A Comprehensive Survey on Nanophotonic Neural Networks: Architectures, Training Methods, Optimization, and Activations Functions

Konstantinos Demertzis [1,2,*], Georgios D. Papadopoulos [1], Lazaros Iliadis [3] and Lykourgos Magafas [1]

[1] Department of Physics, Faculty of Sciences, Kavala Campus, International Hellenic University, St. Loukas, 654 04 Kavala, Greece; gebapad@teiemt.gr (G.D.P.); lmagafas@otenet.gr (L.M.)

[2] School of Science & Technology, Informatics Studies, Hellenic Open University, 263 35 Patra, Greece

[3] School of Civil Engineering, Faculty of Mathematics Programming and General Courses, Democritus University of Thrace, Kimmeria, 691 00 Xanthi, Greece; liliadis@civil.duth.gr

[*] Correspondence: kdemertz@fmenr.duth.gr or kdemertzis@teiemt.gr

Abstract: In the last years, materializations of neuromorphic circuits based on nanophotonic arrangements have been proposed, which contain complete optical circuits, laser, photodetectors, photonic crystals, optical fibers, flat waveguides and other passive optical elements of nanostructured materials, which eliminate the time of simultaneous processing of big groups of data, taking advantage of the quantum perspective, and thus highly increasing the potentials of contemporary intelligent computational systems. This article is an effort to record and study the research that has been conducted concerning the methods of development and materialization of neuromorphic circuits of neural networks of nanophotonic arrangements. In particular, an investigative study of the methods of developing nanophotonic neuromorphic processors, their originality in neuronic architectural structure, their training methods and their optimization was realized along with the study of special issues such as optical activation functions and cost functions. The main contribution of this research work is that it is the first time in the literature that the most well-known architectures, training methods, optimization and activations functions of the nanophotonic networks are presented in a single paper. This study also includes an extensive detailed meta-review analysis of the advantages and disadvantages of nanophotonic networks.

Keywords: nanophotonic neural networks; photonic neural networks; optical neural networks; optical interference unit; optical non-linear unit; optical activation function

1. Introduction

Artificial intelligence (AI) [1] enables machines to be trained so as to perform particular tasks, learn from experience, adapt to or interact with the environment and perform realistic anthropomorphic tasks [2]. Contemporary AI is one of the fastest evolving fields of information technology, in which high-level algorithmic approaches and tools descending from applied math's and engineering are used [3–5]. Most AI applications—from computers playing chess to automatically driven cars—are based to a great extent on the intelligent technologies of neural networks (NNs) [6] for the processing of multidimensional big data, with a view to revealing the hidden knowledge that is included in these groups [7,8].

In classic von Neumann architecture, where the computations are restrained by the speed of the channel between computation and memory (also known as the von Neumann congestion), even the important innovations on problems such as the shrinking of complete circuits, the reduction in their power needs and the decrease of temperature emitted by them cannot achieve the anticipated increases in their computing power [9,10]. Even with the introduction of a graphics processing unit (GPU) as an extra processor for the improvement of graphic interface and the performance of tasks of high-level processing, or

the introduction of Google's tensor processing unit (TPU) as the most powerful adapted processor for the performance of AI procedures, the capabilities of traditional systems seem unable to cope with the demands of modern technology and the uninterrupted flow of the data produced, even if they have offered some of the most important innovations [10–12].

The biggest challenge in this field is the development of fully functional and utilizable neuromorphic systems-on-chip (NSoC) [10,13], which will be able to approach the biological human intelligence, performing the same tasks that the human brain effortlessly performs in no time at all and without remarkable consumption of resources and energy [14]. The neuromorphic computation comprises the creation of neural networks in matter, where the neurons of a physical device are connected with the corresponding synapses of physical devices [15]. The main motive for the neuromorphic computation is the time needed to process the computations and the energy performance provided by a distributed architecture, which avoids the energy turbulence of data between the memory and the CPU [16].

The NSoCs based solidly on previous computational technology overcome the von Neumann congestion, massively use simultaneous computational procedures and are tolerant to faults [17]. Essentially, they form the way in which neural networks function, conveying information in the same temporal and spatial way as the human brain. Furthermore, taking advantage of techniques such as the memristors [10], they are capable of modeling learning skills; that is, the adjustment ability of synapses in storing and conveying information depending on the evolution of a dynamic situation [16].

However, the most important development in the application efforts and standardization of NSoCs is spotted in the expanded efforts for developing fully optical neural networks (ONNs), also known as photonic neural networks (PNNs) or nanophotonic neural networks (NNNs) [15,18–21]. The previously mentioned systems are based on the evolutions of optical technology and the most recent research concerning photonics. Photonics is the science and technology field that deals with the creation, control and detection of photons, especially in the area of visible light and the near-infrared electromagnetic spectrum (wavelength, polarization, transmission rate, etc.) and the great potentials of their interconnection [20,22,23]. It is directly related, in basic as well as applied research, to quantum optics as to how linear transformations can be applied with the minimum energy consumption and with the slightest latency time on neuron level and with optoelectronics in the study of active and passive materials that interact electrically with light [24,25].

Many efforts have been made in recent years to shift from conventional electronics to optical circuits. This review records the most recent research to clarify how close we are to the complete transition to photonic arrangements and their exceptional prospects. Moreover, the main contribution of this research work is that it is the first time in the literature that the most well-known architectures, training methods, optimization, and activations functions of the nanophotonic networks are presented in a single paper. Additionally, the manuscript includes an extensive detailed meta-review analysis of the advantages and disadvantages of nanophotonic networks.

The rest of the paper is organized as follows: Section 2 contains the principles of light and matter interaction. Section 3 describes the current state of research in neuromorphic processors using photon circuits and Section 4 analyzes these architectures. Section 5 is dedicated to the training operation of PNNs. Section 6 summarizes the most common activation functions that are used. In conclusion, Section 7 presents the final remarks and perspectives.

2. Nature of Light

In general, the interaction of light with matter and its diffusion inside it, is sufficiently described by optics (geometric and wave optics) [22,26]. While studying various optic phenomena when the intensity of impacting radiations is small, the response in general and all the individual optic properties/parameters of materials (e.g., refraction index, absorption factor, polarization, etc.) remain stable and independent of intensity [27]. However, when

the intensity of radiation is high, as it happens with laser, for example, and in particular with a focused-beam laser of great power, it has been proved experimentally that the optic response of matter and the optical parameters are modified, often significantly, and become dependent on the intensity of the radiation [28]. It is then that various extremely interesting phenomena take place, which are not detected in the case of low intensity impacting radiation [29]. These phenomena cannot be interpreted considering the linear response of matter, as it is expressed in the fundamental linear relation between the cause, i.e., the *E* electric field, and the result, i.e., the *P* inductive polarization [22,30,31]:

$$\vec{P} = \varepsilon_0 \chi^{(1)} \vec{E} \tag{1}$$

where $\chi^{(1)}$ is the linear susceptibility and ε_0, ε is the intra-electrical invariant of vacuum and matter, respectively. For the susceptibility $\chi^{(1)}$ and the refraction index *n*, it is true that [24,28]:

$$n^2 = \frac{\varepsilon}{\varepsilon_0} = 1 + \chi^{(1)} \tag{2}$$

In the case of high-intensity radiation, conditions of high rate appear in the expression of polarization, the contributions of which are essential and cannot be omitted. These phenomena are the result of modification of the optic properties of matter due to the powerful electric field, and the non-linearity of these phenomena is attributed to the fact that the response of matter is a non-linear function of the intensity of radiation [32].

One of the most important consequences of the linear response of matter under the influence of intense fields, apart from the alterations of the properties of matter, is that if there are different beams going through the same region of a non-linear medium simultaneously, they can interact with one another through matter [30]. Going this reasoning one-step further and taking into consideration the principle of superposition, according to which a beam can be considered as the superposition of two beams of the same polarization, frequency and direction, we can assume that a beam can interact with itself [25,26,28].

As far as the procedure that causes the appearance of a non-linear optic behavior is concerned, when some radiation impacts on some material, it causes changes in the spatial and temporal distribution of the electric charge, inducing electric dipoles, the macroscopic result of which is the creation or modification of the polarization of the material [33]. For low values of the *E* field, the *P* polarization is analogous to the *E* field that caused it and the elementary dipoles, when oscillated, emit radiation of the same properties as the impacting radiation. Nevertheless, for high intensities of the *E* field, the radiation emitted by the elementary dipoles is not in correspondence with the electric *E* field that caused it. This can be explained by the fact that the captivated electrons of atoms/ions/unitary cells of crystal (or the structural unit in general) are forced to great displacements from their balance position. As a result, the motion of electrons cannot be described by the model of harmonic oscillator of Lorentz. Then, the radiation emitted contains frequencies different from those of the initial stimulating radiation. This practically means that it is possible to modify the impacting radiation itself with the addition of new frequencies, for example. In this way, non-linear phenomena can be interpreted and applied, which can explain how a beam of light can interact with one another (or with itself) creating amplification of light through light, merging of a beam with another one, production of new frequencies, etc. [34].

Based on what was previously mentioned, for high intensity of the electric field (e.g., for $E > 10^5$ V/cm) where the presence of non-linear phenomena becomes significant, in the equation describing the polarization conditions of a higher rate appear, and the polarization is presented as an expansion of the Taylor sequence according to the following form [33,35]:

$$\vec{P}(t) = \varepsilon_0 \left[\chi^{(1)} \vec{E}(t) + \chi^{(2)} \vec{E}^2(t) + \chi^{(3)} \vec{E}^3(t) + \cdots \right] \tag{3}$$

where $\chi^{(2)}$ is the second-rate susceptibility, $\chi^{(3)}$ is the third-rate susceptibility and so on and so forth. The susceptibilities are generally tensors, so, for instance, the first-rate

susceptibility $\chi^{(1)}$ is a second-rate tensor with $3 \times 3 = 9$ elements and the corresponding polarization is presented by the following form [25,34,35]:

$$
\begin{bmatrix} P_x^{(1)} \\ P_y^{(1)} \\ P_z^{(1)} \end{bmatrix} = \varepsilon_0 \begin{bmatrix} \chi_{xx}^{(1)} & \chi_{xy}^{(1)} & \chi_{xz}^{(1)} \\ \chi_{yx}^{(1)} & \chi_{yy}^{(1)} & \chi_{yz}^{(1)} \\ \chi_{zx}^{(1)} & \chi_{zy}^{(1)} & \chi_{zz}^{(1)} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \text{ or } P_i^{(1)} = \varepsilon_0 \sum_j \chi_{ij}^{(1)} E_j \tag{4}
$$

where $i, j = x, y, z$.

Similarly, the non-linear second-rate susceptibility $\chi^{(2)}$ is a third-rate tensor $\chi_{ijk}^{(2)}$, whereas the third-rate susceptibility $\chi^{(3)}$ is a fourth-rate tensor $\chi_{ijkl}^{(3)}$. In the case that the medium displays losses, the susceptibility $\chi^{(1)}$ is a complex quantity with its real part being connected to the linear refraction index n and its imaginary part being connected to the factor of linear absorption through the following relations [21,22,26]:

$$
\chi^{(1)} = Re\left(\chi^{(1)}\right) + i\left[Im\left(\chi^{(1)}\right)\right] \tag{5}
$$

where $Re\left(\chi^{(1)}\right) \propto n_0$ and $Im\left(\chi^{(1)}\right) \propto \alpha_0$.

The equivalent relations also apply to the non-linear high-rate susceptibilities, which are also complex numbers with real and imaginary parts equivalent to the corresponding refraction indexes and absorption factors, which correspond to the non-linear refraction index and the non-linear absorption factor. When an intense laser beam passes through a material, the electric field of the beam can induce a change in the refractive index of the material that is proportional to the intensity of the beam. This non-linear effect is called the Kerr effect. The total refractive index of the material is the sum of the refractive index, n_0, with no laser beam present and the term $n_2 I$, where n_2 is the second-order non-linear refractive index and I is the intensity of the beam. The change in refractive index can be positive or negative.

It is also important to point out that the calculation of an observed value in a system of photonic arrangements disrupts the system and, therefore, it shifts to a quantum condition in which the repetition of calculations of the same property leads to the same result. Thus, the following quantum conditions are observed [21,24,36]:

(1) $|E\rangle$: Quantum condition where, if power is calculated, the result will be E.
(2) $|p\rangle$: Quantum condition where, if momentum is calculated, the result will be p.
(3) $|x\rangle$: Quantum condition where, if position is calculated, the result will be x.

In a general condition $|\psi\rangle$, the possibilities of calculating various physical properties are uncertain, that is there is the possibility $P_1(E)$ of calculating the value of energy as E, $P_2(p)$ is the possibility of calculating the value of momentum as p and so on and so forth. In a $|\psi\rangle$, condition system, after the calculation, for example, of energy with an E_1, result, the wave function is disrupted and collapses (transforms) into a new condition $|E_1\rangle$, so that the repetition of the same calculation gives the same result. Respectively, in a $|\psi\rangle$, condition system, after calculating for example the momentum with the result p_1, the wave function is disrupted and collapses (transforms) into a new condition $|p_1\rangle$, so that the repetition of the same calculation gives the same result [37,38].

The conditions $|x\rangle$ and $|p\rangle$ cannot coincide, because the calculation of position (e.g., with photon scattering of short wavelength) alters the momentum. Consequently, there is no certainty about the momentum and the position of a particle as the values observed are random variables, in the sense that for every value of the spectrum of an observed physical quantity, a quantum width of probability for the calculation of this particular value corresponds to it. The total amount of quantum widths of probability for a spectrum of an observed physical quantity fully determines the quantum condition of the system. In that

sense, one of the targets of photonic systems is the calculation of these widths of probability with the use of results of analytic methods [39,40].

