

Review of Integrated Diffractive Deep Neural Networks

Huayi Sheng *

University of Shanghai for Science and Technology, Shanghai, 200031, China

* Corresponding Author Email: ussthsheng@gmail.com

Abstract. An integrated photonic diffractive deep neural network (ID^2NN) is one of the most exciting cross-discipline fields of artificial intelligence and optical computing, combining deep learning with the power of light-speed processing on an integrated platform. We know that neural network in a digital computer is based on transistors, which have significant challenges in keeping pace with Moore's law and limited real-time processing applications due to the increased computational costs associated with them. However, with remarkable progress and advancement in silicon photonic integrated circuits over the last few decades, ID^2NN hold the promise of on-chip miniaturisation and high-speed performance with low power consumption. This paper covers the essential theoretical background for constructing the ID^2NN and reviews the research status of optical diffractive neural networks in the field of neuromorphic computing. Problems of narrowing down current ID^2NN applications are also included in this review. Finally, future research directions for ID^2NN are discussed, and conclusions are delivered.

Keywords: Optical Diffractive Neural Network; Integrated Photonic Circuit; Image Classification.

1. Introduction

Deep learning[1] is a fast-growing field with many use cases and applications in computer vision [2], speech recognition[3], robotic systems[4], etc. However, the real-world engineering applications of deep neural networks(DNN) have been limited due to the inference time and increased computational power associated with these networks.[5] Although much of the recent advances in deep learning have been enabled by electronic-based hardware, its optical implementation has great significance because of the intrinsic parallelism and light-speed computing with low power consumption. [6, 7] Also, as we move into the future, big challenges appear in making computing chips faster and keeping pace with Moore's law.[8] For instance, as autonomous vehicles rely on cameras and artificial intelligence image processing to make decisions at high speed, conventional digital processors show significant incompetence, whereas the optic-driven computing method is poised to address the challenges.

The speed of computation under von Neumann architecture is limited by two factors: the efficiencies of data transfer and the speed of processing(data computation).[9, 10] Electronic circuits operate based on millions of transistors that induce latency and generate heat. Integrated photonic circuits accomplished traditional computing through optical logic processing units and have dramatically improved the performance and reliability and enabled the reduction of size, power consumption and cost per function. In recent years, researchers have leveraged the power of photonics to accelerate the inference of neural networks.[11–13] The main ones are photonic chips-based and optical diffractive neural networks[14]. The former uses optical interference unit to process meshes of light, and the latter performs analogue computing through wave propagation. Diffractive deep neural network (D^2NN) focus on the paradigm that uses light waves as the computing medium to perform the functionality of neural networks. The interference pattern of waves will determine the result, which allows for instantaneous computation. In other words, the signals are processed in the optical domain while it is in motion. Thus, computing at the speed of light can be achieved, and notably, it will produce massive performance and efficiency gains.

Developing an integrated deep diffractive neural network (ID^2NN) will provide an alternative machine learning architecture for on-chip optical computing that may address the needs of artificial intelligence systems. The outline of this review is as follows: it begins with key concepts and essential

terminologies for D^2NN in section 2, followed by section 3 with the mathematical model of D^2NN . Next, section 4 reviews the research status of D^2NN and moves into the domain of ID^2NN in section 5. Section 6 extracts design problems and challenges for ID^2NN . The future directions and conclusions are delivered together in section 7.

2. Key Concepts in Diffractive Deep Neural Network

In classical optics, there are three common ways to model light. The simplest is the ray model, where the light travels in the form of rays. When light waves propagate through and around objects whose dimensions are much small than the wavelength of the light, the rays obeying a set of geometrical rules cannot adequately describe the phenomenon of interference and diffraction.[15] The wave nature of light is discerned in the scope of wave optics, and light is described by a scalar function stating that each object point emits spherical waves, and the light propagates in the directions perpendicular to the wavefronts. Wave optics reaches its limitations when a vector formulation of optical phenomena is required, such as polarisation effects.[15] Electromagnetic optics enters a bigger picture as it provides the most complete and accurate treatment of light in classical optics, where light is described by a vector theory comprising an electric field and a magnetic field, namely the electromagnetic wave that can be derived from Maxwell's equations. However, in the early twentieth century, certain observations (i.e. photoelectric effect,compton scattering, blackbody radiation) indicated that light acts like a particle and comes in discrete energy packets(quanta/photons), which is not predicted by Maxwell's equations.[16, 17] These failures ultimately led to the formalisation of quantum optics.

The relationship between optic theories is illustrated in Figure 1, ray optics is the limit of wave optics when the wavelength is infinitesimally small. Electromagnetic optics encompasses wave optics, and it turns out that wave optics is a scalar approximation of electromagnetic optics.[15] The theory of quantum optics encompassing classical optical theories, provides an explanation for virtually all optical phenomena.

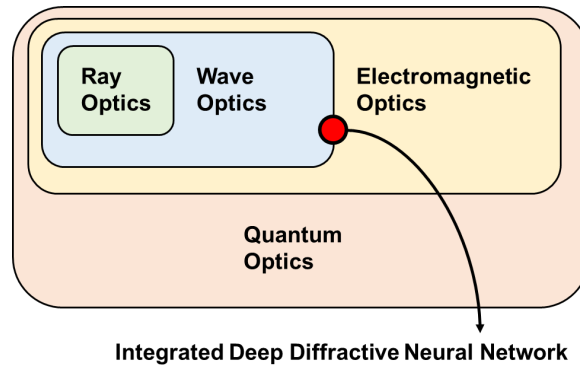


Figure 1. The relationship between optic theories

In this paper, the design theory of the integrated deep diffractive neural network is based on wave optics with a focus on diffraction and will be discussed in detail in the next section.

2.1 Diffractive Optics

The way a 1-D wave ψ evolves in space and time is determined by the wave equation[15, 16]:

$$\frac{\partial^2 \psi(x, t)}{\partial^2 x^2} - \frac{1}{c^2} \frac{\partial^2 \psi(x, t)}{\partial^2 t^2} = 0 \quad (1)$$

where x denotes position, t denotes time, and c denotes the propagation speed of light. The general solution to the wave equation is given as:

$$\psi(x, t) = \cos(kx - \omega t) \quad (2)$$

Here, $k = \frac{2\pi}{\lambda}$ is the wave number, and $\omega = \frac{2\pi}{T}$ is the angular frequency. The wave is periodic in space with period λ and periodic in time for a period of T . Interpretively, k converts a distance x to a phase shift and ω converts a time interval t to a phase shift. More generally, a plane wave travelling in the positive x -axis direction can be described as $\cos(kx - \omega t + \phi)$, where ϕ is the initial phase shift. However, to make the computations simpler, as we will see later, we can convert them to complex notation using Euler's identity ($e^{i\theta} = \cos(\theta) + i\sin(\theta)$):

$$\cos(kx - \omega t + \phi) = \text{Re}\{e^{i(kx - \omega t + \phi)}\} \quad (3)$$

