Time-series classification with an all-optical recurrent neuron

G. Mourgias-Alexandris(1)(2), N. Passalis(1), G. Dabos(1)(2), A. Totović(1)(2), A. Tefas(1) and N. Pleros(1)(2)

(1) Department of Informatics, Aristotle University of Thessaloniki, 54621 Thessaloniki, Greece
(2) Center for Interdisciplinary Research and Innovation (CIRI-AUTH), Balkan Center, Buildings A & B, Thessaloniki, 10th km Thessaloniki-Thermi Rd, P.O. Box 8318, GR 57001. mourgias@csd.auth.gr

Abstract *We present an all-optical recurrent neuron comprising a novel space rotator and a sigmoid activation unit and demonstrate experimentally time-series classification. Real-time recognition of both RZ and NRZ 3-bit-long bit sequences with 100psec optical pulses is successfully demonstrated, revealing an average accuracy of >91%.*

Introduction

During the last decade neuromorphic photonics has gained a lot of attention since they offer orders of magnitude higher computational speeds compared to their electronic counterparts. The superiority of photonic technology over electronics refers to the deployment of a whole new class of ultra-fast, compact and low-energy optical modules empowering a rich portfolio of key-building blocks to build photonic neurons with unmatched performance, energy and footprint breakthroughs[1–8]. Several architectures of neuromorphic photonic layouts[9–17] have been developed, so far, yet mostly for feed-forward[9], convolutional[16] and spiking[11] neural networks. In stark contrast, Recurrent Neural Networks (RNNs) and their variants Long-Short-term-Memory (LSTM) and Gated Recurrent Unit (GRU), which typically form the cornerstone of classical Deep Learning (DL)-based time-series analysis applications (e.g. speech recognition and financial forecasting, are still attempting their very first deployment steps towards migrating from electronic to photonic hardware implementations [6,18–20]. Towards that direction, only photonic Reservoir Computing (RC)[12–15] based alternatives have been reported and in some cases also employed in real AI tasks[14,21], still requiring, however, additional complex signal post-processing of the outcoming optical data for declaring task completion. However, RC circuits are well-known to comprise a special category of recurrent neurons that is not widely adopted by the AI community and can hardly migrate to more advanced deployments such as LSTM and GRUs[22,23].

Herein, we experimentally demonstrate for the first-time, to the best of our knowledge, a Photonic Recurrent Neuron (PRN) that provides successful time-series classification at speeds up to 10Gb/s, without requiring any post-processing. The proposed PRN comprises a novel space rotator module followed by a neural classifier that is formed by a sigmoid activation element incorporated into a recurrent feedback loop. The space rotator transforms the incoming optical time-series vector in order to avoid the use of negative weights. Successful time-series classification has been experimentally presented for a 3-bit data sequence with 100psec-long optical pulses and for both a RZ modulation scheme at 3.3Gb/s as well as a NRZ format at 10Gb/s, revealing an average accuracy of >91%.

PRN architecture

The layout of the proposed PRN is illustrated in Fig. 1(a). The PRN comprises a space rotator and a recurrent neuron, which in turn includes a weighting bank and the optical sigmoid activation unit reported in[24] incorporated into a recurrent loop. The time-series x_t and the rotation mask m_t are both forwarded as input signals into the space rotator unit, providing at its output a rotated version $\tilde{x}_t$ of the input time-series x_t that equals:

$$\tilde{x}_t = g(x_t) = 1 - x_t \qquad (1)$$

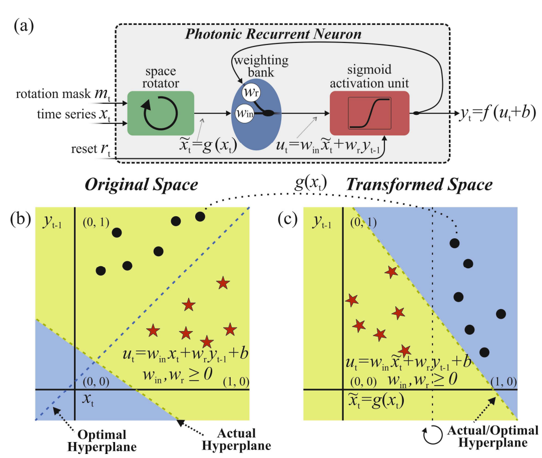

Fig. 1. (a) Layout of the proposed PRN, (b) the original sample space that supports positive weights and (c) the transformed sample space.