In conclusion, taking advantage of the binary nature of light and all the other characteristics that render it the fastest means of communication, the verge of modern investigation is focused on developing photonic neuromorphic processors [38,41].

3. Photonic Neuromorphic Processors

The investigation into the development of photonic neuromorphic processors with passive optic circuits, focuses on advantages such as the following [21,22,24]:

(1) Significant reduction of energy consumption in the applications of logical circuits as well as in data transfer.
(2) Exceptionally high operating speeds with no energy consumption other than on the transmitters and the receivers.
(3) Distribution of the computing power in the whole network, with each neuron performing simultaneously small parts of the whole computational activity.

Nevertheless, a big obstacle of photonic circuits has been the great volume of optical devices and the absence of susceptibility in contrast to the traditional integrated electronic circuits [13,18,21]. The materializations mainly on silicon (Si) and additionally on indium phosphide (InP) constitute a great innovative breakthrough in the materialization of integrated photonics, which is a reality nowadays. Especially, the materializations of integrated photonics with silicon as the construction material have proved to be excellent as they are transparent for wavelengths of 1270 to 1625 nm that are used in communications, and the refraction index with a breadth of 3.48 in 1550 nm guarantees great resistance, while at the same time it can be checked thermally, electrically, mechanically or chemically [16,17,32].

Taking advantage of the properties above, silicon has been widely used for the materializations of passive elements, like waveguides, modulators, splitters, couplers and filters. On the other hand, indium phosphide allows for the materialization of monolithically integrated solutions, which include a combination of passive and active devices, such as lasers and amplifiers. Moreover, the integrated photonic technology offers the prospect of reducing the order of magnitude of the integration into nano-levels with all the significant advantages that the aforementioned reduction of size brings about, such as the reduction of energy footprint, smaller size, etc. [40,42].

Therefore, a multitude of arrangements of optical spare parts and integrated photonic circuits is already available, which results in the appearance of significant progress in the field of photonic neuromorphic, with the development of nanophotonic neural networks with either applications or waveguides or free-space optic [20,36,43].

4. Architectures

The philosophy behind the use of photonic circuits [22,44,45] in nano-arrangements [46–48] and materials is based on the need for significant improvement of the speed of transmitting and processing data and for an improvement of the energy efficiency of devices. The PNN materializations based on the aforementioned materials, which are present to this day, are classified into two main categories: with memory (stateful) and without memory (stateless), as they are concisely presented in Figure 1 [18,21,49]:

Moreover, the PNNs are classified according to design (integrated or free-space optic) and according to optical training ability (trainable) or reference only (inference). The types of networks that have been applied in modern neuromorphic technology and their respective modifications are thoroughly explained below [18,49].
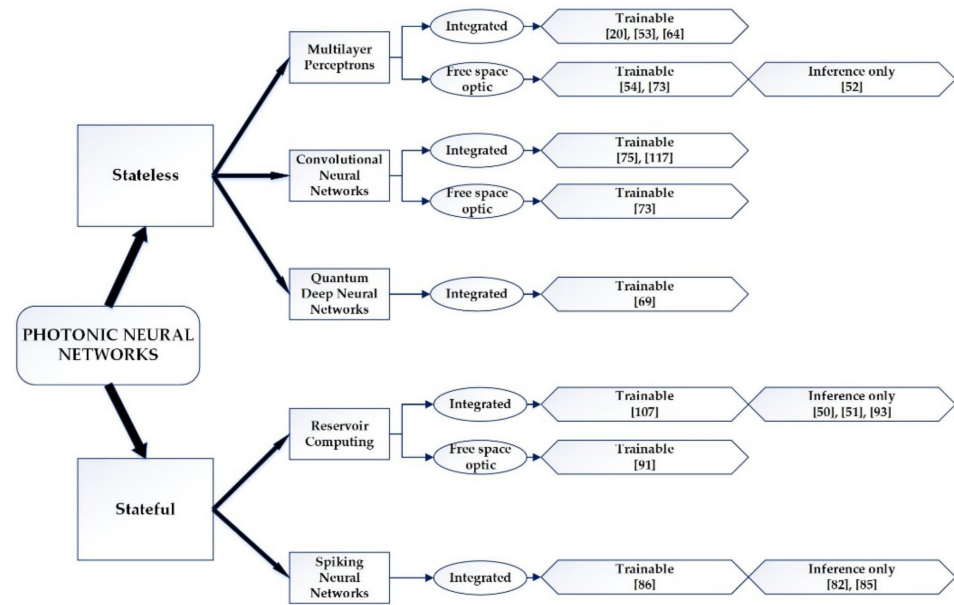
Figure 1. Photonic neural networks classification according to their architecture (stateless or stateful), their design (integrated or free-space optic) and their training ability, presented until 2019.

4.1. Perceptron

The perceptron model consists of a single neuron, is the simplest autonomous system in existence and performs a particular task. This unique neuron of the system has a particular number of connections deriving from other neurons. The perceptron's development into ONNs is the most fundamental scientific field, with many articles having been published with respective materializations [20,50–53]. An all-optical neural network (AONN) architecture with a hidden layer is presented in Figure 2 [54].
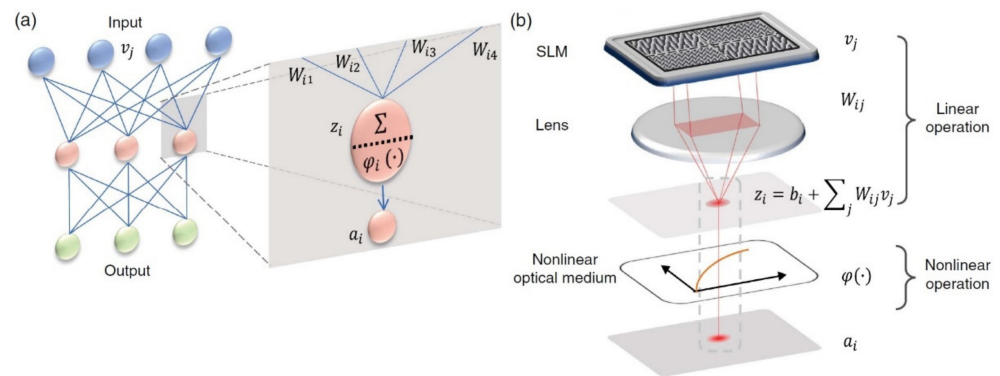

Figure 2. (**a**) A neural network with two layers and a detailed view of one of its neurons. (**b**) Implementation of an optical neuron with linear operation (SLM and lens units) and non-linear operation (activation function φ) [54].

It is based on free-space optics, without the use of light wave guidance and integrated circuits, encoding the input signals with alterations in the illuminating power. During its linear operation, the light impacts on different areas of the surface of a spatial light modulator (SLM) representing the knots v_i of the input layer of a NN. With a special grid coating, the impacting light beam can be split in different j directions with weight W_{ij}. The SLM is placed in the rear focal layer of the lenses, which apply Fourier transformation and sum up all the diffracted beams on the focal point as follows [49,51,54]:

$$z_i = \sum_j W_{ij} v_j \tag{6}$$

This is as it happens with every knot of a conventional NN.

The non-linear operation is accomplished through electromagnetically induced transparency (EIT), which is based on quantum phenomena and is produced by laser-cooled atoms ^{85}Rb, in a magneto-optical trap (MOT) [55,56]. The materialization of this particular architecture is shown in Figure 3 [40,42,54,57,58]:
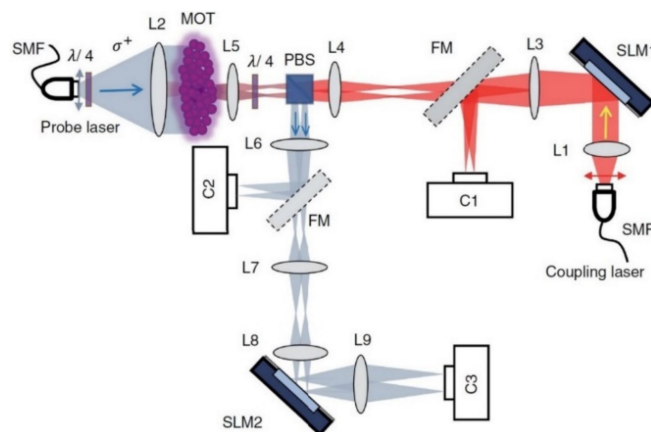

Figure 3. Implementation of the all-optical neural network (AONN) based on free optics [54].

The light beam from the laser single-mode fiber (SMF), which constitutes the encoded input layer, is aligned (L1) and impacts on the first modulator (SLM1), which, in turn, emits four different beams. These are directed towards the L3 lens as v_j inputs, while at the same time the C1 camera, through a special flip mirror (FM), records and calculates their values. Through the L4 and L5 lens system, the non-linearity is introduced by the MOT and afterwards the beams are directed towards the SLM2, after being recorded by the C2 camera first. Finally, the next layer (output layer), which consists of the SLM2 and the L7, L8 and L9 lenses, transforms the four beams into two, which are recorded by the C3 camera [27,29,54]. It is important to mention that the single-layer perceptron's optically implement matrix multiplications. Implementation of matrix multiplication in the optical domain has been a topic of research for decades, and has been shown in free space through the use of beam splitters or Mach–Zehnder interferometers as well as in integrated photonic circuits through the same mechanisms, for application in optical signal processing and reconfigurable optical neural networks. Recently, diffractive neural network architectures have been proposed, in which these matrix multiplications are performed by diffractive elements. This marked the beginning of optical data processing through diffractive neural network inference, although the fabrication methods applied are only suitable for devices operating with a low neuron density.

For the evaluation of this architecture, a classification of the different stages of an Ising model [59] has been carried out, giving similar results compared to a NN created by a computer as these are represented in Figure 4 [37,54].
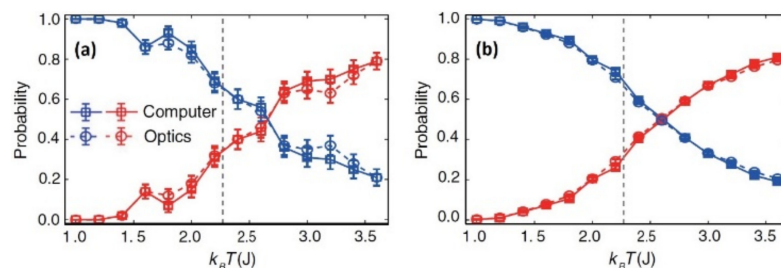

Figure 4. Average possibility of right (blue) and wrong (red) classification of this stage subject to temperature T (K) for 100 (**a**) and 4000 (**b**) settings [54].

It is obvious that this particular materialization can fully substitute a NN created by computers, with the only exception being the fact that commercial SLMs are not fast enough compared to a modern computer [39,54,60].

4.2. Multilayer Perceptrons

A modified model of multiple sensors is called multilayer perceptron, in which between the input and output layers intervene in one or more hidden layers. The data flow in such a network is always from its inputs to its outputs and there is no feedback loop. We also assume that the neurons in every layer interact only with those neurons that belong to their directly adjacent layers. In other words, the first hidden layer accepts the values of the input layer, and the results of the first hidden layer go through the second hidden layer, whose results then go through to the third layer until they finally reach the final output layer. The materialization of the nanophotonic [46,48] multilayer perceptron of Figure 5 is based on the use of nanophotonic circuits that process coherent light [49,61–64].
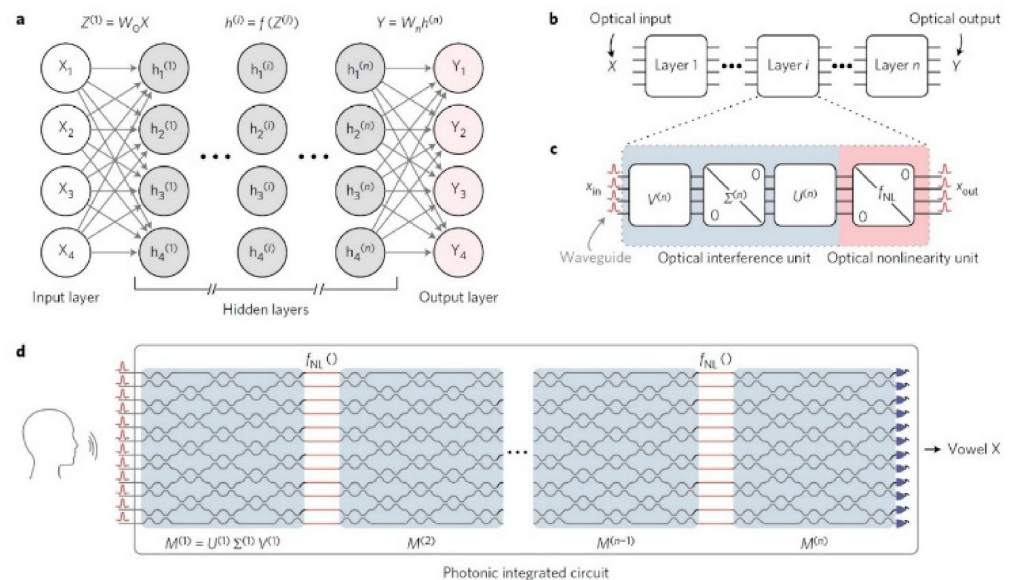

Figure 5. Nanophotonic multilayer perceptron architecture: (**a**) A typical NN with its input–output layers and *n* hidden layers. (**b**) Hidden layers in optical implementation. (**c**) The optical units in each hidden layer. (**d**) The final arrangement in an integrated circuit [64].