The advantage of writing the expression in the exponential form is allowing us to separate time-dependence from space-dependence and straightforwardly compute the intensity of the superposition of two fields:

$$e^{i(kx - \omega t + \phi)} = e^{ikx} e^{i\phi} e^{-i\omega t} \quad (4)$$

Now, we can write the time-harmonic field in three dimensions:

$$\begin{aligned} \psi(x, y, z, t) &= A(x, y, z) \cos(\phi(x, y, z) - \omega t) \\ &= \text{Re}\{A(x, y, z) e^{i\phi(x, y, z)} e^{-i\omega t}\} \\ &= \text{Re}\{U(x, y, z) e^{-i\omega t}\} \end{aligned} \quad (5)$$

The spatial part of the field $U(x, y, z)$ is factorised out in Equation(5), which is a complex-valued function. It has both an amplitude and a phase which vary as a function of position. The amplitude indicates how large the oscillation is at a certain point, while the phase difference indicates the delay accumulated from the source point. The wave described by the scalar function:

$$\psi(x, y, z, t) = U(x, y, z) e^{-i\omega t} \quad (6)$$

should satisfy the 3D wave equation:

$$\nabla^2 \psi(x, y, z, t) - \frac{1}{c^2} \frac{\partial^2 \psi(x, y, z, t)}{\partial t^2} = 0 \quad (7)$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ (the Laplacian).

Plugging Equation (6) into Equation (7) and dividing out $e^{-i\omega t}$ yields the Helmholtz equation:

$$(\nabla^2 + k^2)U(x, y, z) = 0 \quad (8)$$

It turns out that $U(x, y, z)$ must obey this time-independent equation. Plane waves (Equation(9)) and spherical waves (Equation(10)) are the two solutions to the Helmholtz equation:

$$U_P(r) = e^{ik \cdot r} = e^{i(k_x x + k_y y + k_z z)} \quad (9)$$

$$U_S(r) = \frac{e^{ik|r|}}{4\pi|r|} \quad (10)$$

where $k^2 = k_x^2 + k_y^2 + k_z^2$ and r denotes the position vector (x, y, z) .

The time-averaged intensity of the optical field can be calculated as follows:

$$I(r) = |U(r)|^2 \quad (11)$$

2.1.1 Propagation of Wave Fields

To model the propagation of monochromatic fields in free space, there are two ways to conceptualise this propagation. The first one is the angular spectrum method which is the addition of decomposed plane waves. The second is the Rayleigh-Sommerfeld formula which is the addition of impulse responses from point sources.[15] The following will discuss and analyse the two different propagation methods along with their computational method.[16–18]

2.1.1.1 Angular Spectrum Propagation

Supposing a monochromatic complex-valued field $U_0(x, y)$ is in the plane $z = 0$ to compute the propagated field $U_z(x, y)$ by a distance of z , the angular spectrum method yields the angular spectrum of the field by decomposing the initial field into plane waves. Each of the plane waves can propagate separately, as described in Equation (9). Then, by adding the plane waves together, we can

find the propagated field $U_z(x, y)$. The computational recipe of the angular spectrum method contains the following steps:

- Decompose the initial plane into plane waves by taking the Fourier transform, which gives the spectrum $\hat{U}_0(v_x, v_y)$:

$$\begin{aligned}\hat{U}_0(v_x, v_y) &= \iint U_0(x, y) e^{-2\pi i(v_x x + v_y y)} dx dy \\ &= \mathcal{F}\{U_0(x, y)\}(v_x, v_y) \\ &= \mathcal{F}\{U_0(x, y)\}\left(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}\right)\end{aligned}\quad (12)$$

- Calculate the field in the plane z by adding separately propagated plane waves together with weights $\hat{U}_0(v_x, v_y)$ through the inverse Fourier transform:

$$\begin{aligned}U_z(x, y) &= \iint \hat{U}_0(v_x, v_y) e^{2\pi i(v_x x + v_y y + v_z z)} dx dy \\ &= \mathcal{F}^{-1}\left\{\hat{U}_0(v_x, v_y) e^{2\pi i z \sqrt{\frac{1}{\lambda^2} - v_x^2 - v_y^2}}\right\}(x, y) \\ &= \mathcal{F}^{-1}\left\{\hat{U}_0\left(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}\right) e^{i k_z z}\right\}(x, y), \text{ where } k_z = \sqrt{k^2 - k_x^2 - k_y^2}\end{aligned}\quad (13)$$

As there is no restriction on the values for k_x and k_y , when $k_x^2 + k_y^2 > k^2 = \left(\frac{2\pi}{\lambda}\right)^2$, k_z is imaginary, and Equation (13) will become:

$$U_z(x, y) = \mathcal{F}^{-1}\left\{\hat{U}_0\left(\frac{k_x}{2\pi}, \frac{k_y}{2\pi}\right) e^{-|k_z|z}\right\}(x, y)\quad (14)$$

This leads to an evanescent field where higher spatial frequency (k_x, k_y) can not propagate because it decays exponentially with the propagation distance z . [16] Therefore, we can also conclude that the resolution of an image is limited by the wavelength of the illuminating light λ . Features in an image carried by wavelength which is less than λ cannot be captured during propagation.

2.1.1.2 Rayleigh-Sommerfeld Propagation

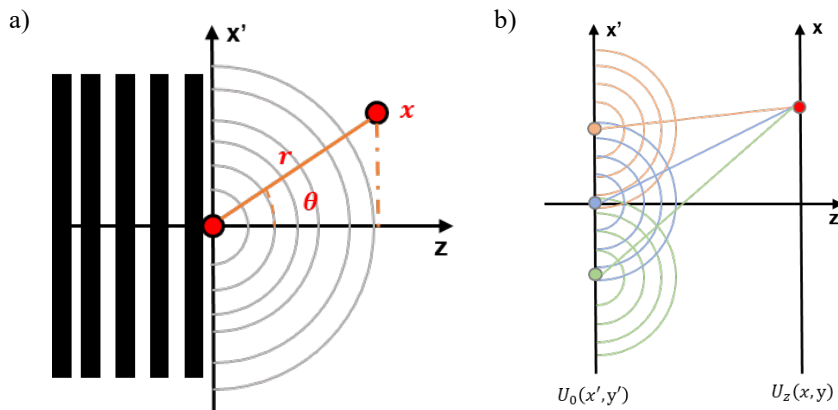


Figure 2. Field propagation based on Rayleigh-Sommerfeld integral
a) The field that propagates through a small pinhole is the impulse response (Equation(15)).
b) The Rayleigh-Sommerfeld integral states that to propagate a field from an initial plane to another plane, all the impulse responses that originate from the initial plane must be summed together (Equation(16)).