with g denoting the transfer function of the space rotator. The rotated time-series $\widetilde{x}_{\mathrm{t}}$ gets then weighted by the input weight value w_{in} and is then multiplexed with the weighted recurrent signal $w_{\mathrm{r}}y_{\mathrm{t-1}}$ that equals zero at t=0, with $y_{\mathrm{t-1}}$ comprising the recurrent signal and w_{r} its weight. The summed output of the weighting bank $u_{\mathrm{t}} = \widetilde{x}_{\mathrm{t}}w_{\mathrm{in}} + y_{\mathrm{t-1}}w_{\mathrm{r}}$ is then injected into the sigmoid activation unit, where a bias signal b is additionally applied so that the output signal of the activation unit equals $y_{\mathrm{t}} = f(u_{\mathrm{t}} + b)$, with f denoting the transfer function of the sigmoid activation unit of[24]. This signal is then split into 2 identical signals with the first copy providing the output of the PRN and the second signal being injected into a feedback loop so as to get weighted by w_{r} and delayed by one-time instance t in order to realize $y_{\mathrm{t-1}}$. The weighted recurrent signal $w_{\mathrm{r}}y_{\mathrm{t-1}}$ is then multiplexed with the next time instance of the weighted time-series $w_{\mathrm{in}}\widetilde{x}_{\mathrm{t}}$ prior re-entering the activation unit. The output of the proposed PRN is governed by the following equation:

$$y_{\mathrm{t}} = f(\widetilde{x}_{\mathrm{t}}w_{\mathrm{in}} + y_{\mathrm{t-1}}w_{\mathrm{r}} + b) \qquad (2)$$

Before launching a new time-series into the PRN, the memory of the system discards any stored instance of the previous time-series by means of a reset signal r_{t}.

The space rotator module is responsible for rotating the time-series input vector so as to reside on the first quadrant of the corresponding coordinate system, where the subsequent optically implemented neural classifier can be successfully applied for classifying the input sequence without using negative weights. Fig. 1(b) illustrates a 2-D space where the decision boundary is implemented by the deployed photonic neuron based only on positive weights. The decision boundary is described by:

$$y_{t-1} = \left(-\frac{w_{in}}{w_r}\right)x_t + \left(-\frac{b}{w_r}\right) \qquad (3)$$

However, the resulted decision boundary denoted as “actual hyperplane” cannot separate the data that are depicted as black circles from the red stars. To overcome this hurdle, we propose the rotation function $g(x_{\mathrm{t}})$ that rotates the sample space by transforming the input data. The new decision boundary is depicted in Fig. 1(c) and separates successfully the red stars from black circles. Moreover, the proposed method requires no changes to the training process.

All-optical 3-bit time-series classification

The experimental validation of the all-optical time-series classification using the proposed photonic neuron has been carried out for three-symbol binary time sequences, where both the input and mask vectors comprise binary digits “0” and “1” with a pattern length of three bits. The PRN was trained to identify and classify three successive ones with a dataset that contains all the possible 3-bit sequences. The PRN was trained using the back propagation through time method[25] employing the squared loss function. The Adam algorithm with a learning rate of η=0.0001 was used for the optimization[26], while the optimization ran for 10,000 iterations.