As shown in Figure 5, the basic theoretic block of NN with the hidden layers (grey) (a), is transferred to optical level operation (b), using two basic optical parts (c), and in particular, an optical interference unit (OIU), which performs multiplication of matrixes, and an optical non-linear unit (ONU), which materializes the activation function. All the above are given briefly and in an integrated circuit form (d). The OIU consists of ranks of special programmed Mach–Zehnder interferometer (MZI). The MZIs convert the phase differences of light into amplitude differences (modulation). The structure of a MZI is shown in Figure 6 [18,51,52,61,64].

Finally, the modification of input M matrix of every i^{th} rank into a matrix product is accomplished as shown in Equation (7) according to the singular-value decomposition (SVD) [18,49,64,65]:

$$M^{(i)} = U^{(i)} \Sigma^{(i)} V^{*(i)} \tag{7}$$

where U is a $m \times m$ real or complex unitary matrix, Σ is a $m \times n$ orthogonal diagonal matrix with no negative values in the diagonal and V^* is the conjugate transpose of V, which is a $n \times n$ real or complex unitary matrix. It must be highlighted that through the M matrix, the weight matrix W_i of the NN is transferred to the optical circuit [22,66,67].
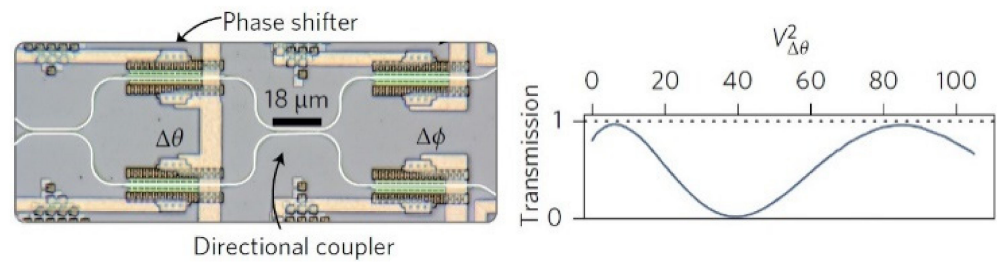
Figure 6. The programmable phase shifter creates modifications in the phase, which, in turn, are converted to amplitude modifications in the directional coupler [64].

The unit of the transfer function in this particular research paper has not yet been experimentally materialized and can only be simulated in a computer, with transformation of the signals from the optical to the electrical layer of operation and vice versa.

The experimental study and the training of ONN was put into practice on a computer first, initially through the application of voice recognition with 76.7% accuracy. Then, the already familiar diagnostic tool of digital identification, the Modified National Institute of Standards and Technology database (MNIST), was used, in which accuracy of results reach 95%, with the highest known value being 99%. This last conclusion shows the potentials and dynamics of mechanical learning in this particular field [18,48,52,68].

4.3. Deep Photonic Neural Networks

Deep neural networks (DNNs) are used in solving complex problems of high complexity like medical image analysis, speech recognition, language translation, image classification and many more [52]. However, as the number of layers increases, its structure becomes more complex and this result in the input of a great computational load on the processor. Consequently, the training time increases and so does the energy consumption. These restrictions created the need for the materialization of PNNs of many layers (deep photonic neural networks—DPNNs), since the advantages of photonic transmission speed and the minimum energy consumption are indisputable [47,49,52].

The construction of standardized, fully optical circuits, with many layers is a true challenge nowadays. An arrangement for the materialization of a DPNN is shown in Figure 7 [69].
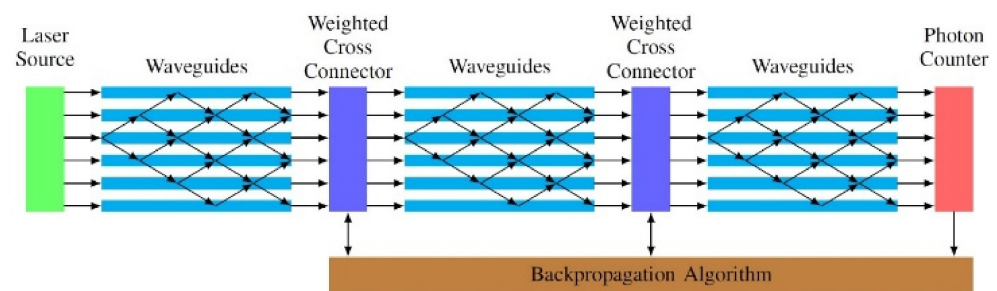

Figure 7. The architecture of a deep photonic neural network (DPNN) [69].

In this particular architecture, the layers of NN are substituted with photonic grids, in which instead of nodes with neurons we have waveguides. The interconnection between layers is accomplished through coupling devices with weighted cross connectors so that the desired output from the network can be achieved. The coupling devices, which are responsible for the control of photons, consist of optical splitters and optical combiners for which the following relation is in effect:

$$c_i = \sum_j w_{ij} \cdot s_j \tag{8}$$

The weights w_{ij} are controlled by external parameters until the network reaches the ideal output and be led to a condition of stable weights (training), where we obtain the following [18,19,22,40,42,46,49,52,69]:

$$\alpha_i^{(l)} = \sum_{j=1}^{N_{(l-1)}} w_{ij} f_j^{(l-1)} \left(a^{(l-1)} \right) \tag{9}$$

4.4. Convolutional Neural Networks

The convolutional neural networks (CNNs) [70] adopt a different approach in their organization as they take advantage of the hierarchical standard of the input data, creating more complex, but fewer and simpler patterns, in their architecture. The nanoscale neuron size not only provides the advantage of a high neuron density, but also results in a short distance (the operative distance, i.e., the distance between the input and output planes, is one to three orders of magnitude smaller than that in other implementations) and more connections between the neurons due to the increased diffraction angles. These features lead to three orders of magnitude increase in the operational frequency, and thus in the operations per second (FLOPS) compared with the devices in the THz region. In this regard, smaller feature sizes can be achieved (<10 nm), potentially creating a completely new platform for smart systems based on CNN.

The architecture of a CNN is analogous to the one of the convertibility patterns of neurons of the human brain and was inspired by the organization of the optical cortex. More analytically, a CNN is a deep learning algorithm, which can take an image at the input, assign the appropriate weights to some of its various characteristics and, consequently, be able to differentiate one from the other. In other words, it has the ability to successfully record the spatial and temporary dependencies in an image through the application of relevant filters. Thus, a better adjustment to the total data is accomplished due to the decrease in the number of parameters that are involved and the reuse of weights [71]. In other words, the network can be trained to better comprehend the structure of an image for example, while the preprocessing that is needed in a CNN is smaller when compared to other classification algorithms. The outcome is that CNNs have an advantage over the NN with perceptron's because the latter are prone to data overload due to the full connection of their knots.

There are several suggestions with CNNs that have been published such as [10,53,72]. A hybrid multilayer optical-electrical NN based on an optical matrix multiplier is presented in Figure 8 [73].

In every one of the network's layers, the inputs $x^{(k)}$ are multiplied with the corresponding weights $W_i^{(k)}$, which are encoded as optical signals with homodyning between each pair of signal weight. The electronic signals that derive are then subject to a non-linear transition function f and are converted to serial signals. Then, they are converted once again to optical signals and are sent to the input of the next layer. This optical system can be used for fully connected as well as CNNs and allows for the inference of conclusions as well as the training in the same optical device.

Another suggestion of a CNN with full use of optical convolutional neural networks (OCNN) is presented in Figure 9 [18,49,70,74,75].

The architecture consists of layers separated in an OIU based on MZI, which performs linear operations on the center panel (convolutional and pooling), one part for the input of the non-linearity unit and a splitters network of 3 dB, for the reorganization of data that the CNN is processing (re-shuffling).

The separators are programmed to introduce the appropriate time lag so that, at the output of the network layers, the signals could synchronize in time and form a new data entry for the input into the matrix nucleus of the next layer. It can be calculated that with this particular architecture the processing would be 30 times faster than that of an especially purpose-built electronic processor for CNNs with the same power consumption. As a result,

such a system could play a significant role in the processing of thousands of terabytes of image and video data that are produced every day on the internet [41,52].
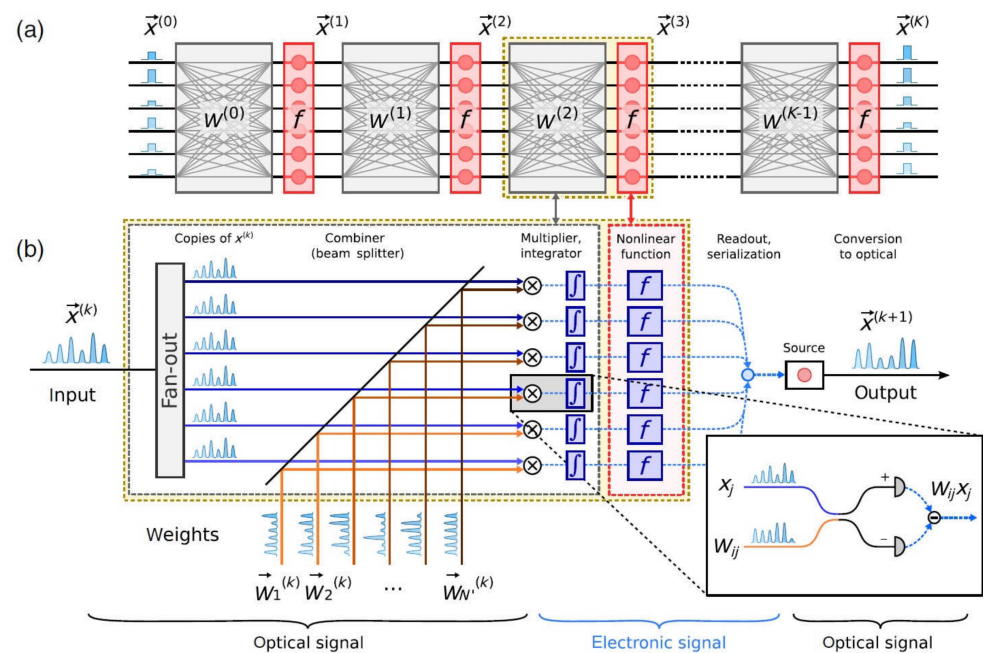

Figure 8. (**a**) Schematic diagram NN of K-layers consisting of a multiplier (grey) and an element for the activation function (red). (**b**) The multiplication performs a combination of inputs with the weight signals using homodyning [73].
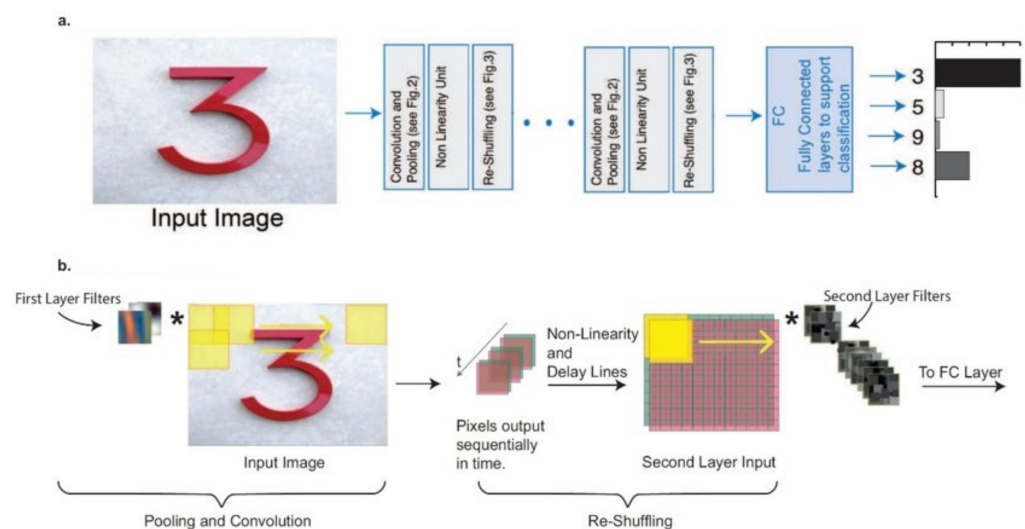

Figure 9. The suggested architecture for a fully optical CNN. (**a**) Logic Block Diagram and (**b**) Schematic Illustration [75].

4.5. Spiking Neural Networks

The *spiking neural networks* (SNNs) [76–78] are networks that imitate more than any other the biological NNs. Apart from the neural and synaptic condition, the SNNs incorporate the concept of time in their operating model. The idea behind this is that the neurons in a SNN should not trigger and be triggered in every propagation circle, as in standard networks of multiple layers with perceptron's. As it happens with the biological neurons, when the dynamics of their cell membrane reaches a particular value, which is called action potential, then the neuron triggers and produces a signal that travels to other neurons,

which, in turn, increase or decrease the dynamics of their cell membrane according to this particular signal. The SNNs use peak sequences as mechanisms of internal information presentation, in contrast to the usual continuous variables, while at the same time having equal, if not better, performance in computational cost to the traditional NNs [79–81].

In the field of optical SNNs, many studies have been conducted in the past years [82,83], initially taking advantage of the fast optical elements used in the construction of big systems with optical fibers. Despite the significant advances to build active optical artificial neurons using for example phase-change materials, lasers, photodetectors and modulators, miniaturized integrated sources and detectors suited for few-photon spike-based operations and of interest for neuromorphic optical computing are still lacking. The successful applications finally led to the completion of arrangements, aiming for greater scalability, increase of energy efficiency, reduction of cost and flexibility in the environmental fluctuations.

