Photonic Spiking Neural Networks with Highly Efficient Training Protocols for Ultrafast Neuromorphic Computing Systems

Dafydd Owen-Newns[1], Joshua Robertson[1], Matěj Hejda[1], and Antonio Hurtado[1]

[1]Institute of Photonics, SUPA Department of Physics, University of Strathclyde, Glasgow, UK.

[*]Corresponding author. Email: dafydd.owen-newns@strath.ac.uk

Abstract

Photonic technologies offer great prospects for novel, ultrafast, energy-efficient, and hardware-friendly neuromorphic (brain-like) computing platforms. Moreover, neuromorphic photonic approaches based on ubiquitous, technology-mature, and low-cost vertical-cavity surface-emitting lasers (VCSELs) (devices found in fiber-optic transmitters, mobile phones, and automotive sensors) are of particular interest. Given that VCSELs have shown the ability to realize neuronal optical spiking responses (at ultrafast GHz rates), their use in spike-based information-processing systems has been proposed. In this study, spiking neural network (SNN) operation, based on a hardware-friendly photonic system of just one VCSEL, is reported alongside a novel binary weight 'significance' training scheme that fully capitalizes on the discrete nature of the optical spikes used by the SNN to process input information. The VCSEL-based photonic SNN was tested with a highly complex multivariate classification task (MADELON) before its performance was compared using a traditional least-squares training method and an alternative novel binary weighting scheme. Excellent classification accuracies of $>94\%$ were achieved by both training methods, exceeding the benchmark performance of the dataset in a fraction of the processing time. The newly reported training scheme also dramatically reduces the training set size requirements and the number of trained nodes ($\leq 1\%$ of the total network node count). This VCSEL-based photonic SNN, in combination with the reported 'significance' weighting scheme, therefore grants ultrafast spike-based optical processing highly reduced training requirements and hardware complexity for potential application in future neuromorphic systems and artificial intelligence applications.

This is a peer reviewed, accepted author manuscript of the following research article: Owen-Newns, D., Robertson, J., Hejda, M., & Hurtado, A. (2023). Photonic spiking neural networks with highly efficient training protocols for ultrafast neuromorphic computing systems. *Intelligent Computing*, [0031]. https://doi.org/10.34133/icomputing.0031

1 Introduction

1.1 Photonic ANNs and Spiking VCSELs

Artificial neural networks (NNs) appear more frequently in the fast-developing field of artificial intelligence (AI), owing to their proven ability to provide high-level performance in numerous complex information processing tasks [1, 2] such as computer vision and data classification. NNs, which draw inspiration from networks of biological neurons in the brain, are formed of highly parallel structures of nodes (neurons) that realize numerous nonlinear transformations to achieve efficient processing and decision-making. The desire to move towards a more-than-Moore era of computing and diverge from traditional computing paradigms has also spurred reports of brain-inspired computing, where multiple systems of large-scale electronic artificial neural network systems, such as Loihi [3], TrueNorth [4] and BrainScaleS [5], exist. However, the number of studies demonstrating optical approaches to beyond-von Neumann computing is also growing given the appeal of the inherent properties of optical media. Optical-based systems can operate with very low component crosstalk, perform across large data bandwidth ranges, and utilize non-interacting frequencies of lights to achieve highly parallel division-multiplexing applications [6]. More importantly, photonic platforms can allow NNs to operate at low power consumption and ultrafast speed, overcoming the squeeze of power efficiency and clock speed faced by state-of-the-art chip-scale electronic components [7, 8]. The realization of light-based NNs, or the so-called photonic neural network (PNN) systems, is therefore of key interest to all future AI applications, where the benefits of increased operation speed and power efficiency are directly felt in the training and processing of large data volumes.

Despite the relative infancy of brain-inspired (neuromorphic) photonic systems, the field is advancing rapidly because of the attractive properties of optics that have inspired the production of multiple PNN accelerators [9–16]. These systems, based on technologies such as phase-change materials [9, 10], optical modulators [11, 16], and micro-ring weight banks [12–15], have been used to demonstrate different information processing tasks with photonic components. Neuromorphic photonic systems built using semiconductor laser (SL) technologies have also demonstrated intriguing neuron-like nonlinear dynamics that are key to the operation of PNN systems [17, 18]. Vertical-cavity surface-emitting lasers (VCSELs), which are well-established, commercially available, and increasingly ubiquitous devices (found in mobile phones, data centers, supermarket barcode scanners, etc.), are a type of neuromorphic SL technology. The ability of VCSELs to imitate the behavior of biological neurons was first proposed in [19], where it was shown that controllable spike activation could be achieved with incoming optical signals by exploiting the nonlinear dynamics surrounding the injection locking condition. It has also been shown that electronic modulation of a VCSEL can cause similar spiking behavior [20]. Neuron-like spiking responses were achieved at ultrafast (GHz) rates with spike widths of approximately 100 ps, making the spiking responses > 6 orders of magnitude faster than those of biological neurons and multiple orders of magnitude faster than those of some electronic spiking systems. VCSELs have exhibited several common neuronal behaviors (such as integration and fire spiking [21, 22], refractoriness, and rate coding [23]) in experimental realizations. Further, early demonstrations of neuromorphic information processing functionalities with VCSELs have been achieved using the spiking neuronal behaviors of VCSELs both experimentally,

including image processing (edge feature detection) [24–26], pattern recognition [22] and exclusive OR (XOR) operation [27], as well as theoretically through simulations of the Yamada and Spin-flip models (see [28] for a review). Moreover, single VCSEL devices have exhibited processing functionality when operated in combination with software-based Spiking Neural Networks (SNNs), achieving high classification accuracy in the MNIST handwritten digit recognition task [29]. VCSELs therefore represent an increasingly promising technological approach to spike-based PNNs that operate with ultrafast telecommunication wavelengths and optical signals, and are hardware-friendly, low-power, and low-cost.

Beyond the photonic imitation of spiking neuronal behaviors, reservoir computing (RC) has been demonstrated to be a powerful method for creating PNNs, yielding excellent performance in complex tasks, while also benefiting from hardware-friendly architectures. First developed in the early 2000s [30, 31], it was shown that in RC architectures (NN architectures with fixed-weight hidden layer connections), only the output layer requires training to achieve successful performance. In these RC architectures, nodes within the hidden layer are referred to as the "reservoir". Therefore, the reservoir is composed of unknown, fixed-strength, interconnected, nonlinear elements coupled to an output layer. Given the fixed nature of connections, RC limits the training requirements of NNs and significantly reduces the resources (computational power and time) required to train large networks of nodes for successful operations [30]. Numerous theoretical [32–34] and experimental [35–37] reports have been published on reservoir computers based on photonics platforms with VCSELs, which frequently feature [38–43]. VCSELs are photonic devices in which both time-delay reservoirs (TDRs) [40–43], which multiplex nodes over time to construct virtual networks and create memory through feedback connections, and spatial-temporal reservoirs [44, 45], which spatially multiplex nodes, have demonstrated successful operations. These systems have demonstrated impressive performance in numerous complex processing tasks, utilizing both off-the-shelf [43] and bespoke large-area (LA)-VCSEL designs [44] to realize photonic computing with continuous (nonspiking) VCSEL signals (see [46] for a review). More recently, we demonstrated for the first time a combination of a reservoir computing technique with a neuromorphic spiking VCSEL neuron [47]. In that report, we revealed that by using a time-multiplexing technique and masking inputs, we could interpret a single VCSEL neuron as a fixed-weight interconnected virtual SNN. This photonic system realizes both the reduced training requirements of an RC system and the sparse, all-or-nothing, binary spike-based representation of a full SNN. The VCSEL-based SNN operated with short temporal nodes (250 ps), fast optical spikes (100 ps), and a configurable network node number, which allowed it to successfully operate on the benchmark iris flower classification task [48] with excellent performance.

In this study, we show that we can not only use time multiplexing to create a VCSEL-based photonic SNN capable of a highly complex processing task, but more importantly, we can achieve high performance with significantly reduced training requirements (i.e., reduced training set sizes and number of trained nodes). This paper adopts the following structure: In Section 1.2 we introduce the experimental setup used in this study to create a photonic SNN and discuss its operating principles. In Section 1.3 we introduce the experimental classification task and explain the preparation of the data for injection into the VCSEL-based SNN. In Section 2.1 we discuss the results of the classification task and explore the training using a standard linear least-squares regression method. In Section 2.2