The experimentally implemented setup of the PRN is depicted in Fig. 2 and its performance was evaluated with every bit being represented by a 100-psec long optical pulse, both when an RZ as well as when an NRZ modulation scheme at 10Gb/s was employed. The Signal Generation Unit was responsible for generating the required signals that have been used for the experimental evaluation, with the sequence of 3-bit-long rotation mask m_{t} signals being imprinted on λ_0 and the sequence of 3-bit-long input vector x_{t} imprinted on λ_2. Moreover, an optical clock signal with 100-psec long optical pulses was generated at λ_1 that was responsible for carrying the result of the rotating function between the mask and the input time vector, while an additional signal carried by λ_3 had the role of the reset signal. The space rotator was realized by means of a Semiconductor Optical Amplifier-Mach-Zehnder Interferometer (SOA-MZI) where the m_{t} and x_{t} optical signals were launched as control signals into the respective SOAs of the two SOA-MZI branches through SOA-MZI ports A and D, respectively. The space rotator output was imprinted on λ_1 that was subsequently launched into the weighting

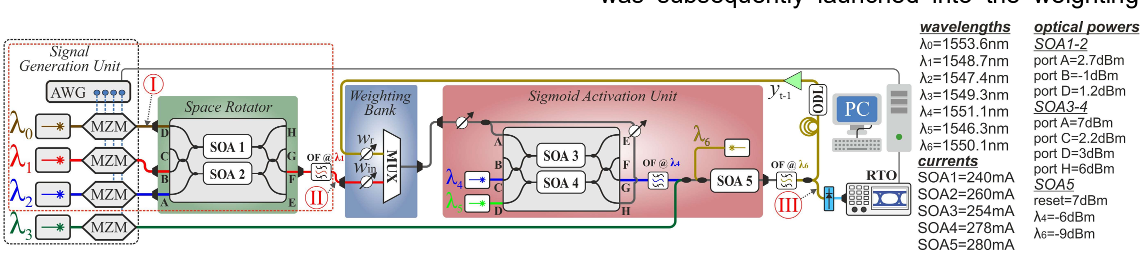

Fig. 2. Experimental setup used to validate the classification of 3-bit RZ & NRZ time-series through the proposed PRN as well as the functionality of the space rotator.

bank module, after being filtered in an Optical band-pass Filter (OF) with a 3dB-bandwidth of 0.8nm. This signal was then weighted by w_{in} in a variable optical attenuator prior being multiplexed with the optical recurrent signal in an optical wavelength multiplexer (MUX). The MUX output enters then the all-optical sigmoid activation function identical to the one demonstrated in[24]. The output of the PRN is imprinted on λ_6 after being filtered by a 0.8nm OF and is split into 2 identical signals, with the first one being injected into a photodiode for being captured and analyzed in a Keysight DSOZ632A RTO with 33 GHz bandwidth and 80GSa/s sampling rate. The second signal constituent was fed into a fiber-based feedback loop that comprises a fiber length of 61m and an optical delay line (ODL), so as to introduce a time delay of $D=T\times(N\times K-1)$ bits, with T denoting the bit period, K the number of bits contained within a signal period, and N being an integer. The delayed signal was forwarded into the weighting bank, effectively forming the y_{t-1} recurrent neuron signal that was subsequently weighted by w_r via a variable optical attenuator prior being multiplexed with the respective weighted x_t signal. Note that the weight values according to the training are equal to w_{in}=10.7dB and w_r=5.5dB.

Real-time time-series classification operation has been realized by using 100-psec long optical pulses within a bit-period of T=300psec for the optical rotation mask m_t, input time-series x_t and input clock signals, are shown in Fig. 5(a)-(f). Figure 5(a) illustrates the bit time-series x_t that contains all the 8 different 3-bit patterns, while the clock and the rotation mask m_t signals have a periodic content of "1110" and "$X_1X_2X_3$0", respectively. When a mask signal of m_t="000" is used, the output of