The impulse response derived from the Helmholtz equation is shown in Equation(15) below. [16] Each point in the field is assumed to emit a spherical wave $\frac{e^{ikr}}{r}$ whose amplitude is the same as that point in the field, but it is multiplied by a factor of $\frac{z}{r} = \cos(\theta)$, which means the spherical wave attenuates when it propagates towards sideways.

$$h(r) = \frac{ze^{ikr}}{r^2} \quad (15)$$

The Rayleigh-Sommerfeld integral can then be applied to find the total field at one point in the final plane[16]:

$$U_z(x, y) = \frac{1}{i\lambda} \iint U_0(x', y') \frac{ze^{ikr}}{r} dx' dy' \quad (16)$$

where r denotes the distance between the point in the initial plane $(x', y', 0)$ and the point (x, y, z) in the final plane:

$$r = \sqrt{(x - x')^2 + (y - y')^2 + z^2} \quad (17)$$

To calculate the Rayleigh-Sommerfeld integral computationally, a direct discretisation is performed[19]:

$$U_z(x, y) = \frac{1}{i\lambda} \sum_{m=1}^M \sum_{N=1}^N U_0(x', y') \frac{ze^{ikr}}{r} \Delta x \Delta y \quad (18)$$

The complex transform of the field between the two planes can be formulated as a matrix-vector multiplication:

$$\begin{aligned} \mathbf{U}_z &= \mathbf{W} \mathbf{U}_0 \\ \begin{bmatrix} U_z(1,1) \\ \vdots \\ U_z(M,N) \end{bmatrix} &= \begin{bmatrix} \frac{1}{i\lambda} \cdot h(r_1) & \cdots & \frac{1}{i\lambda} \cdot h(r_{M*N}) \\ \vdots & \ddots & \vdots \\ \frac{1}{i\lambda} \cdot h(r_{M*N}) & \cdots & \frac{1}{i\lambda} \cdot h(r_1) \end{bmatrix} \begin{bmatrix} U_0(1,1) \cdot \Delta x \Delta y \\ \vdots \\ U_0(M,N) \cdot \Delta x \Delta y \end{bmatrix} \end{aligned} \quad (19)$$

where $\mathbf{U}_0$ and $\mathbf{U}_z$ represents the vectorised initial field and propagated field; $\mathbf{W}$ is the matrix of multiplicative factors determined by impulse responses. It is important to note that $\mathbf{W}$ is a symmetric matrix ($\mathbf{W} = \mathbf{W}^T$), which can be leveraged to reduce the computational effort by making replicas below the diagonal.

The Rayleigh-Sommerfeld diffraction(RSD) integral has been proved to be the only method that forecasted phase shift from the input to the output plane with no error, while the angular spectrum method distorted the phase map because of the corruptions of more minor features by applying Fourier transform.[20, 21] The main setback of RSD is the calculation complexity order of $O((M \times N)^2)$. Nonetheless, we have implemented a parallel-computed RSD that can be used to reduce the elapsed computation time. Taking advantage of the workstation installed with 8 Intel Core Xeon Gold 5218R, the parallelisation of the RSD can achieve a 210-time reduction of computation time compared with a standard CPU-based calculation on an Intel Core i7-8550U.

2.2 Complex-Valued Neural Network

The complex-valued neural networks(CVNNs) are a powerful modelling tool for domains where input data or signal can be naturally interpreted in terms of complex numbers, such as optical signal processing.[19] In CVNNs, the weights, threshold values, input, and output signals are all complex numbers, and the activation function and its derivatives have to be well-behaved everywhere in the complex plane. They have also shown more powerful capability than real-valued neural networks in processing real-world signals. With complex weights, the neural network makes a significant step in gaining a higher functionality because multiplying by a complex number changes a value's both magnitude and direction, which implies a mathematical rotation. The CVNNs rotate the signals as they pass through. As shown in Figure 2.b, in the feed-forward complex-valued neural network, the inputs z_n are sent to a complex-valued neuron, each corresponding to a complex weight; then we

weight and sum the inputs, perform the non-linear activation, and finally, obtain the output y_k [19, 22]. The purpose of training is to get the output close to the ground truth[23]. It is constantly updated to optimise the process of training until the network is converged. In the following statement, the form of y_k is as follows:

$$y_k = \sigma \left(\sum_{i=1}^n w_i * z_i \right) \quad (20)$$

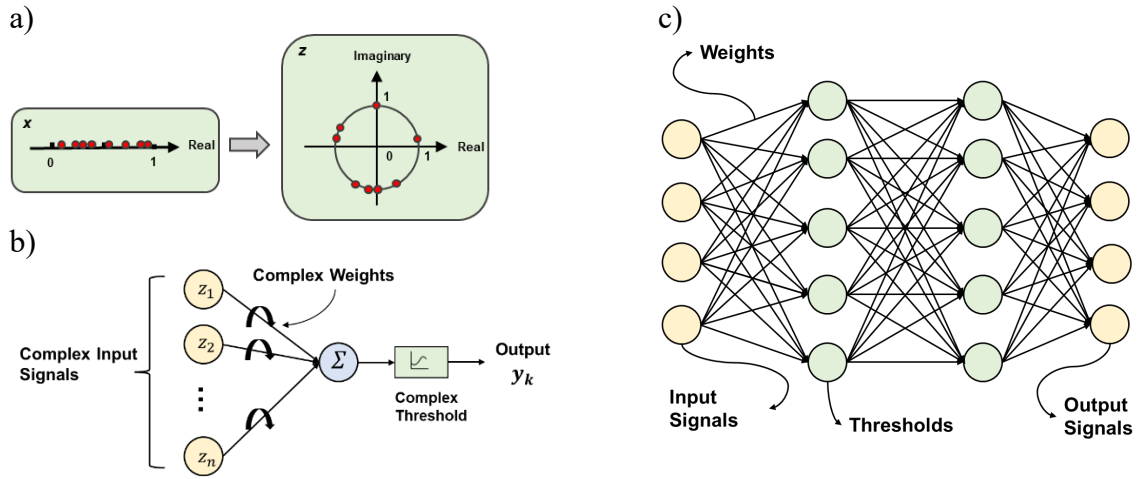


Figure 3. Mechanism of CVNNs

- a) Mapping normalised inputs to the complex unit circle.
- b) Structure of a complex-valued neuron.
- c) The model of feed-forward complex-valued neural network.

In order to work with inputs that are from the real world but not complex numbers like image pixels and want the network to give real number answers or even classification categories, we need to prepare inputs and also map the outputs. Normalised real number inputs can be transformed into the complex unit circle through:

$$z = e^{j2\pi x}, x \in [0,1] \quad (21)$$

Through this transformation, the incoming data can be best spread out over the complex domain with relatively high variance, giving the network the best chance of learning. Similarly, the outputs can be reversed back from complex values to real numbers by:

$$y = \frac{\arg(z)}{2\pi} \quad (22)$$

3. Model Establishment of Diffractive Deep Neural Network

The all-optical diffractive deep neural network(D^2NN) was first proposed by Lin Xing in 2018[14], which focused on the application of image classification. The diffractive deep neural network is comprised of multiple diffractive layers, where each diffractive layer is equated to a hidden layer, and each cell on the diffractive layer is analogous to a single neuron in a traditional neural network. The input layer encodes the image into the amplitude or phase channel of the incident field, while the output plane measures the intensity of the field. When a coherent light beam hits the first layer of the network, each point on that layer will serve as a secondary wave source, according to Huygens' principle[15]. Its amplitude and phase depend on the superposed incident waves and are modulated by the transmission coefficient at that point. A neuron in the diffractive neural network is thus accordingly linked to other neurons in the subsequent layer. Obviously, the structure of the entire diffractive deep neural is similar to a fully connected neural network.