In a survey, the use of a graphene laser is recommended as an artificial neuron, which is the fundamental element for the processing of information in the form of spikes. Moreover, the integrated layer of graphene is used as an optical absorber for the materialization of the non-linear activation function. The following Figure 10 presents the application with the use of circuits of free optics for the creation of a series of current peaks with adjustable characteristics of width and breadth [49,82,84,85].
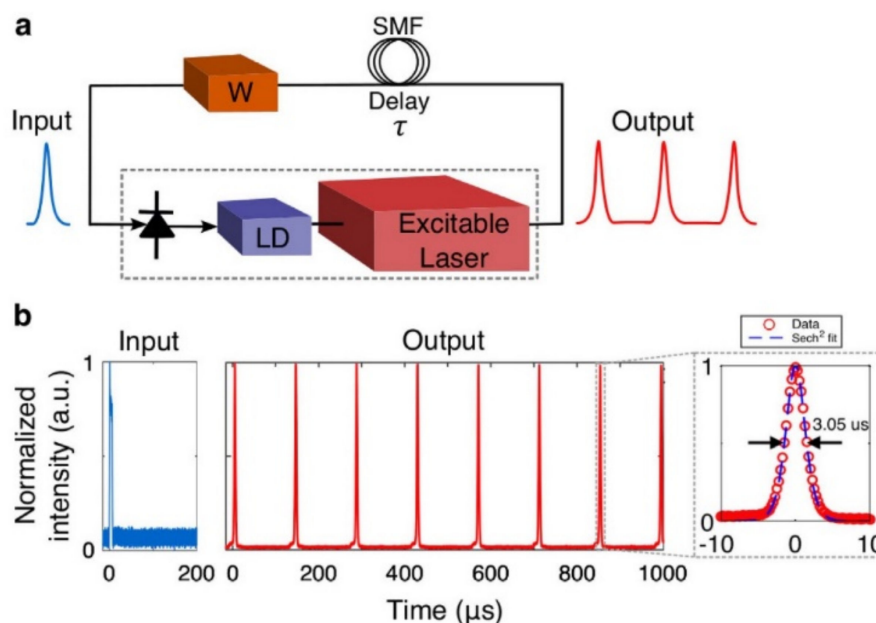

Figure 10. (**a**) The circuit for the creation of repeated current peak. (**b**) The waveforms of the implementation. One pulse of the output is led to the input via single-mode fiber (SMF), which acts as a delay element [82].

In another survey, the fundamental neuron is based on distributed feedback (DFB) laser of semi-conductors of indium phosphide [86]. The use of this type of laser devices is very common in the construction of SNNs. The laser possesses two photodetectors (PD), which allow for inhibitory as well as excitatory stimuli. The recommended device is very fast, reaching 1012 MACs/sec (MAC—Multiply Accumulate Operations) [87,88].

4.6. Reservoir Computing

The use of recurrent neural networks (RNNs) [89,90] has attracted researchers' interest because of their dynamics. The traditional RNNs, however, present some problems in training and designing, so an evolution has been suggested, namely reservoir computing (RC). It is virtually a neural network of feedback, where the input signals are dependent on time and present maximum efficiency compared to any other architecture in applications of sequence signals such as voice recognition, time series prediction, etc. An RC system

consists of a reservoir through which a recording of inputs is conducted in a *n*-dimensional space and a readout layer, where the analysis of standards introduced in the reservoir is performed.

Optical applications with photonic reservoir computing (PRC) architecture are presented in several research projects. In one of them, as shown in Figure 11, passive optical elements are used for the materialization of the reservoir, which consists of a $4 \times 4 = 16$ node system with splitters, couplers and waveguides, creating in this way a complex interferometer that operates in a random way. The fact that it comprises only passive elements renders it perfect from an energy efficiency point of view, but it displays solely a linear behavior. This can be offset in the readout layer with the introduction of a photodiode as a non-linear element [50,51,91,92].
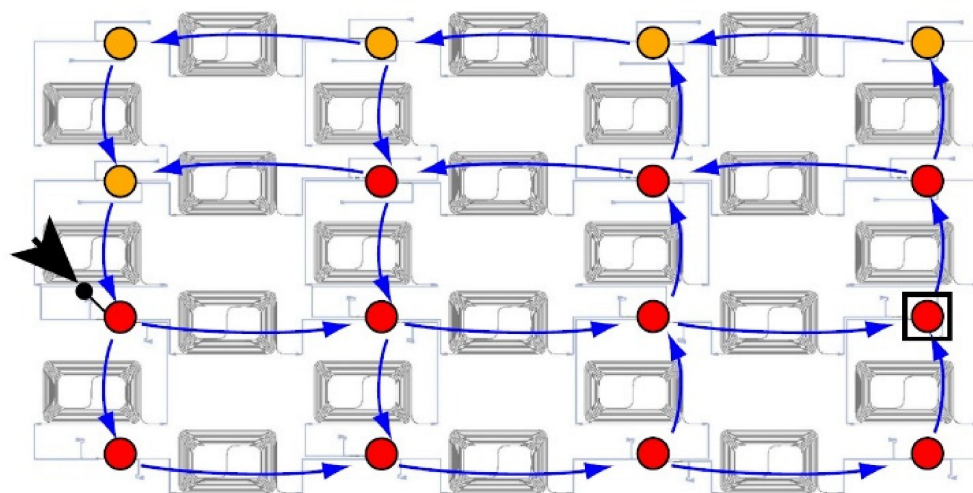

Figure 11. The reservoir structure in optical materialization (chip). It is consisted of interferometers for coupling and splitting between the nodes. Blue arrows represent the specific light flow, if for input is used the node indicated with black arrow. Nodes with yellow dots have output powers below the noise floor. Red ones have an amplitude above noise floor and were measured and used for offline training. For testing the device, an example waveform with sequences of bits with "1" and "0" were collected in the black square with a rounded red dot [50].

In Figure 12 is presented a new topology for the reservoir, based on micro-ring resonators (MR), which are non-linear elements and can cover the need for a non-linear transition function, simplifying in this way the readout layer to the fullest extent [51].
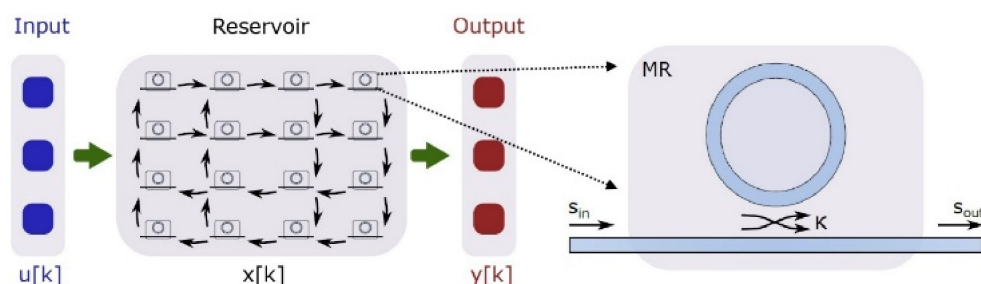
Figure 12. The reservoir with the 16 nodes made from silicon on insulator (SOI) MR [51].

The topology in Figure 12, displayed a better error rate compared to others where the reservoir consists of passive linear elements. The reservoir model of our proposed photonic neuron, on the other hand, can change due to collective and synchronous dynamics of the network for spontaneous information processing because the reservoir dynamics can be controlled by tuning optical-pump amplitudes. Network experiments with reservoir

neurons revealed that input signals from the correlating neurons can induce an effective change in the pump amplitude. The effective change depends on the increase in the order parameter of synchronization, and it causes spontaneous changes in reservoir modes and firing rates of the networked neurons.

An alternative suggestion for PRC is based on the use of photonic crystal cavity (PCC) in the shape of an ellipse quadrant. This particular architecture proves quite useful for projects of processing optical signals dependent on memory such as the header recognition of digital signals [40,74,91,93].

5. Training Methodologies

Training is an important aspect of the neural networks, since it does not only influence the behavior of the network, but also its overall efficiency. In supervised training, the training procedure uses an objectively calculated operation, where the distance (or error) between the desired and the real value is calculated. This operation is used to regulate the internal parameters of the NN, which are the synapses/connections weights of neurons. In order to minimize the deflection between desired and real value, a gradient vector is calculated so that the way of how the error is influenced by any weight shift can be assessed. [49,52,83].

Every time there is a change in the nature of input data in the network, the network needs to be retrained. This retraining can be done gradually as the network performs inference (online learning) [83] or it can be done independently, so that the network can adapt to a new input of training data (offline learning).

Given that the training includes gradient calculation, or even more complex calculations, it is a stage of resources and time consumption. In contrast, the inference (the classification stage by the NN) is a much simpler procedure since the weights are already known in this stage. For this reason, many materializations of PNNs support only the inference stage and the weights are taken with the use of software applications on the level of electronic operation. Moreover, some applications cannot be trained at all, as in [40,42,52,94,95]. These architectures are very fast and efficient as far as energy consumption is concerned, but they are not flexible, as they are especially designed for specific applications as their weights consolidate in the material during their construction.

When in a NN the training is electronic, two main disadvantages appear and, in particular, the physical system dependence on the accuracy of the model is added and the improvement of speed and the efficiency already accomplished with the optical part is eliminated. In order for the training, though complicated, to take full advantage of the photonic technology, it must be specifically adapted to optical architectures.

5.1. Propagation

The ONNs offer many advantages as far as the training of NNs is concerned. In a conventional computer, the training is done using the backpropagation error method and the gradient descent application [96]. Nevertheless, in some NNs where the active number of parameters (which are being calculated in every circle) far surpass the number of distinct parameters (as in RNNs and CNNs), the training with backpropagation is definitely ineffective. In particular, the repeated nature of RNNs virtually makes them an extremely deep NN, whereas in CNNs, something relevant happens, since the same weight parameters are used repeatedly in different parts of an image for the output of its characteristics [18,41,97].

For the training of the network with forward propagation, and also for the calculation of gradient in a particular modification step Δ_{wij} of the weight w_{ij} of a NN, calculations of quantities are needed using the finite difference method (FDM) [98–100]:

$$f\left(w_{ij} + \delta_{ij}\right) \text{ and } f\left(w_{ij}\right) \tag{10}$$

After these two arithmetical operations, we calculate the weight change as follows:

$$\Delta w_{ij} = \frac{f\left(w_{ij} + \delta_{ij}\right) - f\left(w_{ij}\right)}{\delta_{ij}} \tag{11}$$

In a conventional computer, the above procedure is computationally costly. On the other hand, in the field of photonic applications, there are suggestions in ONNs that are better at the immediate calculation of the gradient, as every one of the aforementioned steps of propagation is calculated in stable time, which is restricted only by the rate of photo detection, which reaches 100 GHz, and the energy consumption is analogous only to the number of neurons [64].

This particular architecture is capable of reaching performance rates similar or even faster than backpropagation with conventional computers (e.g., in very deep RNNs). Moreover, with the training procedure in the material (on chip), one can easily parameterize and train unitary matrixes, an approach that is particularly useful in deep NNs [72].

Furthermore, in ONNs there is a possibility of training with the backpropagation method, based on the architecture where OIU with MZI are used for the linear operations of multiplications of matrixes [64]. The algorithm of backpropagation training generally operates in a circular mode between two stages, where in the first stage the error propagation is from the end of the network to its beginning, and in the second stage, there is a recalculation of the weights to check the contribution of each one to the output of the network.

In optical materializations, some basic restrictions to the control of weights are present, which musttake into account that $w_{ij} \geq 0$. There cannot be a negative weight value since there is no negative light intensity value [42,49,72]:

$$\sum_i w_{ij}^{(l)} = 1 \tag{12}$$

The initial light beam is split into waveguides so that the total of their intensities is stable. These particular restrictions are incorporated with the use of functions, which transform the weights w to the desired breadth of activation function values, such as softmax [52,101,102]:

$$w_{ij}^{(l)} = \frac{e^{w_{jj}^{(l)}}}{\sum_i e^{w_{ij}^{(l)}}} \tag{13}$$

In order for backpropagation to be applied, the physical materialization of adjoint variable method (AVM) [103,104] is needed, which allows for the reverse designing of photonic structures. According to this, at first, the adjoint of the initial field is created, the complementary one is propagated in the network reversely to the initial one and the initial field contributes with a replica of the reverse time of the complementary field. After all these, the conditions that yield gradient in every spot are expressed as the solutions of a classical conjugate electromagnetic problem and can be retrieved with an on-the-spot calculation of the field's intensity. A visualization of the operation of this particular method is presented in Figure 13 [18,41,49,96,97].