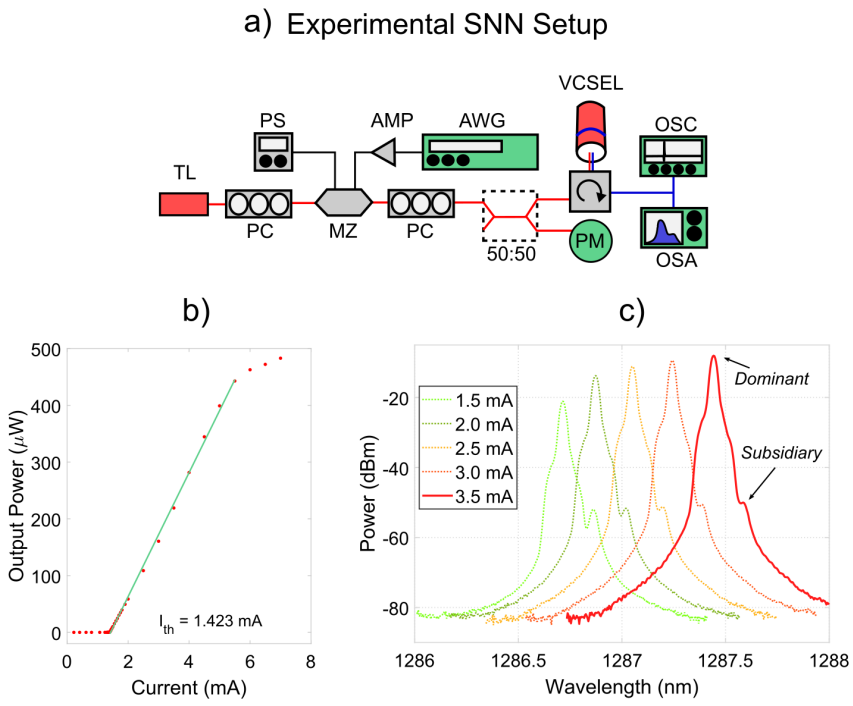

Figure 1: a) Experimental setup used to create the photonic SNN with a single VCSEL device. b) The lasing threshold (Light-Current) characteristics of the experimental VCSEL-neuron reveals a threshold current of 1.423 mA. c) The spectra of the experimental VCSEL-neuron at different bias currents. At the operating current (3.5 mA - red) the device exhibited two orthogonal polarization peaks at 1287.44 nm and 1287.59 nm. Measurements were taken with a room temperature (293 K) stabilized device.

we introduce an alternative binary weight training scheme and discuss its performance with respect to various training set sizes and node requirements. Finally, we present our conclusions in Section 3.

1.2 Photonic SNN Setup and Operation

In this paper, we propose a hardware-friendly photonic SNN based on a single time-multiplexed VCSEL-neuron. The experimental setup used to build and investigate the photonic SNN architecture is illustrated in Fig. 1 (a). The photonic SNN was built using a single telecom-wavelength VCSEL and commercially-sourced fiber-optic telecom components. The selected VCSEL exhibited wavelength emission within the key optical telecommunication O-band window (centered at approximately 1,300 nm) to showcase the full compatibility of the system with optical networking and data center technology. In the experimental setup, external light injection was provided by a tunable laser (TL) source, which was intensity-modulated using a 10 GHz Mach Zehnder modulator (MZ). The MZ is responsible for encoding the input information in the intensity of the optical injection before it enters the VCSEL-neuron. The MZ is driven by a DC power supply (PS) and a fast 12 GSa/s

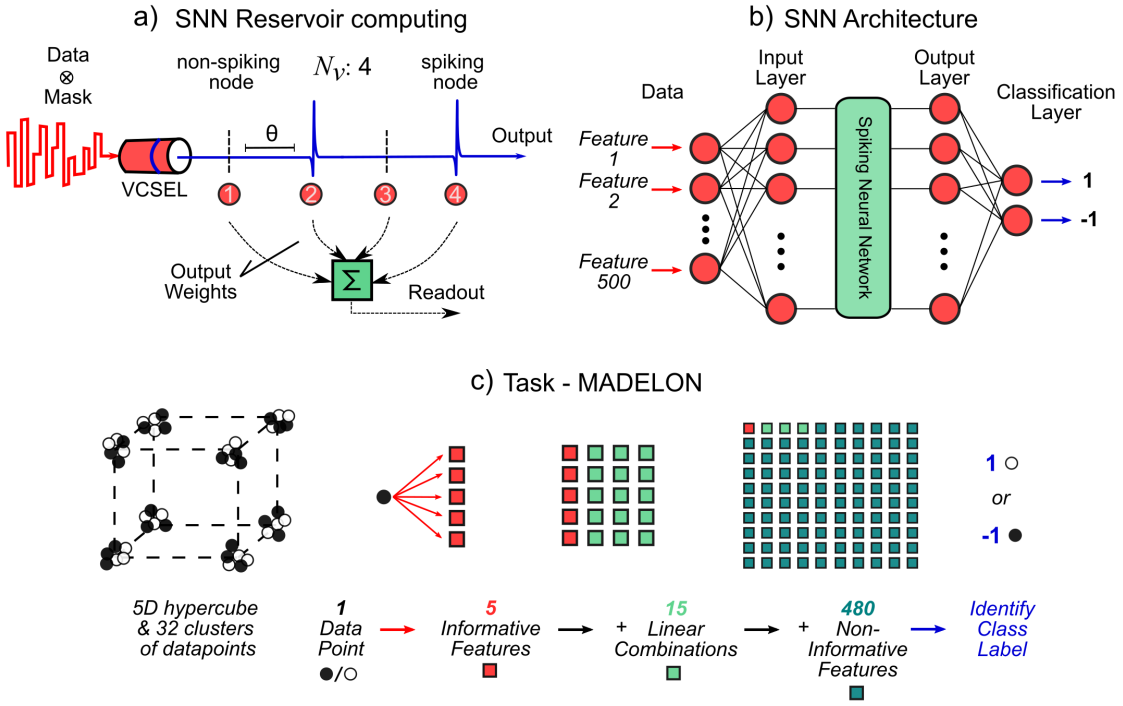

Figure 2: a) Photonic SNN operating principle. Masked input data is injected into the VCSEL forming the photonic SNN. The input information is time-multiplexed with each θ time-slot representing a virtual node (neuron) in the SNN. The time-multiplexed output of the VCSEL is interpreted as a binary node output, either spiking (1) or non-spiking (0). The output layer nodes are weighted offline to provide the photonic SNN readout. b) The architecture of the photonic SNN. 500 features per datapoint are input into the SNN system. In this experimental demonstration, the hidden layer was selected to be made up of either 2,048 or 4,096 virtual nodes. The output layer is trained offline to provide a datapoint classification (either +1 or -1). c) The MADELON classification task. Datapoints are placed on the corners of a 5-dimensional hypercube and randomly labeled -1 or +1. Each artificial data point has 5 informative features (each corresponding to 1 of 5 dimensions). Fifteen linear combinations of the original features are made creating 20 total features. Four hundred and eighty non-informative features (that carry no information of the label) are then added to the feature set. The task requires all 500 data features to be used to classify the initial data point label -1 or +1.

5 GHz arbitrary waveform generator (AWG) that generates the required input information. A 10 dB electrical amplifier (AMP) was used to increase the signal from the AWG, and a polarization controller (PC) was used to maximize TL light coupling into the MZ. The modulated injection was split by a fiber coupler to measure the injection power before the TL light entered the circulator and VCSEL. The output of the VCSEL was captured by the circulator and fed into the analysis, where it was measured using an optical spectrum analyzer (OSA) and a fast 40 GSa/s 16 GHz real-time oscilloscope (OSC). The threshold and spectral characteristics of the O-band VCSEL used in the photonic SNN system are shown in Figs 1 (b) & (c). When stabilized at room temperature (293 K), the VCSEL exhibited a threshold current of 1.423 mA and produced two orthogonal polarization

peaks in the fundamental transverse mode. In this study, with the selected operating current of 3.5 mA, these orthogonal polarization peaks occurred at 1287.44 nm and 1287.59 nm, producing a total output power of ∼220 μW. Using a PC, the optical injection was polarization-matched to the subsidiary (1287.59 nm) polarization mode of the VCSEL, with injection made using negative frequency detuning to induce polarization switching and injection locking. The modulating intensity is responsible for triggering high-speed (approximately 100 ps-long) neuron-like spiking responses near the injection-locking boundary.

In this study, as in [47], we combined the concepts of reservoir computing with a photonic spiking VCSEL-neuron to create a fully photonic SNN using significantly reduced hardware (a single laser device). As discussed in Section 1.1, reservoir computers are a type of artificial NN that host a hidden layer of interconnected nodes whose connections and weights are fixed. This means that in reservoir computers, only the output layer of the network must be trained to achieve high performance [30]. Furthermore, reservoir computers can use time multiplexing to interpret the output of one node as the output of several nodes by sampling at set times (θ). Here, using the VCSEL neuron, we apply the same concepts to interpret the output of a spiking VCSEL as an entire photonic SNN, see Fig 2 (a). The output of the VCSEL was time-multiplexed at intervals of θ to produce a controllable number of network nodes (N_v). The coupling of network nodes is achieved when the duration of the nodes (θ) is less than the timescale of the neuron-like spiking nonlinearity in the VCSEL (spikes can typically be triggered by strong modulations as short as 200 ps, with a refractory period of$\sim$ 1 ns [23]). This condition allows the leaky integrate-and-fire and refractory behaviors of the VCSEL neuron to link multiple nodes coupled in time, thereby forming a complex interconnected virtual network structure. As in [47] (and other VCSEL-based reservoir computers [43]), the input data were prepared by randomly masking the data point features and injecting them into the VCSEL as continuously modulated light. Following injection, each θ-long node delivers an all-or-nothing binary output, determined by the presence (or absence) of a fast 150 ps optical spike, effectively realizing the output of a fast photonic SNN. In this system, the output weights (the weights applied to each node) are trained and calculated offline to provide a readout of the system for the prescribed processing task (the task and training schemes are discussed in Sections 1.3 & 2).