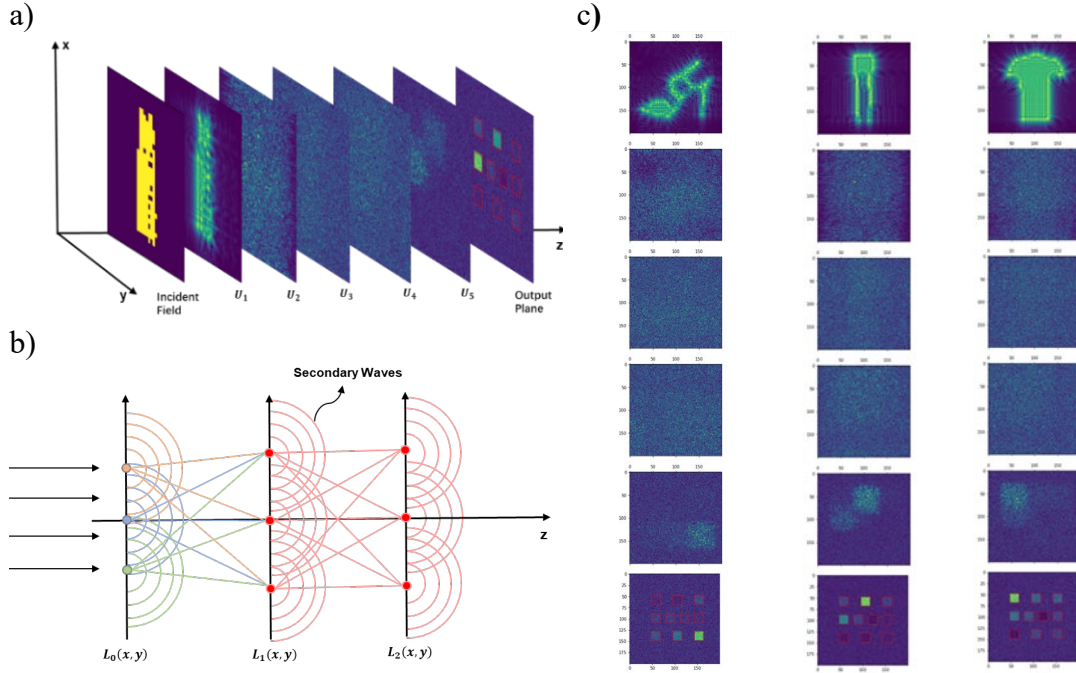


Figure 4. Propagation process of optical diffractive deep neural network

- a) The framework of the three-dimensional D^2NN .
b) The forward propagation model of D^2NN .
c) The classification results of D^2NN (3 examples).

3.1 Architecture of Diffractive Deep Neural Network

According to the Rayleigh-Sommerfeld diffraction equation (Equation 16), when waves propagate to the position (x, y, z) in the l^{th} layer, it will generate the following optical field at point i located at (x_i, y_i, z_i) in the next layer:

$$\begin{cases} U(x_i, y_i, z_i) = w_i^l \cdot U(x, y, z) \cdot \Delta x \Delta y \\ w_i^l = \frac{z_i - z}{r^2} \left(\frac{1}{2\pi r} + \frac{1}{j\lambda} \right) e^{\frac{j2\pi r}{\lambda}} \end{cases} \quad (23)$$

here $\Delta x, \Delta y$ are the quantities of dimensionality and w_i^l is defined as the propagation coefficient. λ is the illumination wavelength and $r = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}$ represents the Euclidean distance between the two points. The process of forward propagation of D^2NN can therefore be summarised as follows:

$$\begin{cases} n_i^{l+1} = t_i^l \cdot m_i^l \\ m_i^l = \sum_k w_k^l n_k^l \cdot \Delta x \Delta y \\ t_i^l = \alpha_i^l e^{j\phi_i^l} \end{cases} \quad (24)$$

where n_i^{l+1} is the output function of the i^{th} neuron located at (x_i, y_i, z_i) and t_i^l is the transmission coefficient of that neuron. The two latent variables α_i^l and ϕ_i^l should be confined to the range $(0, 1)$ and $(0, 2\pi)$ respectively. The forward propagation of the field between two successive layers in D^2NN can be calculated by matrix multiplication[5, 24]:

$$U^{l+1} = W U^l \odot T^l$$

$$\begin{bmatrix} U_1^{l+1}(x_1, y_1, z_1) \\ \vdots \\ U_k^{l+1}(x_i, y_i, z_i) \end{bmatrix} = \begin{bmatrix} w_1(r_1) & \cdots & w_k(r_k) \\ \vdots & \ddots & \vdots \\ w_k(r_k) & \cdots & w_1(r_1) \end{bmatrix} \begin{bmatrix} U_1^l(x_1, y_1, z_1) \cdot \Delta x \Delta y \\ \vdots \\ U_k^l(x_i, y_i, z_1) \cdot \Delta x \Delta y \end{bmatrix} \odot \begin{bmatrix} t_1^l(x_1, y_1, z_1) \\ \vdots \\ t_k^l(x_i, y_i, z_1) \end{bmatrix} \quad (25)$$

Supposing the entire D^2NN consists of M layers, then the detectors measure the intensity of the resulting field in predefined regions:

$$s_i^{M+1} = |m_i^{M+1}|^2 \quad (26)$$

In contrast to the complex-valued neural network, D^2NN works on the physical plane using the interference of waves. The output of each neuron in a CVNN discussed in the last section is $y = \sigma(\omega x)$, where weight ω is the trainable parameter. But for a diffractive deep neural network, the output of a cell is given as $y = \omega x t$, where ω is a hyperparameter determined by diffraction that can be tuned by adjusting the distance between two layers, while the multiplicative bias t is the trainable parameter that can be trained through backpropagation algorithms. The trained neuron' phase value can be mapped onto the different phase delays in terms of physical structures, and the corresponding phase delays are produced by varying the height of the unit cell. In material with refractive index n_1 and height h , the accumulated phase is $\phi_1 = \frac{2\pi n_1}{\lambda} \cdot h$. With the same height in the air, the accumulated phase is $\phi_0 = \frac{2\pi n_0}{\lambda} \cdot h$. The phase difference between the two is $\Delta\phi = \frac{2\pi\Delta n}{\lambda} \cdot h$, and hence the relative height (shown in Figure 5.a) for a unit cell is:

$$\Delta h = \frac{\lambda\phi}{2\pi\Delta n} \quad (27)$$

3.2 Loss Function of Diffractive Deep Neural Network

From the last section, we know it is necessary to construct the loss function (E) based on the optical field intensity distribution. Mean Square Error (MSE)[25] and Softmax Cross Entropy (SCE)[26] are two commonly used loss functions in the field of optical diffractive neural networks. In the following, a comparison and analysis of the two different loss functions are carried out.