This method allows for the effective materialization of backpropagation in a hybrid optical-electronic network, with its main restriction being that a forward feed system, which is mutual and with no losses, is necessary. Moreover, the fact that this method is based on classical Maxwell electromagnetic equations and not on a particular network form renders it extremely flexible for its application on any photonic platform [49,51,70].
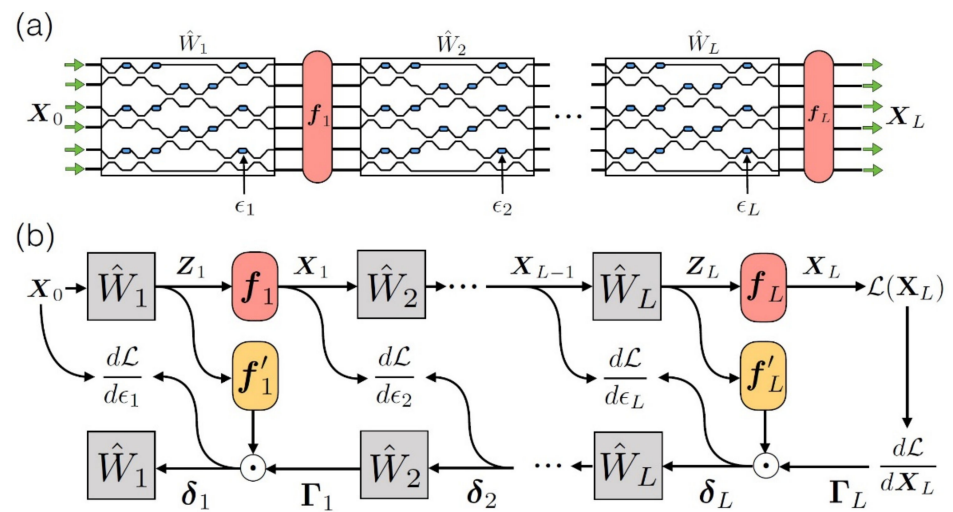
Figure 13. Backpropagation PNN. In stage (**a**), the squares correspond to the OIUs, which materialize the linear operation (matrixes W_L). In blue color, we see the integrated phase shifters for the control of OIU and the training of the network. The red areas correspond to the non-linear activation functions f_L, which are performed through a computer. Respectively, in stage (**b**), the presentation of the operation for the calculation of NN ranks. The route on top corresponds to the anterior propagation and the bottom to the backpropagation [96].

5.2. Non-Linearity Inversion

In RC photonic applications, the training concerns the readout layer [105,106]. Recently, various researchers on RC have focused their attention on the development of the reservoir with several recommended solutions [107]. Nevertheless, the reading level is of fundamental importance because it ultimately determines the behavior of the network and, unlike the reservoir, must be appropriately trained [51,91]. Hitherto, the training and the conjugation of signals on the reading level has taken place in the conventional electric space, and this resulted in the loss of any gain in speed and energy consumption that the optical part of arrangements introduced [50].

For a fully optical solution in the RC networks, only a simple photo detector is required, which will receive the weighted total of all the optical signals. This approach, however, displays a drawback: we lose the ability for direct observation of the conditions of the photonic reservoir, which is necessary in many linear training algorithms. In order to solve this problem, there is a training procedure presented in Figure 14, in which the reservoir's states are estimated through a single photodetector at its output, which includes an approximate inversion of the non-linearity of the photodetector, so it was named non-linearity inversion [42,107–109].
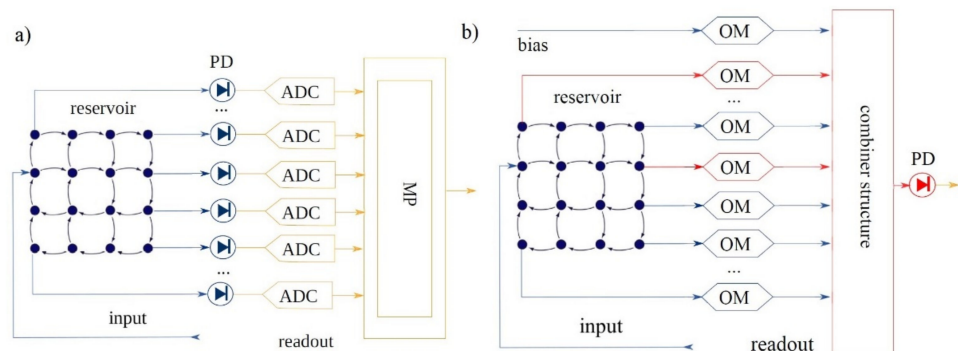

Figure 14. (**a**) The mixed way for training: the optical signal from every node of the reservoir (blue) is transferred through a photodetector (PD) to the electric space (yellow) and through an A/D converter

(ADC) to the microprocessor (MP). (**b**) Non-linearity inversion method: the optical signals are modulated (OM) implementing the weights and summed (combiner structure), before converting to electric signal via PD. The states of the reservoir are estimated by setting the weights (red) according to a certain pattern [107].

This method solves the aforementioned issue of direct observation of the reservoir from the reading layer through a PD, with which a calculation of amplitude and its conditions' phase is materialized. The more complex conditions of the reservoir are observed with the appropriate adaptation of the reading weights, whereas the feedback is achieved through a predetermined input sequence [47,60,89,107].

6. Activation Functions

Neurons are the structural element of the network. Each one of these knots receives a total of arithmetic inputs from different sources (either from other neurons or from the environment), performs some computation based on these inputs and produces an output. This output is either directed to the environment or constitutes input to the other neurons of the network. The computational neurons multiply each one of their inputs by the corresponding synaptic weight and calculate the total sum of the products. This total constitutes the activation function definition, which every knot materializes internally. The value that the function takes for this particular definition is also the output of the neuron for the current inputs and weights.

As a result, an important decision that has to be taken into account for the smooth operation of NNs is the selection of the activation function. In bibliographic references, the use of a powerfully non-linear function based on the electro-optical phenomenon is recommended for better results [110]. Respectively, a plethora of non-linear functions have been materialized, which are presented in the next sections [89,107].

6.1. z–Transform (Complex Non-Linearity)

This function represents the $Z \rightarrow |Z|$ transformation and can be used for full, condense, polar mode. The bilateral z-transform of a sequence of distinguishable time is defined in Equation (14) [111–113]:

$$X(z) = \sum_{n=-\infty}^{+\infty} x(n) \cdot z^{-n} \tag{14}$$

where the complex invariable z is called complex frequency and can be expressed with the use of polar coordinates. The z transformation of a sequence of distinguishable time is a total of infinite terms, which may converge to a real number for some values of the complex z variable and may not converge for some values of the complex z variable. The total of the variable values for which the z transformation exists, that is, for which the total of z transformation converges, constitutes the region of convergence (ROC) [49,114].

The reverse transformation is accomplished by calculating the reverse z transformations in each term of the total using z transformation pairs and, eventually, using the property of linearity of the z transformation. It is materialized with the method of analysis of the rational function in a total of simple fractions as is shown in Equation (15) [17,49]:

$$X(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^{M} b(k) \cdot z^{-k}}{\sum_{k=0}^{N} a(k) \cdot z^{-k}} \tag{15}$$

As an activation function in optical materializations, it is applied in signal analyses and, specifically, in solving linear equations of differences with fixed factors, in calculating the response and in designing linear filters or convolution layers [70].

6.2. Electro-Optical Activation (Complex Non-Linearity)

In NN applications of optical components, there is the possibility of creating non-linearity from the already existing material. The activation function is materialized, converting a small part of the power of the input of the optic signal into electrical voltage. The remaining part of the initial optic signal is developed according to phase and amplitude by this voltage as it goes through an interferometer. A typical example of an electro-optical activation function is presented in Figure 15 [60,85,91,109,110]:
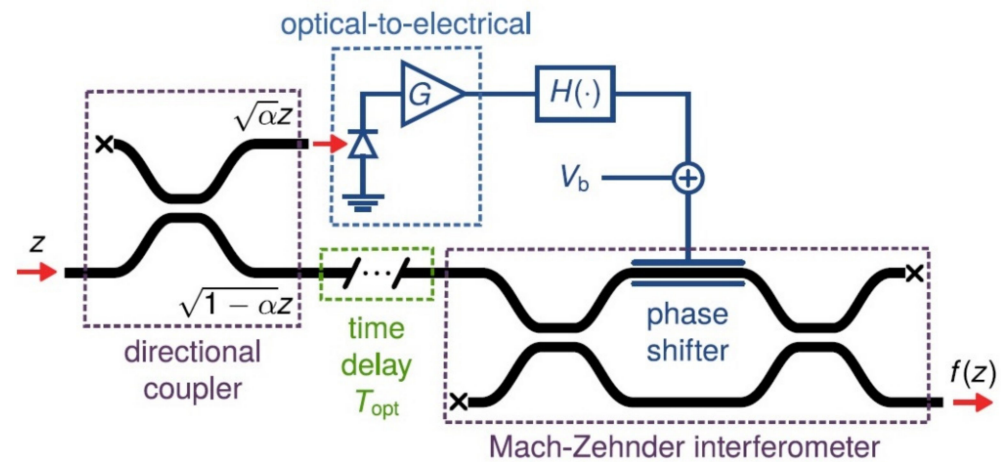

Figure 15. The arrangement for the electro-optical activation function [110].

For an input signal with a *z* value of amplitude, the non-linear activation function *f(z)* happens as the response of the interferometer and of the components throughout the route of the electric signal as is shown in Equation (16) [15,41,52]:

$$f(z) = j\sqrt{1-a} \cdot \exp\left(-j\left[\frac{g_\varphi |z|^2}{2} + \frac{\varphi_b}{2}\right]\right) \cdot \cos\left(\frac{g_\varphi |z|^2}{2} + \frac{\varphi_b}{2}\right) z \qquad (16)$$

where: $\varphi_b = \pi \frac{V_b}{V_\pi}$ και $g_\varphi = \pi \frac{aGRR}{V_\pi}$

(1) α: the factor of input power transformation into an electric signal.
(2) R: the response of the photodetector to the optical to electrical unit.
(3) G: the gain of amplification rate.
(4) V_b: the biasing voltage (bias).
(5) V_π: the required voltage for the π transformation of the phase.

6.3. Sigmoid (Complex Non-Linearity)

The sigmoid activation function is used when a classification between two classes is needed or for a regression of weighted arrangements, as it offers numbers between the space [0, 1] at the output. This can be represented by the transformation shown in Equation (17) [60,77,83]:

$$z \to \frac{1}{1 + e^{-z}} \qquad (17)$$

6.4. Softmax (Complex Non-Linearity)

The Softmax function is represented by the transformation [101,102]:

$$z \to \frac{e^z}{\sum e^z} \qquad (18)$$

It is mainly used for multiclass problems.

6.5. SPM Activation (Non-Linearity)

It represents the transformation [52,110]:

$$Z \rightarrow Z \cdot e^{(-jG|z|^2)} \tag{19}$$

where $G = rad / \left(\frac{V^2}{m^2} \right)$ is the phase transformation for every unitary change of the input voltage.

6.6. zReLU (Non-Linearity)

The *zReLU* is the rectified linear unit function, with which the positive part of its definition is received as follows [10,45,49]:

$$f(z) = \begin{cases} z \text{ if } \mathrm{Re}(z) > 0 \text{ and } \mathrm{Im}(z) > 0 \\ 0 \end{cases} \tag{20}$$

6.7. Cosine Activation Function (Non-Linearity)

Many of the recommended optical architectures for NNs use general-purpose equipment (e.g., for optical communications), whereas, ideally, they should be materialized inside a specific material (hardware). Consequently, there is no general approach to the training method of every recommended technique, as each one of them has its own characteristics that should be taken into account. A familiar problem that photonic architectures display concerns the activation function, due to the limited available choices and the difficulty of its materialization. Most of the suggestions use a combination of optical and electronic elements, such as the Mach–Zehnder optical modulators (MZM) [115,116]. The result is a non-linear activation function of cosine form, which is presented in Equation (21) [60,70,76,117]:

$$P_{out} = P_{in} \sin^2 \left(\frac{\pi}{2} \frac{V_{RF}}{V_\pi} \right) \tag{21}$$

where P_{out} is the output signal, P_{in} is the continuous wave (CW) under modulation signal, V_{RF} is the input signal of the function and V_π is the value of input voltage for a phase shift of π value.