In [47], we recently experimentally demonstrated that a VCSEL-neuron system can successfully realize SNN operation with fast optical spiking signals. We highlight the benchmark nonlinear iris flower classification task, comprising 150 nonlinearly separable data points (flowers), four features per data point (sepal & petal length/height), and three flower classes. However, using our VCSEL-based photonic SNN, we demonstrated that a very high overall average classification accuracy of >97 % could be achieved, demonstrating the powerful flexibility and performance of photonic SNNs. Importantly, in that study, the training of the output weights was achieved using an ordinary least squares regression method; hence, the node outputs were treated as if they were continuous (spiking nodes were assigned the value of 1.0, and others were assigned 0.0). This result required the offline calculation of all node weights as real number (floating point) values. In this report, we focus on solving a much more complex problem and applying an alternative training scheme whereby the discrete nature of binary all-or-nothing spiking responses can be used to improve the speed and efficiency of training. In the following sections, we introduce a new, high-complexity, nonlinear

classification task (MADELON) [49] and demonstrate the successful performance of the photonic SNN with two training methods: the aforementioned least-squares regression training method and an alternative node-significance training method, to highlight the attainment of notable improvements in both training speed and resource requirements.

1.3 Data Set Preparation

We further demonstrate the capability of the VCSEL-based photonic SNN by applying it to a highly complex classification task. The task used in this study is an artificial dataset named 'MADELON,' which was created to test feature selection methods as part of a NIPS 2003 challenge [49]. A simplified schematic of the MADELON data creation process is presented in Fig. 2 (c). The dataset comprises 32 clusters of datapoints placed on the vertices of a 5-dimensional (5D) hypercube. Two classes of data are created by randomly labeling each datapoint with the value -1 or +1. Each data point had five informative features (one for each dimension) that could be used to identify a data class. Using these five informative features, 15 linear combinations were created, increasing the feature set size to 20. Finally, 20 features were randomly scattered among 480 non-informative (no predictive power) features, creating a total of 500 features per data point. Therefore, this task is a multivariate, nonlinear, two-class problem with continuous input variables. The task is highly complex, given that a classification must be made using a feature set in which no feature by itself is informative. This problem is further complicated by the addition of non-informative features, which require the system to eliminate and ignore distractions.

The MADELON dataset contains 2,000 data points intended for training. However, to demonstrate the operation of the VCSEL-based photonic SNN, 300 data points (150 for each class) were used. Limiting the number of input points reduced the length of the time-multiplexed output generated during the experimental runs, thereby allowing us to test different network architectures with diverse node numbers while remaining within the memory of our experimental equipment. As the MADELON task is of much higher complexity than the preliminary iris flower classification task [47], a higher number of virtual nodes (N_v) were selected, namely 2,048 and 4,096, to test the classification performance of the photonic SNN.

The photonic SNN architecture used for this task is illustrated in Fig. 2 (b). Each of the 500 features was fed into the N_v input layer nodes via a random mask. The input data were masked by multiplying a column vector of all 500 features by a $500 \times N_v$ random matrix, resulting in a vector of N_v components. The masked data were fed into the hidden layer of the photonic SNN, where fixed connections and weights were created by the nonlinear temporal dynamic behavior of the VCSEL. Experimentally, randomly masked data were generated by the AWG and modulated by the MZ for optical injection of the VCSEL. The output layer of N_v nodes is then provided by the segmentation of the VCSEL spiking output time series at θ intervals. In this study, a node separation (θ) of 250 ps was used for the time-multiplex network nodes, creating multiple interacting nodes within the 1 ns timescale of nonlinearity in the VCSEL [22, 23]. The final classification layer of the SNN then provides a -1 or 1 label to the data following the offline training and weighting of the output nodes. The following section discusses the two training methods used in this study.

2 Results

First, the MADELON task was executed in a photonic VCSEL-based SNN using a network node count of $N_v = 2048$. This resulted in a total processing time of 512 ns per datapoint in the photonic SNN. The VCSEL was driven with a current of 3.5 mA ($\sim$2.5 times the lasing threshold current) at a temperature of 293 K and the injected light had a mean power of 142.6 μW and was injected into the VCSEL with an initial frequency detuning of -5.4 GHz with respect to the resonant wavelength of the VCSEL's subsidiary polarization mode peak. These conditions put the VCSEL in a state in which it was injection-locked to an externally injected signal. Near the injection-locking boundary, optical spikes could be triggered by varying the intensity of the modulated optical injection light.

In the second experimental run, the number of virtual (spiking) nodes in the network was increased to $N_v = 4096$ (1024 ns of processing time per datapoint). In the second run, the VCSEL was biased with a current of 3.45 mA at a set temperature of 293 K. The injected light signal, encoded with the input data, had a mean optical power of 161 μW, and was injected with an initial frequency detuning of -4.6 GHz (with respect to the resonance of the VCSEL's subsidiary polarization mode).

For a given data point, the output of the photonic SNN is a vector S with length $N_v + 8$. The eight additional virtual nodes were the result of zero padding at the end of the input data sequence, equivalent to a 2 ns reset of the system between inputs. Resetting the system between consecutive datapoint inputs prevented the previous datapoint from influencing the new datapoint and generating undesired spike activations. In the SNN output vector S, the i^{th} element was read as 1 if the i^{th} node spiked within the 250 ps-long time-multiplexed segment, and was assigned value 0 otherwise.

Figure 3 presents the experimentally measured time series depicting the optical input (Fig. 3 (a-b)) and output signals (Fig. 3 (c)) of the photonic SNN. The example shown in Fig. 3 is for a network architecture with 2,048 nodes. Time multiplexing was used to inject each masked datapoint into all (virtual) nodes of the photonic SNN (at a rate of 250 ps per virtual node). This created a time-varying signal for each datapoint, which was applied as the intensity modulation of the light injected into the SNN. During the experiment, all 300 masked datapoints (time-varying signals) were injected into the SNN sequentially, creating a large optical input timeseries. Figure 3 (a) depicts a part of the large optical input time series, showing 50 sequentially encoded masked datapoints from the MADELON dataset. Figure 3 (b) shows the first 512 ns of the optical input time series in Fig. 3 (a), which corresponds to the input signal encoding the first MADELON datapoint. As the amplitude of the optical input signal varies, it triggers the firing of optical spikes from the VCSEL at specific times (virtual nodes). Therefore, different input signals (masked datapoints) entering a photonic SNN produce different optical spike trains in the network. The optical spiking time series measured at the output of the SNN in response to the first injected MADELON dataset point (input signal in Fig. 3b) is shown in Fig. 3 (c).

Figure 4 shows a 2D temporal map that collects in one plot the spiking output from all virtual nodes of the photonic SNN for all 300 consecutively injected data points in the MADELON task. In Fig. 4 a yellow (blue) dot indicates a spiking (non-spiking) response from the VCSEL-based SNN, and the resulting 1 (0) binary output of the virtual node in vector S. The output vector S was

plotted horizontally for each input datapoint, revealing the trends in node activation across inputs and classes. The results for photonic SNNs configured with 2,048 and 4,096 virtual (spiking) nodes are plotted in Figures 4 (a) and (b), respectively. The red line through the center of the map separates the results for each distinct class (-1 & +1) of the input datapoint. Figures 4 (a) and (b) show that each input datapoint of the same class elicits a similar spike train at the output of the photonic SNN. These spike trains were sufficiently distinct from those obtained from other classes of datapoints to allow for a classification operation.

For the SNN with a 2,048-node architecture (Fig. 4 (a)), a total of 298 datapoints (149 per class) were used in the experimental run. This is because the first datapoint in class -1 was affected by an experimental artifact and was removed from the analysis. Therefore, one point from class $+1$ was removed to keep the size of each set equal, resulting in 298 points. For the network architecture with $Nv = 4096$ (Fig. 4 (b)), all 300 data points (150 per class) of the MADELON task were used to test the network operation.

Training was completed using a randomly selected subset of input datapoints and their corresponding spiking responses. The randomly selected subset contains an equal number of points (N_t) for each class for a total training set size of $2N_t$. In our photonic SNN, training is used to calculate the weights applied to the output layer nodes (i.e., the temporal spiking patterns produced by the network and the corresponding vector S). The output weights were determined to achieve the best performance on this set with the assumption that the training set is a good representation of the entire task dataset. The input datapoints not used in the training process were subsequently used to test the performance of the MADELON task. The output layer weights calculated during training were applied to the output layer to infer the class of a given test datapoint. The inferred datapoint class was then compared to the true class label, and the performance was measured using classification accuracy, which is the fraction of correctly classified datapoints across all tested datapoints. The training/testing process was repeated several times, and the training set was randomly selected each time to find an average value for the performance (a process known as random cross-validation).