Specifically, for a phase-only diffractive deep neural network, the backpropagation algorithm is used to train the distribution of ϕ_i^l . When we use MSE as the loss function, the actual light intensity of the output plane is s_k^{M+1} and the expected light intensity is g_k^{M+1} . Then the MSE loss function can be expressed as:

$$E(\phi_i^l) = \frac{1}{k} \sum_k (s_k^{M+1} - g_k^{M+1})^2 \quad (28)$$

where k refers to the number of measurement points at the output plane.

Compared with the MSE loss function that is based on the light intensity distribution of the entire output plane, the SCE loss function only focuses on the normalised light intensity distribution in the detector region of the output plane, ignoring the light intensity outside the detector region. For the N-class image classification task with N detection regions included, the normalised intensity distribution $\tilde{s}_n$ of the n^{th} detector region can be calculated as:

$$\tilde{s}_n = \frac{s_n}{\max\{s_n\}} \quad (29)$$

Different normalisation methods can also be used before applying the softmax function to get the probability distribution $\hat{s}_n$ for each class. The SCE loss function $E(\phi_i^l)$ can be written as:

$$\begin{cases} \hat{s}_n = \frac{\exp(\tilde{s}_n)}{\sum_n^N \exp(\tilde{s}_n)} \\ E(\phi_i^l) = - \sum_n^N \tilde{g}_n \log(\hat{s}_n) \end{cases} \quad (30)$$

where $\tilde{g}_n$ is the expected normalised field intensity.

Based on the chosen loss function, the optimisation problem for a D^2NN design can be written as:

$$\min_{\phi_i^l} E(\phi_i^l), 0 \leq \phi_i^l \leq 2\pi \quad (31)$$

The weights(phase) of each diffractive layer can be iteratively updated through backpropagation algorithms (i.e. Adam optimiser) so that the loss will continuously drop-down, meaning that the actual intensity distribution is approaching the desired result.

Table 1. Summary of classification accuracy for a 5-layer 200×200 phase-only D^2NN with different loss functions

Loss Function	Dataset	Accuracy
MSE	MNIST	91.75%
SCE	MNIST	97.18%
MSE	Fashion-MNIST	81.13%
SCE	Fashion-MNIST	89.13%

By comparing the normalised intensity distribution of the detector regions, it can be observed that with the softmax cross-entropy loss function, the system sacrifices the signal-to-noise ratio of the output signal but yields higher classification accuracy. In contrast, the mean square error loss function has led to a lower classification performance but a higher signal-to-noise ratio of the output signal. Equation (28) and Equation (30) show the difference between the two loss functions, which is the crucial factor leading to this phenomenon: the MSE affects the intensity distribution among all the detectors while the SCE uses an interclass competition mechanism, focusing on the intensity distribution within the detector area while disregarding the intensity distribution outside of it.

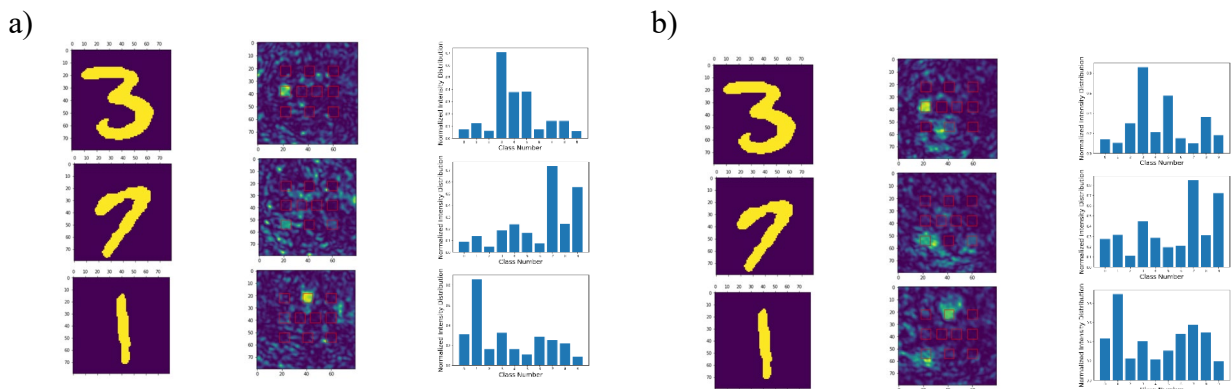


Figure 5. The contrastive analysis between the MSE loss function and SCE loss function

- a) Normalised light intensity distribution based on MSE loss function. b) Normalised light intensity distribution based on SCE loss function.

3.3 Connectivity Analysis of Diffractive Deep Neural Network

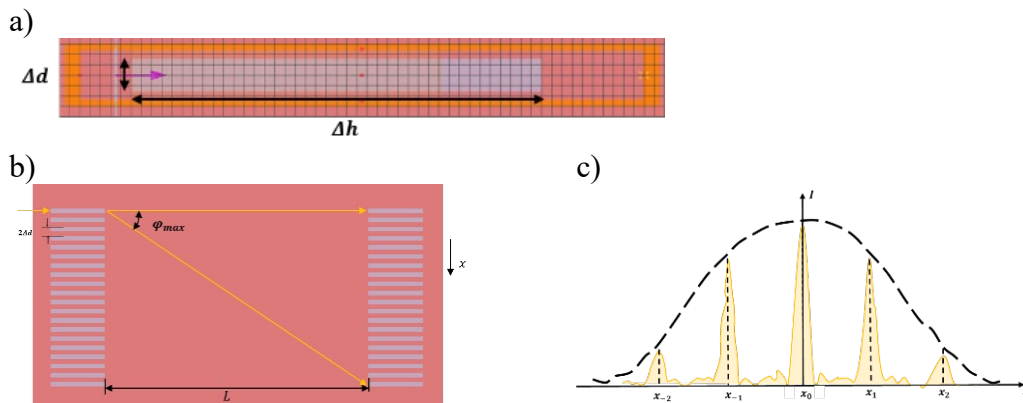


Figure 6. Schematic diagram for connectivity analysis

- a) Micro-structure unit cell of optical diffractive deep neural network.
b) Plane structure of a fully connected optical diffractive deep neural network.
c) Sketch indicating the diffraction pattern.

The fully connected topology is the basis for designing diffractive neural networks[14], which implies that every neuron in a layer can be connected to all neurons in the next layer. In the corresponding optical system, neurons are represented by micro-structure unit cells that have the ability to modulate the incoming light, and the connectivity is realized physically by diffraction. Therefore, the diffraction angle should be designed to be sufficiently large to ensure complete signal transmission between neurons.