Another important problem that must be resolved in PNNs is the initialization of their parameters, such as the choice of the initial values of their weights. In their initial definition, the restrictions that exist in every materialization should be taken into consideration, as, for example, the constant bounded response of the signals that go through all the layers of the network. The topology, with which an optical neuron of a cosine activation function is materialized, is shown in Figure 16 [18,49,51,52,117]:

In this particular materialization, two lasers of a different wavelength, $\lambda_i(+)$ and $\lambda_i(-)$, are used, which, through the MZIs functioning as switches (frame *sign of* $W^{(1)}$), are corresponded to positive and negative values of weights, respectively. Afterwards, the signals are led to the modulators (MOD, frame *Input* $X^{(1)}$) so that the input signal can be "printed" on an optical signal of power $P(X_i^{(1)})$. The next level (frame *Weight* $|W^{(1)}|$) includes a variable optical attenuator (VOA) [104,105], which is responsible for the amplification of signal-weight as is shown in Relation (22) [18,40,49,117]:

$$W_i^{(1)} \cdot P_{x_i^{(1)}} \tag{22}$$

In the next step, the signals are multiplexed (frame *MUX*) and are led in a grade of asynchronous MZI (frame *A-MZI*) for the separation of signals, in signals of positive weight ($\lambda_{1\ldots 9(+)}$) and signals of negative weight ($\lambda_{1\ldots 9(-)}$), and in the end are added up in photodiodes (blue color). In conclusion, the MZM modulator that follows (MOD) and receives the two signals operates in its non-linear area, materializing the transition function of cosine form. This particular architecture, where each neuron produces a signal that is

led to the input of the next neuron, can be completed constructively and constitutes an independent photonic processor (chip) [20,36,84].
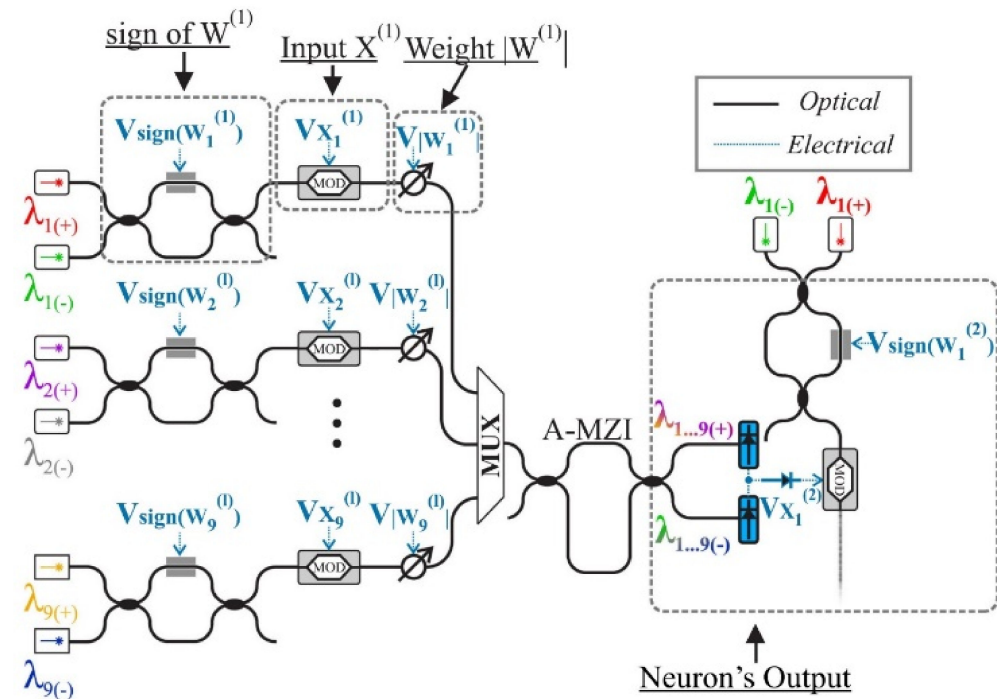

Figure 16. The operation principle of a neuron in an optical materialization [117].

7. Conclusions

In this research paper, we present an overview of the development and materialization methods of neuromorphic circuits of nanophotonic [61] arrangements for every respective contemporary architecture of conventional neural networks, and the advantages and restrictions that arise during the transition from the electronic to the optical materializations are displayed. The aforementioned networks are energy efficient, when compared to the corresponding electronic ones, and much faster due to photons. The reduction of simultaneous processing time radically increases the potentials of modern computational systems, which use optical arrangements, offering a promising alternative approach to micro-electronic and optical-electronic applications.

All these lead to the conclusion that there are potentials for a full transition to optical materializations as these display the following advantages:

(1) Most of the systems do not require energy for the processing of optical signals. As soon as the neural network is trained, the computations on the optical signals are conducted without any additional energy consumption, rendering this particular architecture completely passive.

(2) The optical systems, in contrast to the conventional electronic ones, do not produce heat during their operation and, as a result, they can be enclosed in three-dimensional constructions.

(3) The processing speed in the optical systems is restricted only by the operation frequency of the laser source of light, which reaches 1 THz.

(4) The optical grids enable the multiplication of matrixes with vectors, something which is essential to NNs. The linear transformations (and some non-linear ones) can be performed at the speed of light and detected at a rate of over 100 GHz in photonic networks and, in some cases, with a minimum power consumption.

(5) They are not particularly demanding as far as non-linearities are concerned, since many innate optical non-linearities can be used directly for the application of non-linear operations in PNNs, such as the activation functions.

In conclusion, such a system comprises the most efficient, quick and stable circuits of multiple conventional and high non-confining optical technology components for optimal processing, which mimic the key properties of a real brain.

On the other side, there are some difficulties in the transition to completely PNNs, which are the following:

(1) The dimensions of optical devices are analogous to the light wavelength that they use (400 nm–800 nm).
(2) The mass production of optical devices is limited compared to the electronic ones, since they lack at least 50 years of research and development.
(3) The training of the optical grids is quite difficult because the controlled parameters are active in matrix elements deriving from powerful non-linear functions.
(4) The application of matrix transformations with optical components of mass production (such as fibers and lenses) is a restriction to the spread of ONNs due to the need for stability in the signal phase and to the huge number of neurons, which are required in more complex applications.

To summarize, nanophotonics are more expensive and harder to fix, and waveguides and fibers are harder to use than wires and are characterized by spurious reflections that are more troublesome.

Although there are potentials concerning the materialization of PNNs, there are still some areas that require further research, such as some specific architectures of deep neural nets, specifically Long Short-Term Memory Neural Networks, Generative Adversarial Nets, Geometric Deep Neural Networks, Deep Belief Networks and Deep Boltzmann Machines. Due to the significance of DNNs and the role they play in mechanical learning techniques, the research studies should focus on the question whether every type of conventional DNN can be converted in PNN, performing better and, thus, offering more advantages when compared to electronic arrangements. The ultimate goal in this is to replace the huge energy-consuming NNs, with thousands of knots and multiple interconnections among hidden layers, with very fast optical arrangements.

There are also fields where the research on PNNs should focus on, such as the hyper dimensional learning (HL) [118,119], a modern and very promising approach to NNs, which is still in the development stage. Here, the problem of a photonic materialization lies in the very big size of the internal representation of objects that are used in HL.

A further point that needs to be studied is the application of non-linear functions, which in most of the suggestions are materialized through software outside the optical arrangement. This results in the decline of performance, sometimes of a high rate, given that in multilayer NNs it is necessary to insert non-linearity many times successively.

Many more challenges need to be overcome, such as the many different hardware platforms that have been recommended, which are still under investigation with no clear winner yet. Moreover, we have to improve the already developed hardware as, in many cases, basic elements are still simulated, or classic electronic ones are used. Furthermore, a critical element in a recommended NN architecture is its expandability in various applications, something that must be confirmed with further research studies. Finally, the field of NNs, which is still in early stage, is the massive integration of optical arrangements and, of course, their mass production, which is the last and most fundamental fortress of conventional NN arrangements against the transition to fully optical circuits.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ADC	Analog Digital Converter
AI	Artificial Intelligence
A-MZI	Asynchronous Mach–Zehnder Interferometer
AONN	All Optical Neural Network
AVM	Adjoint Variable Method
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CW	Continuous Wave
DNN	Deep Neural Network
DPNN	Deep Photonic Neural Network
EIT	Electromagnetically Induced Transparency
FDM	Finite Difference Method
FM	Flip Mirror
GPU	Graphics Processing Unit
HL	Hyper-dimensional Learning
MAC	Multiply Accumulate Operations
MNIST	Modified National Institute of Standards and Technology
MOD	Modulator
MOT	Magneto-Optical Trap
MP	Microprocessor
MR	Micro Rings Resonator
MUX	Multiplexor
MZI	Mach–Zehnder Interferometer
MZM	Mach–Zehnder Modulator
NN	Neural Network
NNN	Nanophotonic Neural Network
NSoC	Neuromorphic Systems-on-Chip
OCNN	Optical Convolutional Neural Network
OIU	Optical Interference Unit
OM	Optical Modulator
ONN	Optical Neural Network
ONU	Optical Non-Linear Unit
PCC	Photonic Crystal Cavity
PD	Photodetector
PNN	Photonic Neural Network
PRC	Photonic Reservoir Computing
RC	Reservoir Computing
RNN	Recurrent neural network
ROC	Region Of Convergence
SLM	Spatial Light Modulator
SMF	Single-Mode Fiber
SNN	Spiking Neural Networks
SVD	Singular-Value Decomposition
TPU	Tensor Processing Unit
VOA	Variable Optical Attenuator

References

1. Schrettenbrunnner, M.B. Artificial-Intelligence-Driven Management. *IEEE Eng. Manag. Rev.* **2020**, *48*, 15–19. [CrossRef]
2. Srivastava, S.; Bisht, A.; Narayan, N. Safety and security in smart cities using artificial intelligence—A review. In Proceedings of the 7th International Conference on Cloud Computing, Data Science & Engineering—Confluence, Noida, India, 12–13 January 2017; pp. 130–133. [CrossRef]

3. Tewari, I.; Pant, M. Artificial Intelligence Reshaping Human Resource Management: A Review. In Proceedings of the IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), Buldhana, India, 30 December 2020; pp. 1–4. [CrossRef]

4. Quan, X.I.; Sanderson, J. Understanding the Artificial Intelligence Business Ecosystem. *IEEE Eng. Manag. Rev.* **2018**, *46*, 22–25. [CrossRef]

5. Anezakis, V.-D.; Iliadis, L.; Demertzis, K.; Mallinis, G. Hybrid Soft Computing Analytics of Cardiorespiratory Morbidity and Mortality Risk Due to Air Pollution. In *Information Systems for Crisis Response and Management in Mediterranean Countries*; Dokas, I.M., Saoud, N.B., Dugdale, J., Díaz, P., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 301, pp. 87–105. [CrossRef]

6. Werbos, P. An overview of neural networks for control. *IEEE Control Syst.* **1991**, *11*, 40–41. [CrossRef]

7. Huang, Y.; Gao, P.; Zhang, Y.; Zhang, J. A Cloud Computing Solution for Big Imagery Data Analytics. In Proceedings of the 2018 International Workshop on Big Geospatial Data and Data Science (BGDDS), Wuhan, China, 22–23 September 2018; pp. 1–4. [CrossRef]

8. Mahmud, M.S.; Huang, J.Z.; Salloum, S.; Emara, T.Z.; Sadatdiynov, K. A survey of data partitioning and sampling methods to support big data analysis. *Big Data Min. Anal.* **2020**, *3*, 85–101. [CrossRef]

9. Sperduti, A. An overview on supervised neural networks for structures. In Proceedings of the International Conference on Neural Networks (ICNN'97), Houston, TX, USA, 8–10 June 1997; Volume 4, pp. 2550–2554. [CrossRef]

10. Zhao, C.; Shen, Z.; Zhou, G.Y.; Zhao, C.Z.; Yang, L.; Man, K.L.; Lim, E. Neuromorphic Properties of Memristor towards Artificial Intelligence. In Proceedings of the 2018 International SoC Design Conference (ISOCC), Daegu, Korea, 12–15 November 2018; pp. 172–173. [CrossRef]

11. Kang, Y. AI Drives Domain Specific Processors. In Proceedings of the 2018 IEEE Asian Solid-State Circuits Conference (A-SSCC), Tainan, Taiwan, 5–7 November 2018; pp. 13–16. [CrossRef]

12. Chitty-Venkata, K.T.; Somani, A. Impact of Structural Faults on Neural Network Performance. In Proceedings of the 2019 IEEE 30th International Conference on Application-specific Systems, Architectures and Processors (ASAP), New York, NY, USA, 15–17 July 2019; p. 35. [CrossRef]

13. Li, Z.; Liu, C.; Wang, Y.; Yan, B.; Yang, C.; Yang, J.; Li, H. An overview on memristor crossabr based neuromorphic circuit and architecture. In Proceedings of the 2015 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC), Daejeon, Korea, 5–7 October 2015; pp. 52–56. [CrossRef]

14. Blouw, P.; Eliasmith, C. Event-Driven Signal Processing with Neuromorphic Computing Systems. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 8534–8538. [CrossRef]

15. De Lima, T.F.; Peng, H.-T.; Tait, A.N.; Nahmias, M.A.; Miller, H.B.; Shastri, B.J.; Prucnal, P.R. Machine Learning with Neuromorphic Photonics. *J. Light. Technol.* **2019**, *37*, 1515–1534. [CrossRef]

16. Mead, C.; Moore, G.; Moore, B. Neuromorphic engineering: Overview and potential. In Proceedings of the 2005 IEEE International Joint Conference on Neural Networks, Montreal, QC, Canada, 31 July–4 August 2005; Volume 5, p. 3334. [CrossRef]

17. Zheng, N.; Mazumder, P. *Learning in Energy-Efficient Neuromorphic Computing: Algorithm and Architecture Co-Design*, 1st ed.; Wiley: Hoboken, NJ, USA, 2019. [CrossRef]

18. Sui, X.; Wu, Q.; Liu, J.; Chen, Q.; Gu, G. A Review of Optical Neural Networks. *IEEE Access* **2020**, *8*, 70773–70783. [CrossRef]

19. Tsumura, N.; Fujii, Y.; Itoh, K.; Ichioka, Y. Optical method for generalized Hebbian-rule in optical neural network. In Proceedings of the 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), Nagoya, Japan, 25–29 October 1993; Volume 1, pp. 833–836. [CrossRef]

20. Abel, S.; Horst, F.; Stark, P.; Dangel, R.; Eltes, F.; Baumgartner, Y.; Fompeyrine, J.; Offrein, B.J. Silicon photonics integration technologies for future computing systems. In Proceedings of the 2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC), Fukuoka, Japan, 24 July 2019; pp. 1–3. [CrossRef]

21. Shastri, B.J.; Tait, A.N.; Nahmias, M.A.; de Lima, T.F.; Peng, H.-T.; Prucnal, P.R. Neuromorphic Photonic Processor Applications. In Proceedings of the 2019 IEEE Photonics Society Summer Topical Meeting Series (SUM), Ft. Lauderdale, FL, USA, 8–10 July 2019; pp. 1–2. [CrossRef]

22. Yao, J. Photonic integrated circuits for microwave photonics. In Proceedings of the 2017 IEEE Photonics Conference (IPC) Part II, Orlando, FL, USA, 1–5 October 2017; pp. 1–2. [CrossRef]