2.1 Ordinary Least Squares

In this study, we first used the ordinary least squares (OLS) training method to test the performance of the photonic SNN in the MADELON task, a standard training method applied in other photonic RC systems [47, 50]. Here, the output spiking patterns from the VCSEL (and corresponding vector S) are treated as a binary sequence of floating-point 1.0s and 0.0s. Then, the output layer weights are found via linear regression where the values of each node (1.0 or 0.0) are arranged into matrix X (where the rows of X are the output vectors S). The labels (expected values) of each datapoint form the second-column vector Y. The weight calculation then involves determining the weight matrix W that minimizes Equation (1).

$$E_{rror}(W) = |Y - XW|^2 . \tag{1}$$

Equation (1) is solved using $W = \left(X^T X\right)^{-1} X^T Y$ (the Penrose inverse of X multiplied by Y). For this task, there are two output classes (-1 & +1). Therefore, the labels are two-element vectors

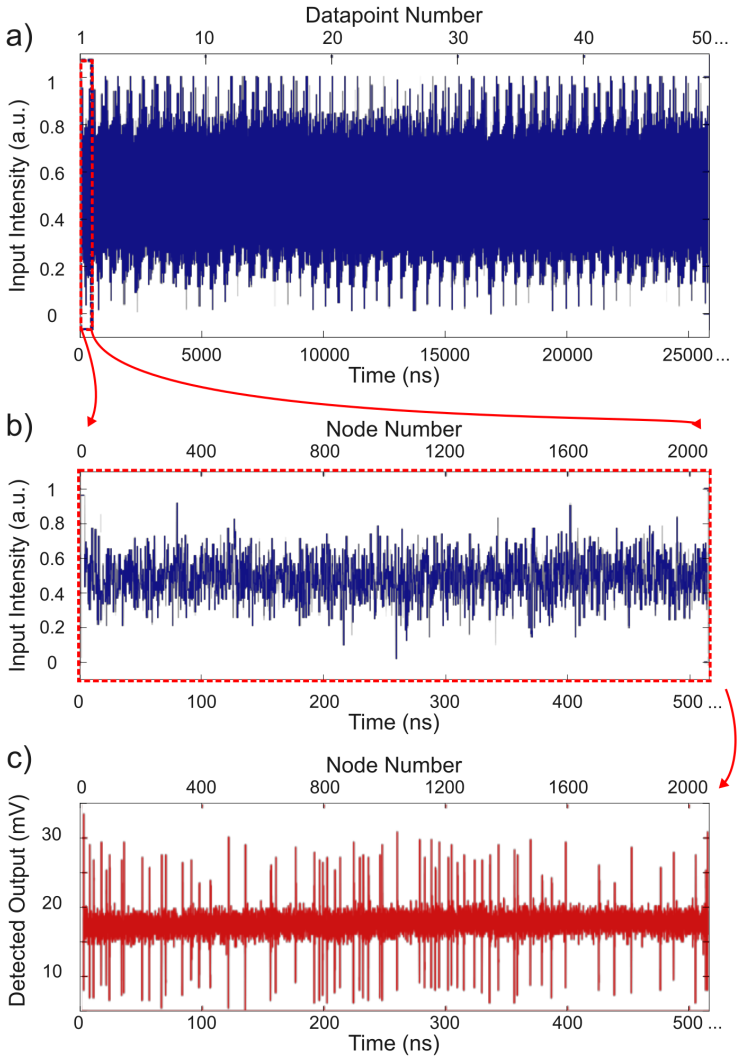

Figure 3: Time series showing the measured input (blue) and output (red) of the photonic SNN for a network architecture of 2,048 virtual nodes. a) A section of injection waveform showing the time-multiplexed intensity variations of 50 consecutive masked datapoints of the MADELON dataset. b) A 512 ns segment of the input timeseries corresponding to the first MADELON datapoint. Each of the 2,048 virtual nodes has a duration of 250 ps, creating the 512 ns input timeseries. c) The corresponding measurement of the SNN output when injected with the first MADELON datapoint.

in which the n^{th} element is set to 1 to denote the n^{th} class, with the other elements being zero. This also implies that the resulting weight matrix W has the dimensions $(N_v \times 2)$. Using W, the prediction of the class of a datapoint (producing a spike pattern S) can be obtained by considering $argmax(S * W)$. Applying weights to the spike pattern vector yielded a score for each class label, where the class with the highest score was the inferred label (prediction) of the photonic SNN system.

Using only 15 training datapoints per class (10 % of the total 300 input datapoints), the OLS

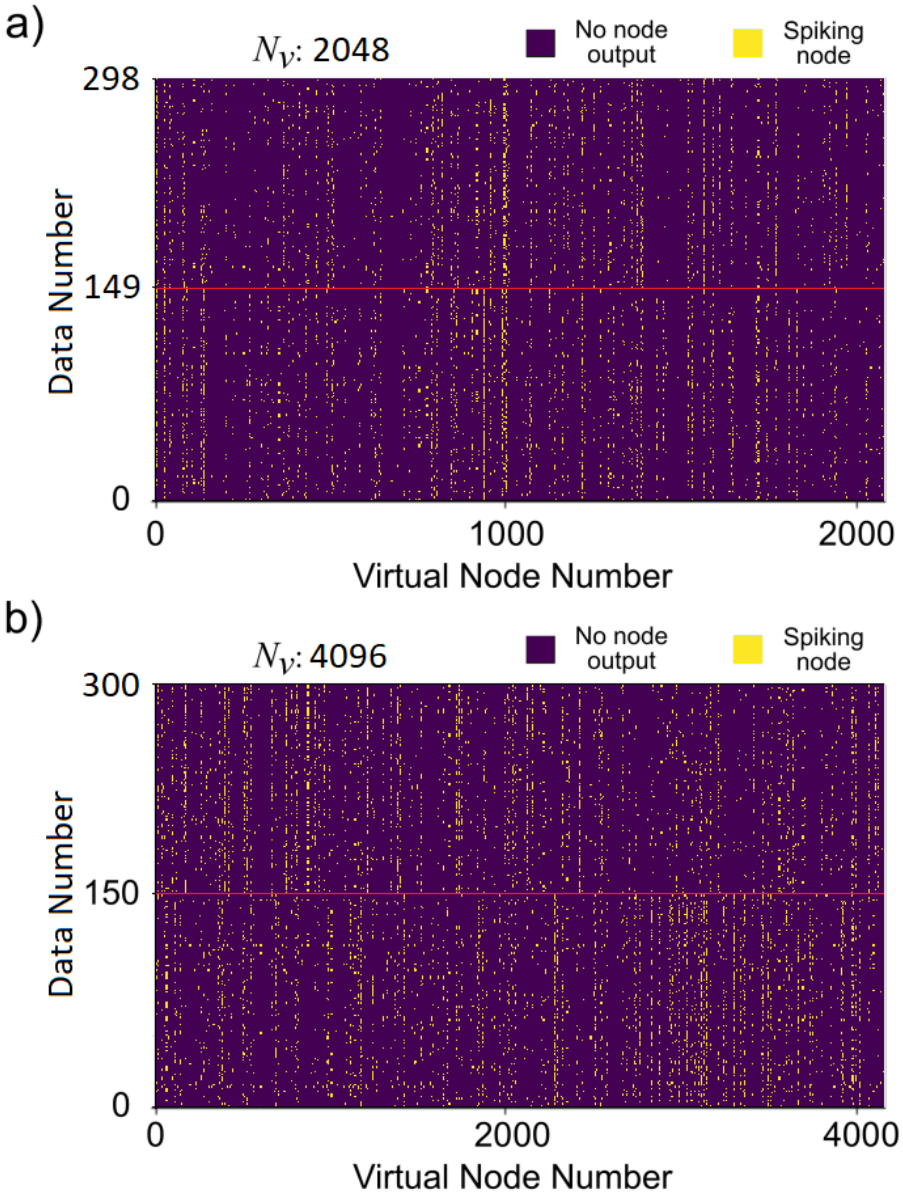

Figure 4: Temporal maps showing the optical spiking patterns produced by the SNN in response to the input dataset using a) 2,048 virtual nodes and b) 4,096 virtual nodes.

training method achieved a very high performance with a peak accuracy of 91% (for a 2,048-node architecture) and 94.5% (for a 4,096-node architecture). Example confusion matrices for $N_v = 2048$ and $N_v = 4096$ trained using this method and $N_t = 15$ are shown in Fig. 5. As expected, the network architecture with a higher number of nodes (N_v) provided a better overall classification performance. A larger number of nodes created more spiking responses from the VCSEL, allowing for more class-specific responses and better classification. The capability to adjust the node number at will is an

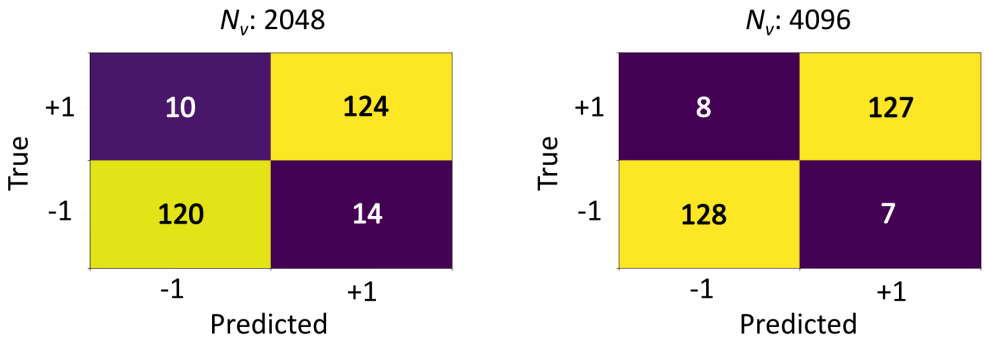

Figure 5: Example confusion matrices depicting the performance of a single instance of the photonic SNN when trained using the ordinary least squares method. A training set size of 15 ($N_t = 15$) was used to achieved accuracies of 91 % and 94.4 % for architectures of 2,048 and 4,096 nodes, respectively.

inherent advantage of the photonic SNN used in this study, as no physical or hardware changes are required to increase the network architecture. Instead, all that is required is the alteration of the input mask and a longer timeseries measurement. This makes the reported VCSEL-based photonic SNN an attractive and flexible system for fast, optical spike-based implementation of numerous processing tasks.

2.2 Binary Weight Training

As discussed in the previous section (Section 2.1) the standard OLS training method delivers a high classification performance during the MADELON task. However, the OLS method does not take full advantage of the discrete binary nature of the fast optical spiking responses in the photonic SNN. For example, it can be observed that in the ideal case of an SNN response, spiking patterns would be perfectly consistent between data of the same class, as well as completely distinct from data of alternative classes. In this ideal case, 2^k classes would only require the use of k nodes to be completely separate and identify datapoints. In the ideal case of the MADELON task, only one node per class is required for identification. In the case of experimental data, initially identifying nodes that have more predictive power (i.e., spike only for a particular class) prior to training would allow for the creation of weight matrices that retain high classification accuracy, despite the use and calculation of only a few nodes (a fraction of the total node count).