When light is normally incident on the micro-structure unit cell, as shown in Figure 6.b, the diffracted light will produce an optical path difference of $2\Delta d \cdot \sin(\varphi_{max})$ which is governed by the grating equations:

$$2\Delta d \cdot \sin(\varphi_{max}) = m \cdot \lambda \quad (32)$$

$$x_m = m \cdot \frac{\lambda L}{2\Delta d} \quad (33)$$

where Δd is the feature size of a unit cell, $m = 0, \pm 1, \pm 2, \dots$ is the integer denoting the diffraction order[16, 17]. The maximum diffraction angle φ_{max} for the optical diffractive deep neural network can be considered when $m = 1$:

$$2\Delta d \cdot \sin(\varphi_{max}) = \lambda \quad (34)$$

The above equation can be rewritten as:

$$\varphi_{max} = \sin^{-1}\left(\frac{\lambda}{2\Delta d}\right) \quad (35)$$

The maximum diffraction angle for the optical diffractive deep neural network φ_{max} thereby can be qualitatively described by Equation (35) with two parameters: λ and Δd . φ_{max} is an indicator of the connectivity between two adjunct diffractive layers, which is proportional to λ and inversely proportional to Δd . Thus, as a general rule, the longer wavelength of the light source and the smaller feature size of the unit cell will provide more sufficiency to achieve a fully connected diffractive deep neural network in an actual optical environment. Additionally, in the scenario with fixed wavelength, the feature size of the unit cell is reduced to the scale of sub-wavelength($\approx \frac{1}{2}\lambda$) or even smaller to realise a 90° maximum diffraction angle:

$$\varphi_{max} = \sin^{-1}\left(\frac{\lambda}{2\Delta d}\right) = \sin^{-1}\left(\frac{\lambda}{2 \cdot (\frac{1}{2}\lambda)}\right) = \sin^{-1}(1) = 90^\circ \quad (36)$$

4. Research Status of Optical Diffractive Neural Networks

In subsequent studies, researchers have built on the D^2NN framework to investigate the range of wavelengths which can be applied to optical diffractive neural networks and their performance on different tasks such as image classification.

Table 2. System design parameters of diffractive deep neural networks at different wavelengths

Source and Publication Year	The wavelength of the Light Source, λ	Neuron Feature Size, Δd
Cui T J et al., Nature Electronics[27], 2022	55.5 mm	20.76 mm ($\approx \frac{1}{2}\lambda$)
Qian C et al., Light: Science & Applications[28], 2020	17.6 mm	10 mm ($\approx \frac{1}{2}\lambda$)
Aydogan Ozcan et al. Science advances[29], 2021	≈ 1 mm	0.5 mm ($\approx \frac{1}{2}\lambda$)
Lin X et al., Science[14], 2018	0.75 mm	0.4 mm ($\approx \frac{1}{2}\lambda$)
Zeng Z M et al., IEEE Photonics Technology Letters, 2019	10.6 μ m	5 μ m ($\approx \frac{1}{2}\lambda$)

The terahertz band is the distinctive frequency range where optical diffraction neural networks were first employed. Once the optical diffractive neural network has been trained, it can be manufactured using nanofabrication technologies. The difficulty of fabrication depends on the feature size of the micro-structure unit cell (the neuron in the optical system), which is generally regulated by Equation (35). As the size of the unit cell increases, the fabrication process becomes more accessible. Subsequently, researchers have broadened the light source for the diffractive neural network to visible and infrared wavelengths. The former makes an effort to reduce the setup complexity of the optical system, while the latter steps into the miniaturisation of the system.

These studies have extended the range of light sources in establishing the optical system and enhanced its performance. However, a diffractive deep neural network based on the long wavelength source is expensive and bulky and can be strongly attenuated by the vapour in the air. [30] Also, as the systems mentioned above use a coherent light source, server interference noise exists in the entire system. The interference noise will cause diffractions to interfere with each other during propagation, resulting in a series of chaotic and disorganised stripes. These stripes can even drown out the desired signals in severe cases.

Chromatic optical diffractive neural networks can effectively suppress the coherent noise caused by interference because the output is the summation of light intensity distributions that are uncorrelated in a narrow band. Though this chromatic system has higher information dimensionality bringing the prospect of multi-channel computing, it weakens the capability of processing complex information because all the computations are based on onefold light intensity.[31, 32]

5. Integrated Diffractive Deep Neural Network

With integrated photonic circuits, light can be confined and propagated through structures on a plate to perform computing. It is a scalable and mature technology available that can do hybrid integration of devices and material technologies. [33] A diffractive deep neural network can be integrated physically on a standard silicon-on-insulator (SOI) substrate. The performance of ID^2NN works [34, 35] proposed in recent years have demonstrated machine learning task of regression and image classification tasks and achieved comparable results to the state-of-the-art.

A diffractive neural network implemented on photonic waveguides or photonic chips is characterized by comprising an input layer, a network layer, and an output layer. The input layer is used to collimate light and encode the information into light signals, which are then loaded into the network; the network layer is a diffraction transmission zone and is used to perform phase adjustment on the received light signals; the output layer is used to merge the light through tapers whose intensity distribution at the end indicate the classification result.

For the guided wave system, the wavelength of light inside the medium depends on the refractive index of the medium:

$$\lambda = \frac{\lambda_0}{n} \quad (37)$$

where λ_0 is free space wavelength, n is the refractive index of the guided material, and λ is the wavelength in the material.[36] The remaining design methodology and mathematical framework are consistent with the discussion in section 2.1.4.

Here, we point out that compared with D^2NN realized in free space, and the integrated diffractive neural network is smaller in size, which has excellent potential application value in other machine learning models(i.e. unsurprised learning, reinforcement learning)[37, 38]. In order to effectively integrate the optical diffractive neural network into a chip, it is necessary to gain a thorough understanding of their non-integrated implementation in free space. The advancement of material technology also allows for the incorporation of non-linear optical materials that have the potential to achieve an activation-like behaviour.