23. Zhuang, L.; Xie, Y.; Lowery, A.J. Photonics-enabled innovations in RF engineering. In Proceedings of the 2018 Australian Microwave Symposium (AMS), Brsibane, QLD, Australia, 6–7 February 2018; pp. 7–8. [CrossRef]

24. Clark, A.S.; Collins, M.J.; Husko, C.; Vo, T.; He, J.; Shahnia, S.; De Rossi, A.; Combrié, S.; Rey, I.H.; Li, J.; et al. Nonlinear Photonics: Quantum State Generation and Manipulation. In Proceedings of the 2014 IEEE Photonics Society Summer Topical Meeting Series, Montreal, QC, Canada, 14–16 July 2014; pp. 140–141. [CrossRef]

25. Abdelgaber, N.; Nikolopoulos, C. Overview on Quantum Computing and its Applications in Artificial Intelligence. In Proceedings of the 2020 IEEE Third International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, USA, 9–13 December 2020; pp. 198–199. [CrossRef]

26. Prucnal, P.R.; Tait, A.N.; Nahmias, M.A.; de Lima, T.F.; Peng, H.-T.; Shastri, B.J. Multiwavelength Neuromorphic Photonics. In Proceedings of the Conference on Lasers and Electro-Optics, San Jose, CA, USA, 9–14 May 2019; p. JM3M.3. [CrossRef]

27. Belforte, D. Overview of the laser machining industry. In Proceedings of the CLEO '99, Conference on Lasers and Electro-Optics (IEEE Cat. No.99CH37013), Baltimore, MD, USA, 28 May 1999; p. 82, In Technical Digest; Summaries of papers presented at the Conference on Lasers and Electro-Optics; Postconference Edition. [CrossRef]

28. Lancaster, D.G.; Hebert, N.B.; Zhang, W.; Piantedosi, F.; Monro, T.M.; Genest, J. A Multiple-Waveguide Mode-Locked Chip-Laser Architecture. In Proceedings of the 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, 23–27 June 2019; p. 1. [CrossRef]

29. Kleine, K.; Balu, P. High-power diode laser sources for materials processing. In Proceedings of the 2017 IEEE High Power Diode Lasers and Systems Conference (HPD), Coventry, UK, 11–12 October 2017; pp. 3–4. [CrossRef]

30. Reithmaier, J.P.; Klopf, F.; Krebs, R. Quantum dot lasers for high power and telecommunication applications. In Proceedings of the LEOS 2001 14th Annual Meeting of the IEEE Lasers and Electro-Optics Society (Cat. No.01CH37242), San Diego, CA, USA, 12–13 November 2001; Volume 1, pp. 269–270. [CrossRef]

31. Renner, D.S.; Jewell, J.; Carlson, N.; Lau, K.; Zory, P. Semiconductor laser workshop—an overview. In Proceedings of the LEOS 93 LEOS-93, San Jose, CA, USA, 15–18 November 1993; p. 724. [CrossRef]

32. Washio, K. Overview and Recent Topics in Industrial Laser Applications in Japan. In Proceedings of the 2007 Conference on Lasers and Electro-Optics (CLEO), Baltimore, MD, USA, 6–11 May 2007; p. 1. [CrossRef]

33. Barrera-Singana, C.; Valenzuela, A.; Comech, M.P. Dynamic Control Modelling of a Bipole Converter Station in a Multi-terminal HVDC Grid. In Proceedings of the 2017 International Conference on Information Systems and Computer Science (INCISCOS), Quito, Ecuador, 23–25 November 2017; pp. 146–151. [CrossRef]

34. Ghannoum, E.; Kieloch, Z. Use of modern technologies and software to deliver efficient design and optimization of 1380 km long bipole III 00 kV HVDC transmission line, Manitoba, Canada. In Proceedings of the PES T&D 2012, Orlando, FL, USA, 7–10 May 2012; pp. 1–6. [CrossRef]

35. Suriyaarachchi, D.H.R.; Wang, P.; Mohaddes, M.; Zoroofi, S.; Jacobson, D.; Kell, D. Investigation of paralleling Bipole II and the future Bipole III in Nelson River HVDC system. In Proceedings of the 10th IET International Conference on AC and DC Power Transmission (ACDC 2012), Birmingham, UK, 4–6 December 2012; p. 13. [CrossRef]

36. Bovino, F.A. On chip intrasystem quantum entangled states generator. In Proceedings of the 2017 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, 20–24 June 2017; p. 1. [CrossRef]

37. Lund, A.P.; Ralph, T.C. Efficient coherent state quantum computing by adaptive measurements. In Proceedings of the 2006 Conference on Lasers and Electro-Optics and 2006 Quantum Electronics and Laser Science Conference, Long Beach, CA, USA, 21–26 May 2006; pp. 1–2. [CrossRef]

38. Knight, P.L. Quantum communication and quantum computing. In Proceedings of the Quantum Electronics and Laser Science Conference, Baltimore, MD, USA, 23–26 May 1999; p. 32, Technical Digest; Summaries of Papers Presented at the Quantum Electronics and Laser Science Conference. [CrossRef]

39. Ferrari, D.; Cacciapuoti, A.S.; Amoretti, M.; Caleffi, M. Compiler Design for Distributed Quantum Computing. *IEEE Trans. Quantum Eng.* **2021**, *2*, 1–20. [CrossRef]

40. Silverstone, J.W.; Thompson, M.; Rarity, J.G.; Rosenfeld, L.M.; Sulway, D.A.; Sayers, B.D.J.; Biele, J.; Sinclair, G.F.; Sahin, D.; Kling, L.; et al. Silicon Quantum Photonics in the Short-Wave Infrared: A New Platform for Big Quantum Optics. In Proceedings of the 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, 23–27 June 2019; p. 1. [CrossRef]

41. Arun, G.; Mishra, V. A review on quantum computing and communication. In Proceedings of the 2014 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking, Surat, India, 17–18 December 2014; pp. 1–5. [CrossRef]

42. Ding, Y.; Llewellyn, D.; Faruque, I.I.; Bacco, D.; Rottwitt, K.; Thompson, M.G.; Wang, J.; Oxenlowe, L.K. Quantum Entanglement and Teleportation Based on Silicon Photonics. In Proceedings of the 2020 22nd International Conference on Transparent Optical Networks (ICTON), Bari, Italy, 19–23 July 2020; pp. 1–4. [CrossRef]

43. De Adelhart Toorop, R.; Bazzocchi, F.; Merlo, L.; Paris, A. Constraining flavour symmetries at the EW scale I: The A 4 Higgs potential. *J. High Energy Phys.* **2011**, *2011*, 35. [CrossRef]

44. Khan, M.U.; Xing, Y.; Ye, Y.; Bogaerts, W. Photonic Integrated Circuit Design in a Foundry+Fabless Ecosystem. *IEEE J. Sel. Top. Quantum Electron.* **2019**, *25*, 1–14. [CrossRef]

45. Parhi, K.K.; Unnikrishnan, N.K. Brain-Inspired Computing: Models and Architectures. *IEEE Open J. Circuits Syst.* **2020**, *1*, 185–204. [CrossRef]

46. El-Kady, I.; Taha, M.M.R. Nano Photonic Sensors for Microdamage Detection: An Exploratory Simulation. In Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, HI, USA, 10–12 October 2005; Volume 2, pp. 1961–1966. [CrossRef]

47. Noda, S.; Asano, T.; Imada, M. Novel nanostructures for light: Photonic crystals. In Proceedings of the 2003 Third IEEE Conference on Nanotechnology, IEEE-NANO 2003, San Francisco, CA, USA, 12–14 August 2003; Volume 2, pp. 277–278. [CrossRef]

48. Vuckovic, J.; Yoshie, T.; Loncar, M.; Mabuchi, H.; Scherer, A. Nano-scale optical and quantum optical devices based on photonic crystals. In Proceedings of the 2nd IEEE Conference on Nanotechnology, Washington, DC, USA, 26–28 August 2002; pp. 319–321. [CrossRef]

49. De Marinis, L.; Cococcioni, M.; Castoldi, P.; Andriolli, N. Photonic Neural Networks: A Survey. *IEEE Access* **2019**, *7*, 175827–175841. [CrossRef]

50. Vandoorne, K.; Mechet, P.; Van Vaerenbergh, T.; Fiers, M.; Morthier, G.; Verstraeten, D.; Schrauwen, B.; Dambre, J.; Bienstman, P. Experimental demonstration of reservoir computing on a silicon photonics chip. *Nat. Commun.* **2014**, *5*, 3541. [CrossRef]

51. Coarer, F.D.-L.; Sciamanna, M.; Katumba, A.; Freiberger, M.; Dambre, J.; Bienstman, P.; Rontani, D. All-Optical Reservoir Computing on a Photonic Chip Using Silicon-Based Ring Resonators. *IEEE J. Sel. Top. Quantum Electron.* **2018**, *24*, 1–8. [CrossRef]

52. Lin, X.; Rivenson, Y.; Yardimci, N.T.; Veli, M.; Luo, Y.; Jarrahi, M.; Ozcan, A. All-optical machine learning using diffractive deep neural networks. *Science* **2018**, *361*, 1004–1008. [CrossRef]

53. Shi, B.; Calabretta, N.; Stabile, R. Image Classification with a 3-Layer SOA-Based Photonic Integrated Neural Network. In Proceedings of the 2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC), Fukuoka, Japan, 7 July 2019; pp. 1–3. [CrossRef]

54. Zuo, Y.; Li, B.; Zhao, Y.; Jiang, Y.; Chen, Y.-C.; Chen, P.; Jo, G.-B.; Liu, J.; Du, S. All-optical neural network with nonlinear activation functions. *Optica* **2019**, *6*, 1132. [CrossRef]

55. Harris, S.E. Electromagnetically Induced Transparency. *Phys. Today* **1997**, *50*, 36–42. [CrossRef]

56. Raab, E.L.; Prentiss, M.; Cable, A.; Chu, S.; Pritchard, D.E. Trapping of Neutral Sodium Atoms with Radiation Pressure. *Phys. Rev. Lett.* **1987**, *59*, 2631–2634. [CrossRef]

57. Grabowski, A.; Pfau, T. A lattice of magneto-optical and magnetic traps for cold atoms. In Proceedings of the 2003 European Quantum Electronics Conference. EQEC 2003 (IEEE Cat No.03TH8665), Munich, Germany, 22–27 June 2003; p. 274. [CrossRef]

58. Grossman, J.M.; Aubin, S.; Gomez, E.; Orozco, L.; Pearson, M.; Sprouse, G.; True, M. New apparatus for magneto-optical trapping of francium. In Proceedings of the Quantum Electronics and Laser Science Conference (IEEE Cat. No.01CH37172), Baltimore, MD, USA, 6–11 May 2001; p. 220, Technical Digest; Summaries of papers presented at the Quantum Electronics and Laser Science Conference; Postconference Technical Digest. [CrossRef]

59. The Ising Model. Available online: http://stanford.edu/~{}jeffjar/statmech2/intro4.html (accessed on 8 March 2020).

60. Singh, J.; Singh, M. Evolution in Quantum Computing. In Proceedings of the 2016 International Conference System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 25–27 November 2016; pp. 267–270. [CrossRef]

61. Yatsui, T.; Ohtsu, M. Development of nano-photonic devices and their integration by optical near field. In Proceedings of the IEEE/LEOS International Conference on Optical MEMs, Lugano, Switzerland, 20–23 August 2002; pp. 199–200. [CrossRef]

62. Olyaee, S.; Ebrahimpur, R.; Esfandeh, S. A hybrid genetic algorithm-neural network for modeling of periodic nonlinearity in three-longitudinal-mode laser heterodyne interferometer. In Proceedings of the 2013 21st Iranian Conference on Electrical Engineering (ICEE), Mashhad, Iran, 14–16 May 2013; pp. 1–5. [CrossRef]

63. Ren, Z.; Pope, S.B. The geometry of reaction trajectories and attracting manifolds in composition space. *Combust. Theory Model.* **2006**, *10*, 361–388. [CrossRef]

64. Shen, Y.; Harris, N.C.; Skirlo, S.; Prabhu, M.; Baehr-Jones, T.; Hochberg, M.; Sun, X.; Zhao, S.; LaRochelle, H.; Englund, D.; et al. Deep learning with coherent nanophotonic circuits. *Nat. Photon.* **2017**, *11*, 441–446. [CrossRef]

65. Howland, P.; Park, H. Generalizing discriminant analysis using the generalized singular value decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 995–1006. [CrossRef]

66. Zhang, Q.; Qin, Y. Adaptive Singular Value Decomposition and its Application to the Feature Extraction of Planetary Gearboxes. In Proceedings of the 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, China, 16–18 August 2017; pp. 488–492. [CrossRef]

67. Zhou, B.; Liu, Z. Method of Multi-resolution and Effective Singular Value Decomposition in Under-determined Blind Source Separation and Its Application to the Fault Diagnosis of Roller Bearing. In Proceedings of the 2015 11th International Conference on Computational Intelligence and Security (CIS), Shenzhen, China, 19–20 December 2015; pp. 462–465. [CrossRef]

68. Shen, Y.; Bai, Y. Statistical Computing with Integrated Photonics System. In Proceedings of the 2019 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC), Fukuoka, Japan, 16–17 July 2019; p. 1. [CrossRef]

69. Leelar, B.S.; Shivaleela, E.S.; Srinivas, T. Learning with Deep Photonic Neural Networks. In Proceedings of the 2017 IEEE Workshop on Recent Advances in Photonics (WRAP), Hyderabad, India, 18–19 December 2017; pp. 1–7. [CrossRef]

70. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–21. [CrossRef]

71. Saha, S. A Comprehensive Guide to Convolutional Neural Networks—The ELI5 Way. 2018. Available online: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53 (accessed on 11 March 2020).