The following process is introduced to determine the weights for this SNN:

1. Count the number of spikes that occurred for each class in each node, denoted $s_{n,i}$ (# spikes from data of class i occurring in node n).

2. Calculate the significance score of each node towards each class according to Equation (2):

$$z_{n,i} = s_{n,i}^2 / \sum_i s_{n,i} \tag{2}$$

3. For each class, choose the top N_n scoring nodes and set their weights towards that class ($W_{n,i}$) to 1, and all remaining node weights to zero.

Applying the aforementioned algorithm provides the 'significance scores' for all network nodes, with higher scores obtained for nodes that contain not only more spikes, but nodes that have their spikes more concentrated in one specific class. The resulting output weight matrix W contains only nonzero entries on nodes that are strongly indicative of a specific class. The 2D temporal maps in Fig. 6 reveal the spike trains of a subset of network nodes, for the case when the photonic SNN was operated with a total node count of 2,048. This plot shows the significance scoring system applied to selected nodes. For example, Fig. 6 shows that node 844 (marked with an amber arrow) elicited an optical spike for 57 of the total 298 input datapoints used in this experimental run. In this case, the spiking responses from $n = 844$ are evenly distributed across each class (31 spikes for class -1 and 26 for class +1). In this specific case, the significance scores for this node are $z_{844,-1} = 16.9$ (class -1) and $z_{844,+1} = 11.9$ (class +1). In parallel, Fig. 6 shows that Node 885 (marked with a red arrow) fired fewer optical spikes overall. A total of 30 responses were elicited. However, because they were less evenly distributed (and hence more indicative of a particular class), the node scores $z_{885,-1} = 2.1$ (class -1) and $z_{885,+1} = 16.1$ (class +1), a class +1 significance score similar to that of node 844. Finally, when the appearance of fast optical spikes is much more biased towards one specific class and is more likely to occur overall, the resulting significance score z is much higher. For example, in node 853 (marked with a green arrow), 103 spikes occurred, with 88 of them occurring in class -1, yielding scores of $z_{853,-1} = 75.1$ and $z_{853,+1} = 2.1$. With this method, the weight $W_{853,-1}$ of node 853 is set to 1, as if a spike was detected in node 853, and it is likely that the datapoint is of class -1. In this example, the weights of the other highlighted nodes ($W_{844,-1}$ & $W_{885,-1}$) are set to zero.

For a data point to be classified, the spike-output vector S must be multiplied by the weight matrix W_i for each class. The class with the highest sum value was the prediction class of the data point. Because the output layer weights are binary, this is equivalent to applying a filter to the optical spike trains and counting the number of remaining spikes. Using this method, the performance can be tuned by selecting the number of nodes (N_n) for training (training node number). Using a fraction of the total number of available nodes can result in equal (if not better) performance, while keeping the weight matrix and number of weight calculations sparse.

Fig. 7 shows the resulting confusion matrices of the MADELON task when the significance-training algorithm was run with 15 training datapoints ($N_t = 15$). For the photonic SNN architecture with 2,048 nodes, a peak accuracy of 92.5 % is obtained, whereas for the 4,096-node architecture a peak accuracy of 95.2 % is achieved. Again, as expected, the architecture with a larger number of nodes provided a higher classification performance. Comparing this performance to the OLS training method, also run at $N_t = 15$, we found that significance training shows a small improvement in peak accuracy in both cases (2,048- and 4,096-node architectures). This improvement in performance shows that despite the same training set size, by considering only nodes significant to specific classes, we can achieve a similar or higher classification performance. Furthermore, it is important to highlight that the training node number used in the significance scoring method was substantially lower than that used in the OLS method. In the novel training method, only the top 20 most

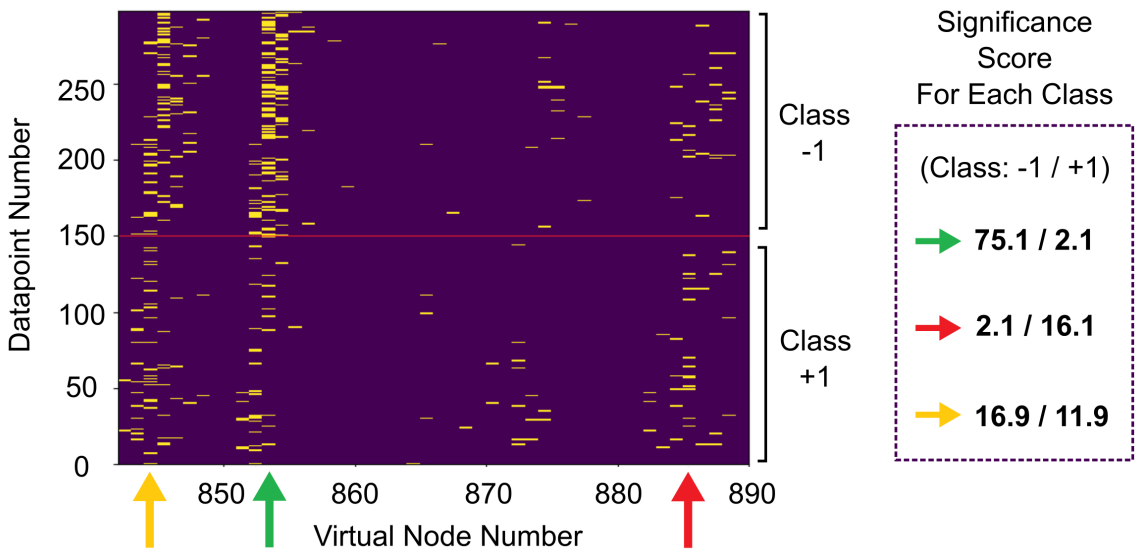

Figure 6: Section of a 2D temporal map for a network architecture of 2,048 nodes (between nodes 842 and 890). The arrows mark nodes 844 (orange), 853 (green) and 885 (red). In node 853, 88 spikes occurred for class -1, and 15 for class +1, giving it a high significance score of 75.1 (for class -1) and 2.1 (for class +1). Node 844 has significance scores of 16.9 (for class -1) and 11.9 (for class +1). Node 885 has significance scores of 2.1 (for class -1) and 16.1 (for class 1). Node 853 has the highest significance score meaning it would likely be used to perform a classification.

significant nodes ($N_n = 20$) were used in both reported SNN architecture cases (2,048 and 4,096). Therefore, the novel training method uses only 1% and 0.5% (for 2,048 and 4,096 nodes, respectively) of the total nodes used for OLS training, where every node in the network is trained ($N_n = N_v$). This result indicates that training fewer nodes overall is not only computationally efficient, but does not directly hinder the performance of the system (and in this case, increases classification performance). However, these newly presented classification performance values are not within the limits of the photonic SNN with this training scheme. In Figs. 8 and 9, we consider a larger range of training set sizes (N_t), and compare their optimal performance as well as their optimal training node number (N_n).

Increasing the number of nodes leads to better performance because of the increased likelihood of a node spiking more strongly for one class than for another. This was caused by the longer-range connections formed in the SNN owing to the refractory behavior of the spiking dynamics in the VCSEL.

Figs. 8 (a) and 9 (a) show the peak accuracy achieved in the MADELON classification task as a function of the training set size used for the 2,048 and 4,096 network architectures, respectively. Figs. 8 (b) and 9 (b) show the optimal training node number (N_n) used to attain the optimal accuracy.

Both Figs. 8 and 9 reveal two key results. First, the photonic SNN achieved very high accuracy levels, despite the complexity of the MADELON task and its large number of features. Specifically, a maximum accuracy of 94.4% was achieved for the 2,048-node architecture, and a maximum accuracy

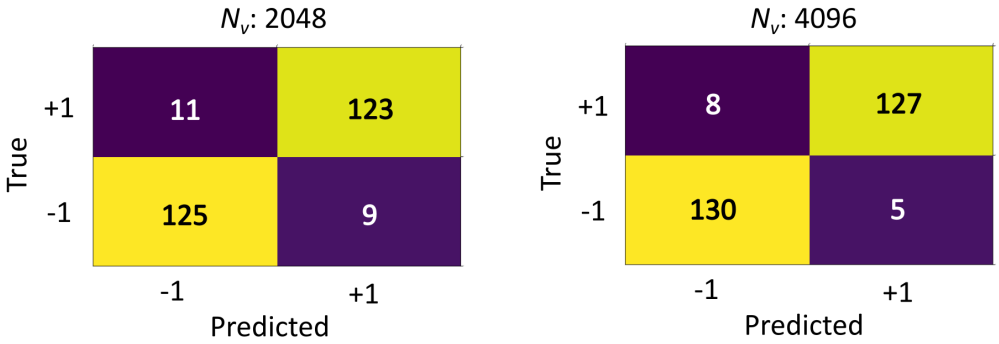

Figure 7: Example confusion matrices depicting the performance of a single instance of the photonic SNN when trained using the significance score method. A training set size of 15 ($N_t = 15$) was used to achieve accuracies of 92.5 % and 95.2 % for architectures of 2,048 and 4,096 nodes, respectively. Only the 20 most significant nodes ($N_n = 20$) were used to train the SNN.