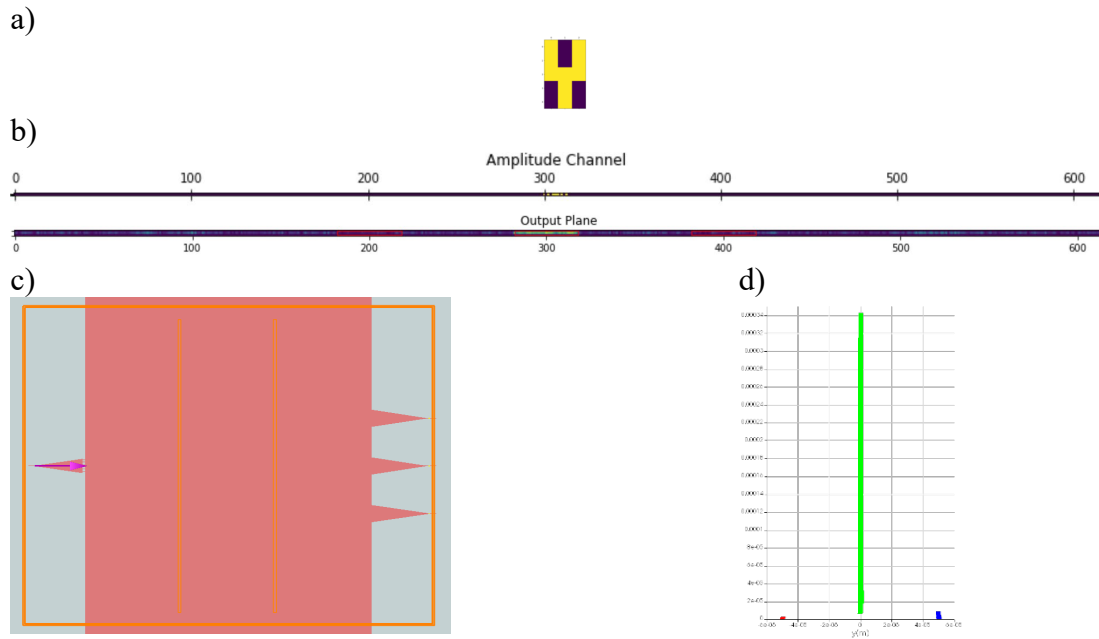


Figure 7. A comparison between FDTD simulation result and python-based simulation result

- a) The input object: Letter "Y".
- b) Flatten input to ID^2NN and the field intensity distribution from the output plane of ID^2NN .
- c) Schematic diagram for ID^2NN in FDTD.
- d) Power distribution in the output tapers.

6. Research Problems of Integrated Diffractive Deep Neural Network

Optical diffractive neural networks have seen significant growth and success since their inception, with researchers in the field achieving a wealth of results. However, as this field is relatively new, much of the research is still in its early stages. This section highlights two main scientific problems in the field.

Firstly, the design approach for physical system parameters in the optical diffractive deep neural network is based on the theory of maximum diffraction angle that lacks quantitative analysis methods. The feature size of the neurons in the system is determined by the order of working wavelength, with an inverse relationship controlled by Equation (35). However, the theory of traditional maximum diffraction angle can only qualitatively describe the relationship between the two, suggesting that reducing the feature size of unit cells to $\frac{1}{2}\lambda$ can lead to a successful optical computing system. When the system is integrated into a chip, the connecting dimensionality is reduced. To maintain a high-accuracy computing performance in processing complex information, the number of neurons in a single layer will significantly increase. Therefore, it is vital to have a theoretical model that quantitatively describes the constraints between system parameters to ensure an optimal design.

The second problem is the technical challenge in nanofabrication. The use of coherent light sources leads to severe interference noise. Increasing the number of I/O ports in the system while suppressing interference noise to maintain high performance requires a high level of accuracy in the fabrication process.[35]

7. Conclusion and Outlook

In this review, we introduce the implementation of an optical diffractive deep neural network in free space and silicon-on-insulator and extract two significant research and design problems. Optical diffractive neural networks have the ability to perform complex tasks at ultrafast speed. However,

their training and optimization still rely on the neural network framework, which is implemented through computer-based systems. Developing a fully-optical training framework, which bypasses the need for a computer, would be of significant importance for the advancement of optical diffractive neural network research. The classification accuracy of images for an optical diffractive neural network is still lower than electronic ones. Hence, improving the classification accuracy of optical diffraction neural networks while maintaining their other advantages could be a key focus for future development.

References

- [1] Deep learning | Nature [online]. [vid. 2022-12-21]. Dostupné z: [https:// www. nature. com/ articles/ nature 14539](https://www.nature.com/articles/nature14539).
- [2] Helping robots see the big picture | Science [online]. [vid. 2022-12-21]. Dostupné z: [https:// www. science. org/ doi/full/10.1126/science.346.6206.186](https://www.science.org/doi/full/10.1126/science.346.6206.186).
- [3] Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups | IEEE Journals & Magazine | IEEE Xplore [online]. [vid. 2022-12-21]. Dostupné z: [https:// ieeexplore. ieee. org/ document/6296526](https://ieeexplore.ieee.org/document/6296526).
- [4] A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots | IEEE Journals & Magazine | IEEE Xplore [online]. [vid. 2022-12-21]. Dostupné z: [https:// ieeexplore. ieee. org/ document/ 7358076](https://ieeexplore.ieee.org/document/7358076).
- [5] CANZIANI, Alfredo, Adam PASZKE a Eugenio CULURCIELLO. An Analysis of Deep Neural Network Models for Practical Applications [online]. B.m.: arXiv. 14. duben 2017 [vid. 2022-12-21]. Dostupné z: [doi:10.48550/arXiv.1605.07678](https://doi.org/10.48550/arXiv.1605.07678). arXiv:1605.07678 [cs].
- [6] Tunable lifetime nanocrystals | Nature Photonics [online]. [vid. 2022-12-21]. Dostupné z: [https:// www. nature. com/articles/nphoton.2013.353](https://www.nature.com/articles/nphoton.2013.353).
- [7] YAN, Tao, Jiamin WU, Tiankuang ZHOU, Hao XIE, Feng XU, Jingtao FAN, Lu FANG, Xing LIN a Qionghai DAI. Fourier-space Diffractive Deep Neural Network. Physical Review Letters [online]. 2019, 123 (2), 023901. Dostupné z: [doi:10.1103/PhysRevLett.123.023901](https://doi.org/10.1103/PhysRevLett.123.023901).
- [8] WALDROP, M. Mitchell. The chips are down for Moore's law. Nature News [online]. 2016, 530(7589), 144. Dostupné z: [doi:10.1038/530144a](https://doi.org/10.1038/530144a).
- [9] EFNUSHEVA, Danijela, Ana CHOLAKOSKA a Aristotel TENTOV. A SURVEY OF DIFFERENT APPROACHES FOR OVERCOMING THE PROCESSOR-MEMORY BOTTLENECK. International Journal of Information Technology and Computer Science [online]. 2017, 151. Dostupné z: [doi:10.5121/ijcsit.2017.9214](https://doi.org/10.5121/ijcsit.2017.9214).
- [10] Beyond von Neumann. Nature Nanotechnology [online]. 2020, 15(7), 507–507. ISSN 1748-3395. Dostupné z: [doi:10.1038/s41565-020-0738-x](https://doi.org/10.1038/s41565-020-0738-x).
- [11] SOLLI, Daniel R. a Bahram JALALI. Analog optical computing. Nature Photonics [online]. 2015, 9(11), 704–706. ISSN 1749-4893. Dostupné z: [doi:10.1038/nphoton.2015.208](https://doi.org/10.1038/nphoton.2015.208).
- [12] SHEN, Yichen, Nicholas C. HARRIS, Scott SKIRLO, Mihika PRABHU, Tom BAEHR-JONES, Michael HOCHBERG, Xin SUN, Shijie ZHAO, Hugo LAROCHELLE, Dirk ENGLUND a Marin SOLJAČIĆ. Deep learning with coherent nanophotonic circuits. Nature Photonics [online]. 2017, 11(7), 441–446. ISSN 1749-4893. Dostupné z: [doi:10.1038/nphoton.2017.93](https://doi.org/10.1038/nphoton.2017.93).
- [13] PSALTIS, Demetri, David BRADY, Xiang-Guang GU a Steven LIN. Holography in artificial neural networks. Nature [online]. 1990, 343(6256), 325–330. ISSN 1476-4687. Dostupné z: [doi: 10.1038/ 3433 25a0](https://doi.org/10.1038/343325a0).
- [14] All-optical machine learning using diffractive deep neural networks | Science [online]. [vid. 2022-11-25]. Dostupné z: <https://www.science.org/doi/10.1126/science.aat8084>.
- [15] BORN, Max. Electromagnetic theory of propagation, interference and diffraction of light. nedatováno.
- [16] Goodman, J.W. (1996). Introduction to Fourier Optics. B.m.: McGraw-Hill Science, Engineering & Mathematics, nedatováno.