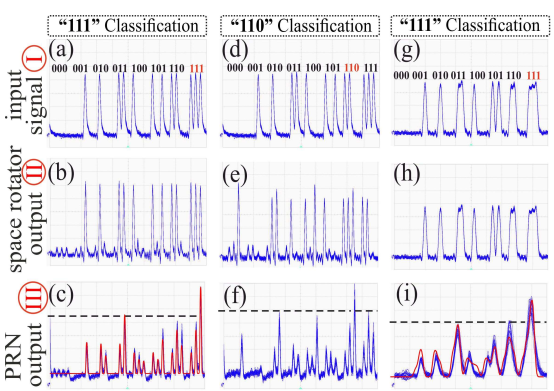

Fig. 3. (a) Time traces from the experimental evaluation of 3-bit classification with space rotation mask: (a)-(c) m_t="000"and (d)-(f) m_t="001". (h)-(i) depicts results for a mask of m_t="000" using NRZ data. y-axis: (2.90mV/div), (a)-(f) x-axis:(980psec/div), (g)-(i) x-axis: (500psec/div).

the space rotator is identical to the input signal x_t of Fig. 5(a), as shown in Fig. 5(b), and the PRN has the role of identifying the incoming bit sequence of "111". The corresponding PRN output is illustrated in Fig. 5(c), clearly revealing that the highest amplitude output pulse emerges at the end of the "111" input time vector, so that a simple thresholding function (shown by the dashed straight line) can validate the successful all-optical recognition of "111". The red line in Fig. 5(c) illustrates the respective output of the software-implemented PRN at a computer, providing almost a perfect match with the corresponding experimentally obtained waveform. Figures 5(d)-(f) depict respective experimental results when the bit pattern "110" has to be identified by the PRN within the sequence of the input time vector, with the rotation mask being equal to m_t="001". In this way, the amplitude pulse across the entire output sequence reveals at the end of "110", as illustrated in Fig. 5(f). Successful 3-bit string recognition has been also performed when using NRZ optical pulses at 10Gb/s. Figures 5(g)-(i) illustrate the obtained results for the NRZ bit sequence classification process when a mask of m_t="000" is employed. Again, the last pulse of the output signal has the highest amplitude verifying the successful classification of NRZ pattens as well. The average 3-bit classification accuracy in case of 10Gb/s RZ optical bit stream was 91.12% when measured over a total number of 80000 bits, with a standard deviation of only 0.78%. Similar results were obtained for 10Gb/s NRZ data, where the average accuracy was found to be 90.13% with a standard deviation of only 0.68%. The operational conditions of the experiment are summarized in Fig. 2.

Conclusion

This work presents all-optical time-series classification using a PRN that requires just a simple thresholding function at its output, negating the need for off-line signal post-processing. Successful 3-bit optical time-series recognition at 3.3 and 10Gb/s for RZ and NRZ formats with an average accuracy of 91.12% and 90.13%, respectively, has been experimentally demonstrated. The proposed layout lays the foreground towards more sophisticated photonic RNNs like LSTMs and GRUs at much higher operational speeds compared to their electronic counterparts.

Acknowledgements

The work was in part funded by the EU-projects PlasmoniAC (871391) and NEBULA (871658).

References

[1] A. N. Tait, T. F. De Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Neuromorphic photonic networks using silicon photonic weight banks," Sci. Rep. **7**, 1–10 (2017).

[2] A. N. Tait, A. X. Wu, T. Ferreira de Lima, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Two-pole microring weight banks," Opt. Lett. **43**, 2276 (2018).

[3] B. Shi, N. Calabretta, and R. Stabile, "Deep Neural Network Through an InP SOA-Based Photonic Integrated Cross-Connect," IEEE J. Sel. Top. Quantum Electron. **26**, 1–11 (2020).

[4] G. Mourgias-Alexandris, A. Tsakyridis, N. Passalis, A. Tefas, K. Vyrsokinos, and N. Pleros, "An all-optical neuron with sigmoid activation function," Opt. Express **27**, 9620 (2019).

[5] M. Miscuglio, A. Mehrabian, Z. Hu, S. I. Azzam, J. George, A. V. Kildishev, M. Pelton, and V. J. Sorger, "All-optical nonlinear activation function for photonic neural networks [Invited]," Opt. Mater. Express **8**, 3851 (2018).