of 95.7% was achieved for the 4,096-node architecture. The performance of this photonic SNN system was higher than that of previous software-based machine learning algorithms reported following the NIPS 2003 challenge [49], which reached accuracy levels of up to 93.78 %. Importantly, a key difference is that a photonic SNN is applied to the MADELON task with a hardware-friendly implementation (using only one VCSEL), low-power operation (sub-pJ energy per spike, $\sim 150\,\mu$W average optical powers, and 3.5 mA of applied bias current), and ultrafast performance (250 ps/node yielding a total processing time of 512 ns and 1024 ns per datapoint for 2,048 and 4,096 nodes, respectively). Second, the performance of the photonic SNN increases with the training set size (associated with the best-fit training), and the optimal value of N_n may slightly decrease with the training set size. This result was somewhat expected, as training with more datapoints typically improves the system performance, resulting in a lower dependence on high node numbers. Thirdly, Figs. 8 and 9 show that the optimal training node number (N_n) remains consistently low while the system retains very high performance, across all training set sizes (N_t) (from 1 to 100). Only small node numbers, as low as < 10 (out of the total node count N_v), must be considered to successfully achieve high performance in the MADELON task. This means that, independent of the training set size, the highly reduced training requirements of the significance scoring method remain a key benefit of the alternative training scheme. Finally, as illustrated in Figs. 8 and 9, very small training set sizes, $N_t \leq 10$ (≤ 3 % of the total dataset inputs), are capable of effectively training the photonic SNN system and achieving high accuracy ($> 90\%$) in the MADELON task. Comparing the previous OLS results for a network architecture of 4,096 (from 2.1) with those in Fig. 9, we find that at training set sizes < 15, the significance scoring method achieves similar levels of classification accuracy. Making use of the novel training approach can, therefore, enable a similar performance to the OLS method, but with the remarkable benefit of being more flexible in terms of the number of training datapoints.

The resulting (binary) weight matrix has several advantages over those produced by typical methods: it is very sparse–all but a few elements are nonzero, making this require very little electronic storage, and as the nonzero elements are only 1 (no floating-point numbers are used), the weights

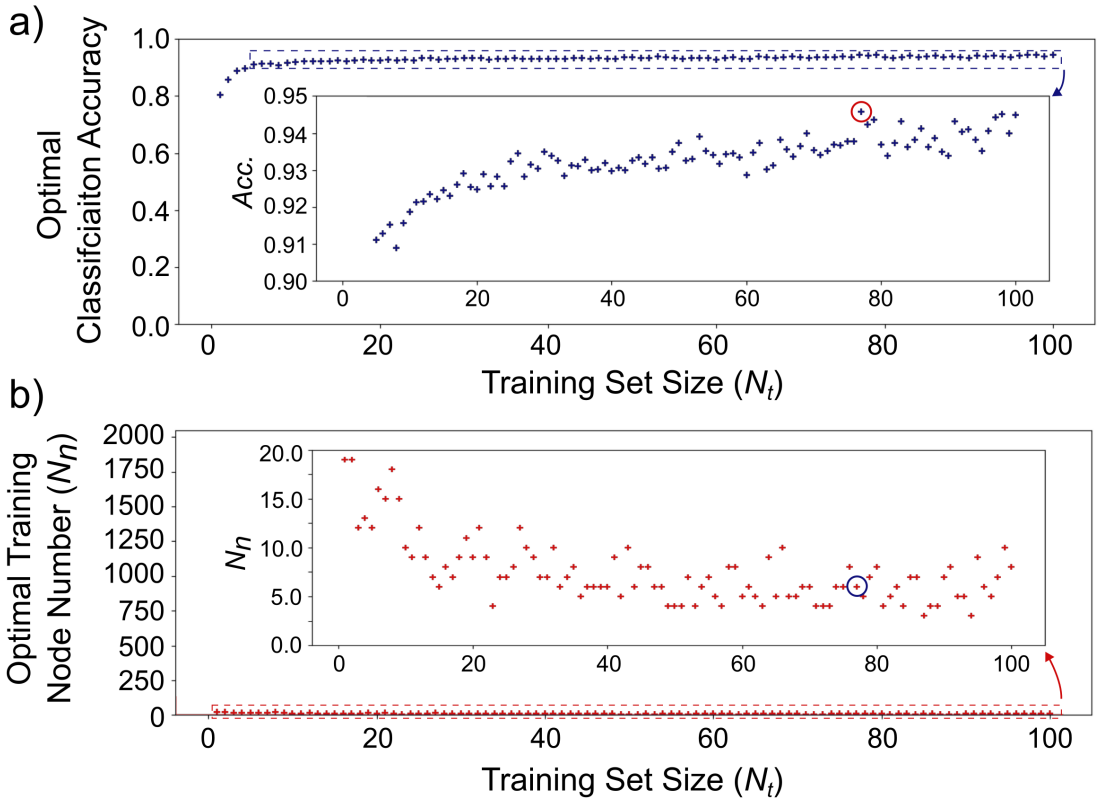

Figure 8: Photonic SNN performance versus the training set size, using a 2,048-node architecture. a) The optimal classification accuracy (averaged over 10 random selections of the training set) is plotted against training set size (N_t, blue). The inset highlights the consistent high performance over various training set sizes (N_t ranges from 1 to 100). b) The number of training nodes (N_n) used to achieve the corresponding optimal accuracy, plotted against increasing training set size (red). The peak accuracy, 94.4%, occurs when $N_n = 6$ and $N_t = 76$ (circled).

could potentially be directly applied in optical hardware without the need for precision modulation of the signal. Compared to traditional training methods, the algorithm for obtaining this weight matrix is much simpler: typical training methods use costly operations, such as matrix inversions, to calculate floating-point weights, whereas this method simply counts the number of spikes occurring at each node using low-cost computations.

3 Conclusion

This study demonstrates the high classification performance of a novel laser-based photonic spiking neural network for tackling a highly complex, multivariate, nonlinear classification task (MADELON) with 500 datapoint features. Importantly, this study also introduces a novel 'significance' training approach that makes use of binary weights (0 or 1), and that leverages on the advantages of the discrete optical spiking signals found in the photonic SNN. The experimental approach to

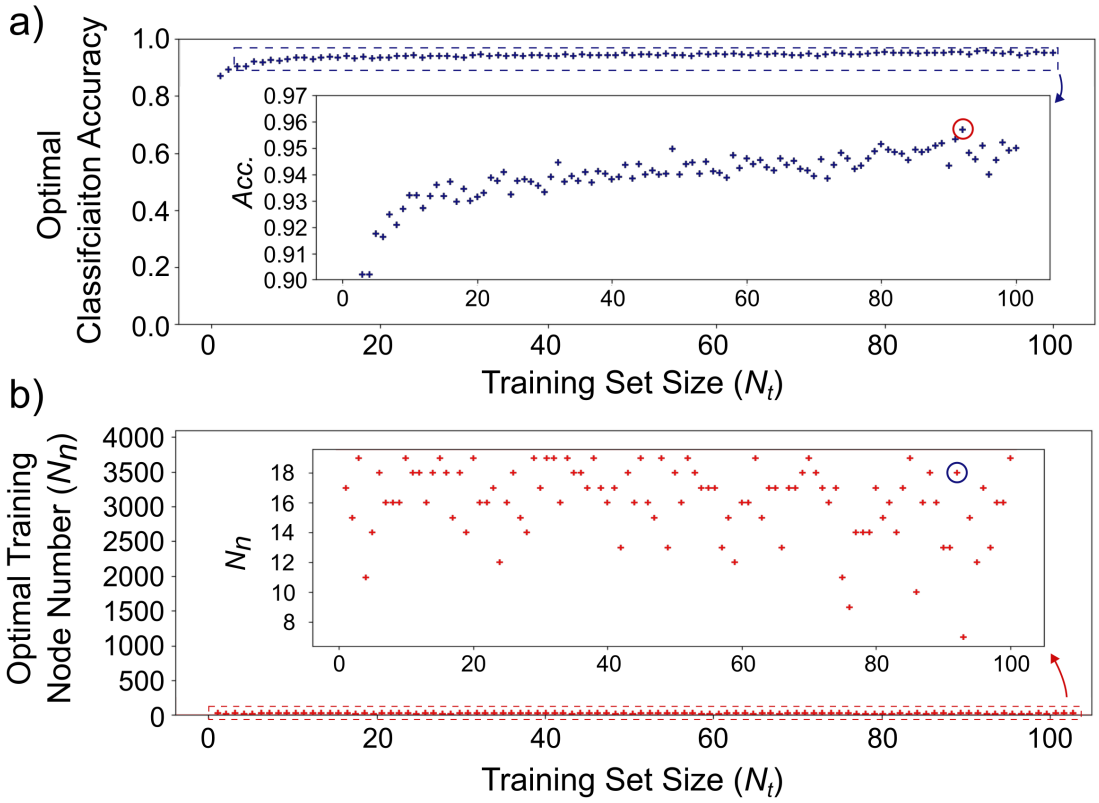

Figure 9: Photonic SNN performance versus the training set size, using a 4,096-node architecture. a) The optimal classification accuracy (averaged over 10 random selections of the training set) is plotted against training set size (N_t, blue). The inset highlights the consistent high performance over various training set sizes (N_t ranges from 1 to 100). b) The number of training nodes (N_n) used to achieve the corresponding optimal accuracy, plotted against increasing training set size (red). The peak accuracy, 95.7%, occurs when $N_n = 18$ and $N_t = 91$ (circled).

an SNN combines the spiking dynamics of a VCSEL with a novel network architecture inspired by the reservoir computing paradigm to process data entirely optically at extremely high speeds (GHz rates). The SNN uses all-optical neuron-like spikes to create a time-multiplexed feed-forward spiking neural network in which the values of each time-multiplexed (virtual) node are linked through the VCSEL's nonlinear temporal dynamics. The computational power of the photonic SNN was demonstrated using an OLS method of weight training. We demonstrated that, by training the output layer weights with this OLS approach, we could achieve high accuracies of up to 91% and 94.4% for the MADELON classification task using SNN architectures with 2,048 and 4,096 nodes, respectively.