- [17] Vasudevan Lakshminarayanan, Hassen Ghalila, Ammar, A. and L. Srinivasa Varadharajan (2018). Understanding Optics with Python. B.m.: CRC Press, nedatováno.
- [18] David George Voelz (2010). Computational Fourier optics: a MATLAB tutorial. B.m.: Bellingham, Wash.: Spie Press, nedatováno.
- [19] BASSEY, Joshua, Lijun QIAN a Xianfang LI. A Survey of Complex-Valued Neural Networks [online]. B.m.: arXiv. 28. leden 2021 [vid. 2022-12-03]. Dostupné z: doi:10.48550/arXiv.2101.12249. arXiv: 2101.12249 [cs, stat].
- [20] MEHRABKHANI, Soheil a Thomas SCHNEIDER. Is the Rayleigh-Sommerfeld diffraction always an exact reference for high-speed diffraction algorithms? Optics Express [online]. 2017, 25(24), 30229. ISSN 1094-4087. Dostupné z: doi:10.1364/OE.25.030229.
- [21] BUITRAGO-DUQUE, Carlos a Jorge GARCIA-SUCERQUIA. Non-approximated Rayleigh-Sommerfeld diffraction integral: advantages and disadvantages in the propagation of complex wave fields. Applied Optics [online]. 2019, 58(34), G11–G18. ISSN 2155-3165. Dostupné z: doi:10.1364/AO.58.000G11.
- [22] Introduction to artificial neural networks | IEEE Conference Publication | IEEE Xplore [online]. [vid. 2022-12-21]. Dostupné z: <https://ieeexplore.ieee.org/document/483329>.
- [23] A Hybrid Training Algorithm for Feedforward Neural Networks | SpringerLink [online]. [vid. 2022-12-21]. Dostupné z: <https://link.springer.com/article/10.1007/s11063-006-9013-x>.
- [24] MOSS, David. Photonic multiplexing architectures for optical neuromorphic networks [online]. 2022 [vid. 2022-12-21]. Dostupné z: <https://hal.archives-ouvertes.fr/hal-03764240>.
- [25] MARMOLIN, Hans. Subjective MSE measures. IEEE Transactions on Systems, Man, & Cybernetics [online]. 1986, SMC-16, 486–489. ISSN 0018-9472. Dostupné z: doi:10.1109/TSMC.1986.4308985.
- [26] CSIRO Research Publications Repository - On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation functions [online]. [vid. 2022-12-21]. Dostupné z: <https://publications.csiro.au/rpr/pub?list=BRO&pid=procite:bflf60dc-ea12-49a7-a14d-79b1419a475e>.
- [27] A programmable diffractive deep neural network based on a digital-coding metasurface array | Nature Electronics [online]. [vid. 2022-12-15]. Dostupné z: <https://www.nature.com/articles/s41928-022-00719-9>.
- [28] Performing optical logic operations by a diffractive neural network | Light: Science & Applications [online]. [vid. 2022-12-15]. Dostupné z: <https://www.nature.com/articles/s41377-020-0303-2>.
- [29] Spectrally encoded single-pixel machine vision using diffractive networks | Science Advances [online]. [vid. 2022-12-21]. Dostupné z: <https://www.science.org/doi/10.1126/sciadv.abd7690>.
- [30] LU, Lidan, Zhoumo ZENG, Lianqing ZHU, Qiankun ZHANG, Bofei ZHU, Qifeng YAO, Mingxing YU, Haisha NIU, Mingli DONG a Guoshun ZHONG. Miniaturized Diffraction Grating Design and Processing for Deep Neural Network. IEEE Photonics Technology Letters [online]. 2019, 31(24), 1952–1955. ISSN 1041-1135, 1941-0174. Dostupné z: doi:10.1109/LPT.2019.2948626.
- [31] LABORDE, Victor, Jérôme LOICQ, Juriy HASTANIN a Serge HABRAKEN. Multilayer diffractive optical element material selection method based on transmission, total internal reflection, and thickness. Applied Optics [online]. 2022, 61(25), 7415–7423. ISSN 2155-3165. Dostupné z: doi:10.1364/AO.465999.
- [32] PSALTIS, Demetri. Coherent optical information systems. Science (New York, N.Y.) [online]. 2002, 298(5597), 1359–1363. ISSN 1095-9203. Dostupné z: doi:10.1126/science.1078823.
- [33] Chrostowski, L. and Hochberg, M. (2015). Silicon Photonics Design. B.m.: Cambridge University Press, nedatováno.
- [34] FU, Tingzhao, Yubin ZANG, Honghao HUANG, Zhenmin DU, Chengyang HU, Minghua CHEN, Sigang YANG a Hongwei CHEN. On-chip photonic diffractive optical neural network based on a spatial domain electromagnetic propagation model. Optics Express [online]. 2021, 29(20), 31924–31940. ISSN 1094-4087. Dostupné z: doi:10.1364/OE.435183.
- [35] LUO, Xuhao, Yueqiang HU, Xiangnian OU, Xin LI, Jiajie LAI, Na LIU, Xinbin CHENG, Anlian PAN a Huigao DUAN. Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible. Light:

Science & Applications [online]. 2022, 11(1), 158. ISSN 2047-7538. Dostupné z: doi:10.1038/s41377-022-00844-2.

- [36] COLDREN, Larry A. Diode Lasers and Photonic Integrated Circuits. In: Optical Engineering [online]. 1997, s. 616 [vid. 2022-12-21]. ISSN 0091-3286. Dostupné z: doi:10.1117/1.601191.
- [37] BARLOW, H. B. Unsupervised learning. Neural Computation [online]. 1989, 1, 295–311. ISSN 1530-888X. Dostupné z: doi:10.1162/neco.1989.1.3.295.
- [38] BUENO, J., S. MAKTOOBI, L. FROEHLI, I. FISCHER, M. JACQUOT, L. LARGER a D. BRUNNER. Reinforcement learning in a large-scale photonic recurrent neural network. Optica [online]. 2018, 5(6), 756–760. ISSN 2334-2536. Dostupné z: doi:10.1364/OPTICA.5.000756.