72. Hamerly, R.; Sludds, A.; Bernstein, L.; Prabhu, M.; Roques-Carmes, C.; Carolan, J.; Yamamoto, Y.; Soljacic, M.; Englund, D. Towards Large-Scale Photonic Neural-Network Accelerators. In Proceedings of the 2019 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 7–11 December 2019; pp. 22.8.1–22.8.4. [CrossRef]

73. Hamerly, R.; Bernstein, L.; Sludds, A.; Soljačić, M.; Englund, D. Large-Scale Optical Neural Networks Based on Photoelectric Multiplication. *Phys. Rev. X* **2019**, *9*, 021032. [CrossRef]
74. Mrozek, T. Simultaneous Monitoring of Chromatic Dispersion and Optical Signal to Noise Ratio in Optical Network Using Asynchronous Delay Tap Sampling and Convolutional Neural Network (Deep Learning). In Proceedings of the 2018 20th International Conference on Transparent Optical Networks (ICTON), Bucharest, Romania, 1–5 July 2018; pp. 1–4. [CrossRef]
75. Bagherian, H.; Skirlo, S.; Shen, Y.; Meng, H.; Ceperic, V.; Soljacic, M. On-Chip Optical Convolutional Neural Networks. *arXiv* **2018**, arXiv:1808.03303.
76. Li, S.-L.; Li, J.-P. Research on Learning Algorithm of Spiking Neural Network. In Proceedings of the 2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing, Chengdu, China, 13–14 December 2019; pp. 45–48. [CrossRef]
77. Stewart, T.C.; Eliasmith, C. Large-Scale Synthesis of Functional Spiking Neural Circuits. *Proc. IEEE* **2014**, *102*, 881–898. [CrossRef]
78. Demertzis, K.; Iliadis, L. A Hybrid Network Anomaly and Intrusion Detection Approach Based on Evolving Spiking Neural Network Classification. In *E-Democracy, Security, Privacy and Trust in a Digital World*; Sideridis, A.B., Kardasiadou, Z., Yialouris, C.P., Zorkadis, V., Eds.; Springer International Publishing: Cham, Switzerland, 2014; Volume 441, pp. 11–23. [CrossRef]
79. Demertzis, K.; Iliadis, L.; Bougoudis, I. Gryphon: A semi-supervised anomaly detection system based on one-class evolving spiking neural network. *Neural Comput. Appl.* **2019**, *32*, 4303–4314. [CrossRef]
80. Demertzis, K.; Iliadis, L.; Spartalis, S. A Spiking One-Class Anomaly Detection Framework for Cyber-Security on Industrial Control Systems. In *Engineering Applications of Neural Networks*; Boracchi, G., Iliadis, L., Jayne, C., Likas, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 744, pp. 122–134. [CrossRef]
81. Demertzis, K.; Iliadis, L.; Anezakis, V.-D. A deep spiking machine-hearing system for the case of invasive fish species. In Proceedings of the 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Gdynia, Poland, 3–5 July 2017; pp. 23–28. [CrossRef]
82. Shastri, B.J.; Nahmias, M.A.; Tait, A.N.; Rodriguez, A.W.; Wu, B.; Prucnal, P.R. Spike processing with a graphene excitable laser. *Sci. Rep.* **2016**, *6*, srep19126. [CrossRef]
83. Lobo, J.L.; Del Ser, J.; Bifet, A.; Kasabov, N. Spiking Neural Networks and online learning: An overview and perspectives. *Neural Networks* **2019**, *121*, 88–100. [CrossRef]
84. Van Vaerenbergh, T.; Fiers, M.; Bienstman, P.; Dambre, J. Towards integrated optical spiking neural networks: Delaying spikes on chip. In Proceedings of the 2013 Sixth "Rio De La Plata" Workshop on Laser Dynamics and Nonlinear Photonics, Montevideo, Uruguay, 9–12 December 2013; pp. 1–2. [CrossRef]
85. Yang, Y.; Deng, Y.; Xiong, X.; Shi, B.; Ge, L.; Wu, J. Neuron-Like Optical Spiking Generation Based on Silicon Microcavity. In Proceedings of the 2020 IEEE 20th International Conference on Communication Technology (ICCT), Nanning, China, 28–31 October 2020; pp. 970–974. [CrossRef]
86. Nahmias, M.A.; Peng, H.-T.; De Lima, T.F.; Huang, C.; Tait, A.N.; Shastri, B.J.; Prucnal, P.R. A TeraMAC Neuromorphic Photonic Processor. In Proceedings of the 2018 IEEE Photonics Conference (IPC), Reston, VA, USA, 30 September–4 October 2018; pp. 1–2. [CrossRef]
87. Spoorthi, H.R.; Narendra, C.P.; Mohan, U.C. Low Power Datapath Architecture for Multiply—Accumulate (MAC) Unit. In Proceedings of the 2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 17–18 May 2019; pp. 391–395. [CrossRef]
88. Stelling, P.F.; Oklobdzija, V.G. Implementing multiply-accumulate operation in multiplication time. In Proceedings of the 13th IEEE Sympsoium on Computer Arithmetic, Asilomar, CA, USA, 6–9 July 1997; pp. 99–106. [CrossRef]
89. Bala, A.; Ismail, I.; Ibrahim, R.; Sait, S.M. Applications of Metaheuristics in Reservoir Computing Techniques: A Review. *IEEE Access* **2018**, *6*, 58012–58029. [CrossRef]
90. Demertzis, K.; Iliadis, L.; Pimenidis, E. Geo-AI to aid disaster response by memory-augmented deep reservoir computing. *Integr. Comput. Eng.* **2021**, *28*, 383–398. [CrossRef]
91. Li, S.; Pachnicke, S. Photonic Reservoir Computing in Optical Transmission Systems. In Proceedings of the 2020 IEEE Photonics Society Summer Topicals Meeting Series (SUM), Cabo San Lucas, Mexico, 13–15 July 2020; pp. 1–2. [CrossRef]
92. Vandoorne, K.; Fiers, M.; Verstraeten, D.; Schrauwen, B.; Dambre, J.; Bienstman, P. Photonic reservoir computing: A new approach to optical information processing. In Proceedings of the 2010 12th International Conference on Transparent Optical Networks, Munich, Germany, 29 June–1 July 2010; pp. 1–4. [CrossRef]
93. Laporte, F.; Katumba, A.; Dambre, J.; Bienstman, P. Numerical demonstration of neuromorphic computing with photonic crystal cavities. *Opt. Express* **2018**, *26*, 7955. [CrossRef]
94. Peng, B.; Özdemir, K.; Chen, W.; Nori, F.; Yang, L. What is and what is not electromagnetically induced transparency in whispering-gallery microcavities. *Nat. Commun.* **2014**, *5*, 5082. [CrossRef] [PubMed]
95. Shi, W.; Lin, J.; Sepehrian, H.; Zhalehpour, S.; Guo, M.; Zhang, Z.; Rusch, L.A. Silicon Photonics for Coherent Optical Transmissions (Invited paper). In Proceedings of the 2019 Photonics North (PN), Quebec City, QC, Canada, 21–23 May 2019; p. 1. [CrossRef]
96. Hughes, T.W.; Minkov, M.; Shi, Y.; Fan, S. Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **2018**, *5*, 864. [CrossRef]
97. Garrett, A.J.M.; Jaynes, E.T. Review: Probability Theory: The Logic of Science. *Law Probab. Risk* **2004**, *3*, 243–246. [CrossRef]

98. Ohnishi, R.; Wu, D.; Yamaguchi, T.; Ohnuki, S. Numerical Accuracy of Finite-Difference Methods. In Proceedings of the 2018 International Symposium on Antennas and Propagation (ISAP), Busan, Korea, 23–26 October 2018; pp. 1–2.

99. Serteller, N.F.O. Electromagnetic Wave Propagation Equations in 2D by Finite Difference Method: Mathematical Case. In Proceedings of the 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 11–13 October 2019; pp. 1–5. [CrossRef]

100. Xu, L.; Zhengyu, W.; Guohua, L.; Yinlu, C. Numerical simulation of elastic wave based on the staggered grid finite difference method. In Proceedings of the 2011 International Conference on Consumer Electronics, Communications and Networks (CECNet), Xianning, China, 11–13 April 2011; pp. 3283–3286. [CrossRef]

101. Hussain, M.A.; Tsai, T.-H. An Efficient and Fast Softmax Hardware Architecture (EFSHA) for Deep Neural Networks. In Proceedings of the 2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems (AICAS), Washington, DC, USA, 6–9 June 2021; pp. 1–4. [CrossRef]

102. Rao, Q.; Yu, B.; He, K.; Feng, B. Regularization and Iterative Initialization of Softmax for Fast Training of Convolutional Neural Networks. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8. [CrossRef]

103. Igarashi, H.; Watanabe, K. Complex Adjoint Variable Method for Finite-Element Analysis of Eddy Current Problems. *IEEE Trans. Magn.* **2010**, *46*, 2739–2742. [CrossRef]

104. Zhang, Y.; Negm, M.H.; Bakr, M.H. An Adjoint Variable Method for Wideband Second-Order Sensitivity Analysis Through FDTD. *IEEE Trans. Antennas Propag.* **2015**, *64*, 675–686. [CrossRef]

105. Walker, E.P.; Feng, W.; Zhang, Y.; Zhang, H.; McCormick, F.B.; Esener, S. 3-D parallel readout in a 3-D multilayer optical data storage system. In Proceedings of the International Symposium on Optical Memory and Optical Data Storage Topical Meeting, Waikoloa, HI, USA, 7–11 July 2002; pp. 147–149. [CrossRef]

106. Arai, Y. Vertical integration of radiation sensors and readout electronics. In Proceedings of the Melecon 2010—2010 15th IEEE Mediterranean Electrotechnical Conference, Valletta, Malta, 26–28 April 2010; pp. 1062–1067. [CrossRef]

107. Freiberger, M.; Katumba, A.; Bienstman, P.; Dambre, J. Training Passive Photonic Reservoirs with Integrated Optical Readout. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *30*, 1943–1953. [CrossRef]

108. Kim, I.; Vassilieva, O.; Akasaka, Y.; Palacharla, P.; Ikeuchi, T. Enhanced Spectral Inversion for Fiber Nonlinearity Mitigation. *IEEE Photon-Technol. Lett.* **2018**, *30*, 2040–2043. [CrossRef]

109. Umeki, T.; Kazama, T.; Ono, H.; Miyamoto, Y.; Takenouchi, H. Spectrally efficient optical phase conjugation based on complementary spectral inversion for nonlinearity mitigation. In Proceedings of the 2015 European Conference on Optical Communication (ECOC), Valencia, Spain, 27 September–1 October 2015; pp. 1–3. [CrossRef]

110. Williamson, I.A.D.; Hughes, T.W.; Minkov, M.; Bartlett, B.; Pai, S.; Fan, S. Reprogrammable Electro-Optic Nonlinear Activation Functions for Optical Neural Networks. *IEEE J. Sel. Top. Quantum Electron.* **2019**, *26*, 1–12. [CrossRef]

111. Li, J.; Dai, J. Z-Transform Implementations of the CFS-PML. *IEEE Antennas Wirel. Propag. Lett.* **2006**, *5*, 410–413. [CrossRef]

112. Watanabe, T. An optimized SAW chirp -Z Transform for OFDM systems. In Proceedings of the 2009 IEEE International Frequency Control Symposium Joint with the 22nd European Frequency and Time forum, Besancon, France, 20–24 April 2009; pp. 416–419. [CrossRef]

113. Zhang, Q.; Zong, Z. A New Method for Bistatic SAR Imaging Based on Chirp-Z Transform. In Proceedings of the 2014 Seventh International Symposium on Computational Intelligence and Design, Hangzhou, China, 13–14 December 2014; pp. 236–239. [CrossRef]

114. Chung, W.; Johnson, C.R. Characterization of the regions of convergence of CMA adapted blind fractionally spaced equalizer. In Proceedings of the Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284), Pacific Grove, CA, USA, 1–4 November 1998; Volume 1, pp. 493–497. [CrossRef]

115. Fan, F.; Hu, J.; Zhu, W.; Gu, Y.; Zhao, M. A multi-frequency optoelectronic oscillator based on a dual-output Mach-Zender modulator and stimulated brillouin scattering. In Proceedings of the 2017 IEEE Photonics Conference (IPC), Orlando, FL, USA, 1–5 October 2017; pp. 667–668. [CrossRef]

116. Magazzu, G.; Ciarpi, G.; Saponara, S. Design of a radiation-tolerant high-speed driver for Mach Zender Modulators in High Energy Physics. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; pp. 1–5. [CrossRef]

117. Passalis, N.; Mourgias-Alexandris, G.; Tsakyridis, A.; Pleros, N.; Tefas, A. Training Deep Photonic Convolutional Neural Networks with Sinusoidal Activations. *IEEE Trans. Emerg. Top. Comput. Intell.* **2019**, *5*, 384–393. [CrossRef]

118. Datta, S.; Antonio, R.A.G.; Ison, A.R.S.; Rabaey, J.M. A Programmable Hyper-Dimensional Processor Architecture for Human-Centric IoT. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2019**, *9*, 439–452. [CrossRef]

119. Kanerva, P. Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cogn. Comput.* **2009**, *1*, 139–159. [CrossRef]