Additionally, we introduced a new 'significance' training approach that assigns binary weights to (optical spiking) nodes according to their overall usefulness and significance score. In this approach, only high-significance scoring nodes, that is, nodes that spike frequently for one class but not others, were considered and used for network training and performance evaluation. We showed that only a very small fraction ($\leq 1\%$ in the presented case) of the total number of nodes (in the output layer)

were required to classify the data successfully. We demonstrated that classification accuracies of 94.4% and 95.7% could be achieved using this new training method. The accuracies provided by the significance-training approach showed an improvement over those achieved by the OLS method, while also significantly reducing the number of training nodes. Moreover, the photonic SNN demonstrated a classification performance that improved upon the benchmark accuracy (93.78%) achieved by software-implemented NNs reported by the dataset authors [49]. Additionally, we demonstrated that a photonic SNN trained with the new significance method could realize high-level performance with small training set sizes (<10 datapoints), further reducing the overall resources necessary for training the optical system.

Finally, the proposed photonic SNN offers several inherent physical and computational benefits over traditional digital semiconductor processing systems, notably ultrafast performance (250 ps/node), low-power usage ($\sim 150\,\mu$W average optical power, and 3.5 mA of applied bias current), and hardware-friendly implementation (using just one VCSEL to process all virtual nodes). Furthermore, a VCSEL-based photonic SNN can adjust the performance and processing rate by changing the number of virtual nodes used in the system, which can be performed arbitrarily and on the fly during pre-processing. In conclusion, we believe these results open possibilities for further photonics-based processing systems that run and operate entirely on optical hardware, and that are capable of solving highly complex tasks with high accuracy and ultrafast, energy-efficient operation.

Acknowledgments

Author Contributions

D. Owen-Newns performed pre- and post-processing of all data. D. Owen-Newns & J. Robertson performed experimental runs on a photonic system. M. Hejda contributed to data postprocessing and visualization. A. Hurtado supervised all research efforts. All the authors contributed equally to the writing of this manuscript.

Funding

The authors acknowledge that this work was supported by the UKRI Turing AI Acceleration Fellowships Programme (EP/V025198/1), by the European Commission (Grant 828841-ChipAI-H2020-FETOPEN-2018-2020), and UK EPSRC (EP/N509760/1, EP/P006973/1).

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this article.

Data Availability

All data underpinning this publication are openly available from the University of Strathclyde KnowledgeBase at https://doi.org/x.xxxxxxx. For the purpose of Open Access, the author has applied a

References

[1] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, p. 53, Dec. 2021, ISSN: 2196-1115. DOI: 10.1186/s40537-021-00444-8. [Online]. Available: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8.

[2] N. L. Kazanskiy, M. A. Butt, and S. N. Khonina, "Optical Computing: Status and Perspectives," *Nanomaterials*, vol. 12, no. 13, p. 2171, Jun. 2022, ISSN: 2079-4991. DOI: 10.3390/nano12132171. [Online]. Available: https://www.mdpi.com/2079-4991/12/13/2171.

[3] M. Davies *et al.*, "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018, ISSN: 0272-1732. DOI: 10.1109/MM.2018.112130359. [Online]. Available: http://ieeexplore.ieee.org/document/8259423/.

[4] M. V. DeBole *et al.*, "TrueNorth: Accelerating From Zero to 64 Million Neurons in 10 Years," *Computer*, vol. 52, no. 5, pp. 20–29, May 2019, ISSN: 0018-9162. DOI: 10.1109/MC.2019.2903009. [Online]. Available: https://ieeexplore.ieee.org/document/8713821/.

[5] C. Pehle *et al.*, "The BrainScaleS-2 accelerated neuromorphic system with hybrid plasticity," Jan. 2022. arXiv: 2201.11063. [Online]. Available: http://arxiv.org/abs/2201.11063.

[6] C. Brackett, "Dense wavelength division multiplexing networks: principles and applications," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 6, pp. 948–964, 1990, ISSN: 07338716. DOI: 10.1109/49.57798. [Online]. Available: http://ieeexplore.ieee.org/document/57798/.

[7] P. R. Prucnal, B. J. Shastri, and M. C. Teich, *Neuromorphic Photonics*, P. R. Prucnal and B. J. Shastri, Eds. CRC Press, May 2017, ISBN: 9781315370590. DOI: 10.1201/9781315370590. [Online]. Available: https://www.taylorfrancis.com/books/9781498725248.

[8] D. A. B. Miller, "Attojoule Optoelectronics for Low-Energy Information Processing and Communications," *J. Lightwave Technol.*, vol. 35, no. 3, pp. 346–396, Feb. 2017. [Online]. Available: http://www.osapublishing.org/jlt/abstract.cfm?URI=jlt-35-3-346.

[9] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature*, vol. 569, no. 7755, pp. 208–214, May 2019, ISSN: 0028-0836. DOI: 10.1038/s41586-019-1157-8. [Online]. Available: http://www.nature.com/articles/s41586-019-1157-8.

[10] J. Feldmann *et al.*, "Parallel convolutional processing using an integrated photonic tensor core," *Nature*, vol. 589, no. 7840, pp. 52–58, Jan. 2021, ISSN: 0028-0836. DOI: 10.1038/s41586-020-03070-1. [Online]. Available: http://www.nature.com/articles/s41586-020-03070-1.

[11] S. Xu, J. Wang, R. Wang, J. Chen, and W. Zou, "High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays," *Optics Express*, vol. 27, no. 14, p. 19 778, Jul. 2019, ISSN: 1094-4087. DOI: 10.1364/OE.27.019778. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=oe-27-14-19778.

[12] A. N. Tait *et al.*, "Neuromorphic photonic networks using silicon photonic weight banks," *Scientific Reports*, vol. 7, no. 1, p. 7430, Dec. 2017, ISSN: 2045-2322. DOI: 10.1038/s41598-017-07754-z. [Online]. Available: http://www.nature.com/articles/s41598-017-07754-z.

[13] T. F. de Lima *et al.*, "Machine Learning With Neuromorphic Photonics," *Journal of Lightwave Technology*, vol. 37, no. 5, pp. 1515–1534, Mar. 2019, ISSN: 0733-8724. DOI: 10.1109/JLT.2019.2903474. [Online]. Available: https://ieeexplore.ieee.org/document/8662590/.

[14] A. Mehrabian, Y. Al-Kabani, V. J. Sorger, and T. El-Ghazawi, "PCNNA: A Photonic Convolutional Neural Network Accelerator," in *2018 31st IEEE International System-on-Chip Conference (SOCC)*, IEEE, Sep. 2018, pp. 169–173, ISBN: 978-1-5386-1491-4. DOI: 10.1109/SOCC.2018.8618542. [Online]. Available: https://ieeexplore.ieee.org/document/8618542/.

[15] F. Ashtiani, A. J. Geers, and F. Aflatouni, "An on-chip photonic deep neural network for image classification," *Nature*, vol. 606, no. 7914, pp. 501–506, Jun. 2022, ISSN: 0028-0836. DOI: 10.1038/s41586-022-04714-0. [Online]. Available: https://www.nature.com/articles/s41586-022-04714-0.

[16] H. Zhang *et al.*, "An optical neural chip for implementing complex-valued neural network," *Nature Communications*, vol. 12, no. 1, p. 457, Dec. 2021, ISSN: 2041-1723. DOI: 10.1038/s41467-020-20719-7. [Online]. Available: http://www.nature.com/articles/s41467-020-20719-7.

[17] Z. Chen *et al.*, "Deep Learning with Coherent VCSEL Neural Networks," Jul. 2022. arXiv: 2207.05329. [Online]. Available: http://arxiv.org/abs/2207.05329.

[18] P. R. Prucnal, B. J. Shastri, T. Ferreira de Lima, M. A. Nahmias, and A. N. Tait, "Recent progress in semiconductor excitable lasers for photonic spike processing," *Advances in Optics and Photonics*, vol. 8, no. 2, p. 228, Jun. 2016, ISSN: 1943-8206. DOI: 10.1364/AOP.8.000228. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=aop-8-2-228.

[19] A. Hurtado and J. Javaloyes, "Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems," *Applied Physics Letters*, vol. 107, no. 24, p. 241 103, 2015, ISSN: 0003-6951. DOI: 10.1063/1.4937730. [Online]. Available: http://aip.scitation.org/doi/10.1063/1.4937730.

[20] J. Robertson, E. Wade, and A. Hurtado, "Electrically controlled neuron-like spiking regimes in vertical-cavity surface-emitting lasers at ultrafast rates," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 6, 2019. DOI: 10.1109/JSTQE.2019.2899040.

[21] J. Robertson, E. Wade, Y. Kopp, J. Bueno, and A. Hurtado, "Towards Neuromorphic Photonic Networks of Ultrafast Spiking Laser Neurons," *IEEE Journal of Selected Topics in Quantum Electronics*, 2019. DOI: 10.1109/JSTQE.2019.2931215.

