massively parallel amplitude-only fourier neural network. Optica 7, 1812–1819 (2020).
27. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).

28. Real-Time Computer Vision group, I. German Traffic Sign Recognition Benchmark. https://www.kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign (2011).

29. Hu, Z. et al. Batch processing and data streaming fourier-based convolutional neural network accelerator. *arXiv preprint arXiv:2112.12297* (2021).

30. Lyubashevsky, V. & Micciancio, D. Generalized compact knapsacks are collision resistant. In *International Colloquium on Automata, Languages, and Programming*, 144–155 (Springer, 2006).

31. Györfi, T., Cret, O., Hanrot, G. & Brisebarre, N. High-throughput hardware architecture for the swiftf/swifftx hash functions. *IACR Cryptol. ePrint Arch. 2012*, 343 (2012).

32. David, V. Computational fourier optics (2011).

5 Acknowledgements

M.S.G acknowledges valuable discussions on algorithmic approaches and simulation details with Jonathan George and Nicholas Gorgone.