[6] A. N. Tait, T. Ferreira De Lima, M. A. Nahmias, H. B. Miller, H. T. Peng, B. J. Shastri, and P. R. Prucnal, "Silicon Photonic Modulator Neuron," Phys. Rev. Appl. **11**, 1–16 (2019).

[7] J. George, R. Amin, A. Mehrabian, J. Khurgin, T. El-Ghazawi, P. R. Prucnal, and V. J. Sorger, "Electrooptic Nonlinear Activation Functions for Vector Matrix Multiplications in Optical Neural Networks," in *Advanced Photonics 2018 (BGPP, IPR, NP, NOMA, Sensors, Networks, SPPCom, SOF)* (OSA, 2018), p. SpW4G.3.

[8] A. Totovic, G. Dabos, N. Passalis, A. Tefas, and N. Pleros, "Femtojoule per MAC Neuromorphic Photonics : An Energy and Technology Roadmap," J. Sel. Toplics Quantum Electron. **26**, 1-15 (2020).

[9] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," Nat. Photonics **11**, 441–446 (2017).

[10] G. Mourgias-Alexandris, A. Totovic, A. Tsakyridis, N. Passalis, K. Vyrsokinos, A. Tefas, and N. Pleros, "Neuromorphic Photonics With Coherent Linear Neurons Using Dual-IQ Modulation Cells," J. Light. Technol. **38**, 811–819 (2020).

[11] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," Nature **569**, 208–214 (2019).

[12] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," Nat. Commun. **5**, 1–6 (2014).

[13] M. Freiberger, S. Sackesyn, C. Ma, A. Katumba, P. Bienstman, and J. Dambre, "Improving Time Series Recognition and Prediction With Networks and Ensembles of Passive Photonic Reservoirs," IEEE J. Sel. Top. Quantum Electron. **26**, 1–11 (2020).

[14] K. Vandoorne, J. Dambre, D. Verstraeten, B. Schrauwen, and P. Bienstman, "Parallel reservoir computing using optical amplifiers," IEEE Trans. Neural Networks **22**, 1469–1481 (2011).

[15] F. Laporte, A. Katumba, J. Dambre, and P. Bienstman, "Numerical demonstration of neuromorphic computing with photonic crystal cavities," Opt. Express **26**, 7955 (2018).

[16] H. Bagherian, S. Skirlo, Y. Shen, H. Meng, V. Ceperic, and M. Soljacic, "On-Chip Optical Convolutional Neural Networks," 1–18 (2018).

[17] 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," Adv. Opt. Photonics **8**, 228 (2016).

[18] H. Peng, T. F. de Lima, M. A. Nahmias, A. N. Tait, B. J. Shastri, and P. R. Prucnal, "Autaptic Circuits of Integrated Laser Neurons," in *Conference on Lasers and Electro-Optics* (OSA, 2019), Vol. 1, p. SM3N.3.

[19] G. Mourgias-Alexandris, G. Dabos, N. Passalis, A. Totovic, A. Tefas, and N. Pleros, "All-Optical WDM Recurrent Neural Networks With Gating," IEEE J. Sel. Top. Quantum Electron. **26**, 1–7 (2020).

[20] G. Mourgias-Alexandris, G. Dabos, N. Passalis, A. Tefas, A. Totovic, and N. Pleros, "All-optical recurrent neural network with sigmoid activation function," in *Optical Fiber Communication Conference (OFC) 2020* (OSA, 2020), p. W3A.5.

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

[22] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput. **9**, 1735–1780 (1997).

[23] G. Zhou, J. Wu, C.-L. Zhang, and Z.-H. Zhou, "Minimal Gated Unit for Recurrent Neural Networks," (2016).

[24] G. Mourgias-Alexandris, A. Tsakyridis, N. Passalis, A. Tefas, K. Vyrsokinos, and N. Pleros, "An all-optical neuron with sigmoid activation function," Opt. Express **27**, 9620 (2019).

[25] P. J. Werbos, "Backpropagation Through Time: What It Does and How to Do It," Proc. IEEE **78**, 1550–1560 (1990).

[26] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," (2014)