[22] J. Robertson, M. Hejda, J. Bueno, and A. Hurtado, "Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons," *Scientific Reports*, vol. 10, no. 1, 2020. DOI: 10.1038/s41598-020-62945-5.

[23] M. Hejda, J. Robertson, J. Bueno, and A. Hurtado, "Spike-based information encoding in vertical cavity surface emitting lasers for neuromorphic photonic systems," *JPhys Photonics*, vol. 2, no. 4, p. 44 001, 2020. DOI: 10.1088/2515-7647/aba670. [Online]. Available: https://doi.org/10.1088/2515-7647/aba670.

[24] J. Robertson, Y. Zhang, M. Hejda, J. Bueno, S. Xiang, and A. Hurtado, "Image edge detection with a photonic spiking VCSEL-neuron," *Optics Express*, vol. 28, no. 25, pp. 37 526–37 537, 2020. DOI: 10.1364/OE.408747.

[25] Y. Zhang, J. Robertson, S. Xiang, M. Hejda, J. Bueno, and A. Hurtado, "All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes," *Photon. Res.*, vol. 9, no. 5, B201–B209, 2021. DOI: 10.1364/PRJ.412141. [Online]. Available: http://www.osapublishing.org/prj/abstract.cfm?URI=prj-9-5-B201.

[26] J. Robertson, Y. Zhang, M. Hejda, J. Bueno, S. Xiang, and A. Hurtado, "Image edge detection with a photonic spiking VCSEL-neuron," *Optics Express*, vol. 28, no. 25, pp. 37 526–37 537, Dec. 2020, ISSN: 1094-4087. DOI: 10.1364/OE.408747. [Online]. Available: https://doi.org/10.1364/OE.408747%20https://opg.optica.org/abstract.cfm?URI=oe-28-25-37526.

[27] Y. Zhang *et al.*, "Experimental demonstration of pyramidal neuron-like dynamics dominated by dendritic action potentials based on a VCSEL for all-optical XOR classification task," *Photonics Research*, vol. 9, no. 6, p. 1055, Jun. 2021, ISSN: 2327-9125. DOI: 10.1364/PRJ.422628. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=prj-9-6-1055.

[28] S. Xiang *et al.*, "A review: Photonics devices, architectures, and algorithms for optical neural computing," *Journal of Semiconductors*, vol. 42, no. 2, p. 023 105, Feb. 2021, ISSN: 1674-4926. DOI: 10.1088/1674-4926/42/2/023105. [Online]. Available: https://iopscience.iop.org/article/10.1088/1674-4926/42/2/023105.

[29] J. Robertson *et al.*, "Ultrafast neuromorphic photonic image processing with a VCSEL neuron," *Scientific Reports*, vol. 12, no. 1, p. 4874, Dec. 2022, ISSN: 2045-2322. DOI: 10.1038/s41598-022-08703-1. [Online]. Available: https://www.nature.com/articles/s41598-022-08703-1.

[30] H. Jaeger, "The" echo state" approach to analysing and training recurrent neural networks-with an erratum note'," *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, 2001.

[31] W. Maass, T. Natschläger, and H. Markram, "Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations," *Neural Computation*, vol. 14, no. 11, pp. 2531–2560, Nov. 2002, ISSN: 0899-7667. DOI: 10.1162/089976602760407955. [Online]. Available: https://direct.mit.edu/neco/article/14/11/2531-2560/6650.

[32] T. Hülser, F. Köster, L. Jaurigue, and K. Lüdge, "Role of delay-times in delay-based photonic reservoir computing [Invited]," *Optical Materials Express*, vol. 12, no. 3, p. 1214, Mar. 2022, ISSN: 2159-3930. DOI: 10.1364/OME.451016. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=ome-12-3-1214.

[33] A. Rohm, L. Jaurigue, and K. Ludge, "Reservoir Computing Using Laser Networks," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, no. 1, pp. 1–8, Jan. 2020, ISSN: 1077-260X. DOI: 10.1109/JSTQE.2019.2927578. [Online]. Available: https://ieeexplore.ieee.org/document/8758193/.

[34] Y. Huang, P. Zhou, Y. G. Yang, D. Y. Cai, and N. Q. Li, "Enhanced performance of reservoir computing using multiple self-injection and mutual injection vcsels," *IEEE Journal of Selected topics in Quantum Electronics*, vol. 29, no. 2, pp. 1–9, 2023. DOI: 10.1109/JSTQE.2022.3216628. [Online]. Available: https://ieeexplore.ieee.org/document/9927344.

[35] K. Vandoorne *et al.*, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nature Communications*, vol. 5, no. 1, p. 3541, May 2014, ISSN: 2041-1723. DOI: 10.1038/ncomms4541. [Online]. Available: http://www.nature.com/articles/ncomms4541.

[36] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nature Communications*, vol. 4, no. 1, p. 1364, Jun. 2013, ISSN: 2041-1723. DOI: 10.1038/ncomms2368. [Online]. Available: http://www.nature.com/articles/ncomms2368.

[37] Q. Vinckier *et al.*, "High-performance photonic reservoir computer based on a coherently driven passive cavity," *Optica*, vol. 2, no. 5, p. 438, May 2015, ISSN: 2334-2536. DOI: 10.1364/OPTICA.2.000438. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=optica-2-5-438.

[38] J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Optics Express*, vol. 25, no. 3, p. 2401, Feb. 2017, ISSN: 1094-4087. DOI: 10.1364/OE.25.002401. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=oe-25-3-2401.

[39] A. Argyris, J. Bueno, and I. Fischer, "PAM-4 Transmission at 1550 nm Using Photonic Reservoir Computing Post-Processing," *IEEE Access*, vol. 7, pp. 37 017–37 025, 2019, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2019.2905422. [Online]. Available: https://ieeexplore.ieee.org/document/8669764/.

[40] J. Vatin, D. Rontani, and M. Sciamanna, "Enhanced performance of a reservoir computer using polarization dynamics in VCSELs," *Optics Letters*, vol. 43, no. 18, p. 4497, Sep. 2018, ISSN: 0146-9592. DOI: 10.1364/OL.43.004497. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=ol-43-18-4497.

[41] J. Vatin, D. Rontani, and M. Sciamanna, "Experimental reservoir computing using VCSEL polarization dynamics," *Optics Express*, vol. 27, no. 13, p. 18 579, Jun. 2019, ISSN: 1094-4087. DOI: 10.1364/OE.27.018579. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=oe-27-13-18579.

[42] J. Vatin, D. Rontani, and M. Sciamanna, "Experimental realization of dual task processing with a photonic reservoir computer," *APL Photonics*, vol. 5, no. 8, p. 086 105, 2020. DOI: 10.1063/5.0017574. eprint: https://doi.org/10.1063/5.0017574. [Online]. Available: https://doi.org/10.1063/5.0017574.

[43] J. Bueno, J. Robertson, M. Hejda, and A. Hurtado, "Comprehensive Performance Analysis of a VCSEL-Based Photonic Reservoir Computer," *IEEE Photonics Technology Letters*, vol. 33, no. 16, pp. 920–923, Aug. 2021, ISSN: 1041-1135. DOI: 10.1109/LPT.2021.3075095. [Online]. Available: https://ieeexplore.ieee.org/document/9415868/.

[44] X. Porte, A. Skalli, N. Haghighi, S. Reitzenstein, J. A. Lott, and D. Brunner, "A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser," *Journal of Physics: Photonics*, vol. 3, no. 2, p. 024 017, Apr. 2021, ISSN: 2515-7647. DOI: 10.1088/2515-7647/abf6bd. [Online]. Available: https://iopscience.iop.org/article/10.1088/2515-7647/abf6bd.

[45] A. Skalli, X. Porte, N. Haghighi, S. Reitzenstein, J. A. Lott, and D. Brunner, "Computational metrics and parameters of an injection-locked large area semiconductor laser for neural network computing," *arXiv e-prints*, arXiv:2112.08947, arXiv:2112.08947, Dec. 2021. arXiv: 2112.08947 [cs.ET].

[46] A. Skalli *et al.*, "Photonic neuromorphic computing using vertical cavity semiconductor lasers," *Optical Materials Express*, vol. 12, no. 6, p. 2395, Jun. 2022, ISSN: 2159-3930. DOI: 10.1364/OME.450926. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=ome-12-6-2395.

[47] D. Owen-Newns, J. Robertson, M. Hejda, and A. Hurtado, "GHz Rate Neuromorphic Photonic Spiking Neural Network With a Single Vertical-Cavity Surface-Emitting Laser (VCSEL)," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 29, no. 2: Optical Computing, pp. 1–10, Mar. 2023, ISSN: 1077-260X. DOI: 10.1109/JSTQE.2022.3205716. [Online]. Available: https://ieeexplore.ieee.org/document/9887865/.

[48] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936. DOI: https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

[49] I. Guyon, J. Li, T. Mader, P. A. Pletscher, G. Schneider, and M. Uhr, "Competitive baseline methods set new standards for the nips 2003 feature selection benchmark," *Pattern Recognition Letters*, vol. 28, no. 12, pp. 1438–1444, 2007, ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2007.02.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865507000694.

[50] J. Bueno, J. Robertson, M. Hejda, and A. Hurtado, "Experimental Implementation of a Photonic Neural Network with a 1550nm-VCSEL subject to Optical Injection and Delayed Optical Feedback," in *2020 IEEE Photonics Conference, IPC 2020 - Proceedings*, 2020, ISBN: 9781728158914. DOI: 10.1109/IPC47351.2020.9252399.