Cascading photonic reservoirs with deep neural networks increases computational performance

Ian Bauwens[a,b], Guy Van der Sande[a], Peter Bienstman[b], and Guy Verschaffelt[a]

[a]Applied Physics Research Group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

[b]Photonics Research Group, Department of Information Technology, Ghent University-IMEC, Technologiepark Zwijnaarde 126, 9052 Ghent, Belgium

ABSTRACT

Deep neural networks (DNNs) have been successfully applied to solve complex problems, such as pattern recognition when analyzing big data. To achieve a good computational performance, these networks are often designed such that they contain a large number of trainable parameters. However, by doing so, DNNs are often very energy-intensive and time-consuming to train. In this work, we propose to use a photonic reservoir to preprocess the input data instead of directly injecting it into the DNN. A photonic reservoir consists of a network of many randomly connected nodes which do not need to be trained. It forms an additional layer to the deep neural network and can transform the input data into a state in a higher dimensional state-space. This allows us to reduce the size of the DNN, and the amount of training required for the DNN. We test this assumption using numerical simulations that show that such a photonic reservoir as preprocessor results in an improved performance, shown by a lower test error, for a deep neural network, when tested on the one-step ahead prediction task of the Santa Fe time-series. The performance of the stand-alone DNN is poor on this task, resulting in a high test error. As we also discuss in detail in [Bauwens et al, Frontiers in Physics 10, 1051941 (2022)], we conclude that photonic reservoirs are well-suited as physical preprocessors to deep neural networks for tackling time-dependent tasks due to their fast computation times and low-energy consumption.

Keywords: Preprocessor, deep neural network, delay-based reservoir computing, machine learning, semiconductor laser, feedback

1. INTRODUCTION

Deep neural networks (DNNs) have proven invaluable in the field of machine learning and artificial neural networks due to their versatility across a wide array of tasks such as natural language processing, drug discovery, medical image detection, and big data analysis. Their adeptness at recognizing inherent patterns and features enables them to attain remarkably high accuracy in these tasks, which explains their widespread adoption. The functioning of artificial neural networks involves training a network by fitting a large dataset of input samples to a model through the optimization of internal network weights. However, as the complexity and depth of these networks increase, often necessary for achieving high accuracies, the number of parameters and weights can grow substantially, making them computationally challenging to train. This increased complexity necessitates significant memory, time, and energy resources for effective training.

To address this issue, we propose the integration of a photonics-based preprocessor, placed in front of of the DNN architecture.[1] This integration allows for the transformation of input data into a higher-dimensional state-space, capitalizing on photonics' advantages, including low-energy consumption, inherent parallelism, and fast computation times.[2,3] The photonic preprocessor explored in this study is based on reservoir computing (RC), which are recurrent neural networks (RNNs) and which consist of three different layers: an input layer, a reservoir, and an output layer. An illustration of an RC system is shown in Fig. 1 The input layer serves to introduce input data into the system, while the output layer is used for predictions, typically utilizing a linear

Further author information: (Send correspondence to I.B.)

: E-mail: ian.bauwens@vub.be

read-out with linear weights. The reservoir consists of many randomly interconnected nonlinear nodes. Unlike DNNs, the internal weights of the reservoir are fixed and remain constant, with training solely occurring in the output layer, leading to significantly reduced training times compared to DNNs.[4] RC systems have already demonstrated success in various prediction tasks, including speech recognition,[5,6] time-series predictions,[7–9] and nonlinear channel equalization tasks.[10]

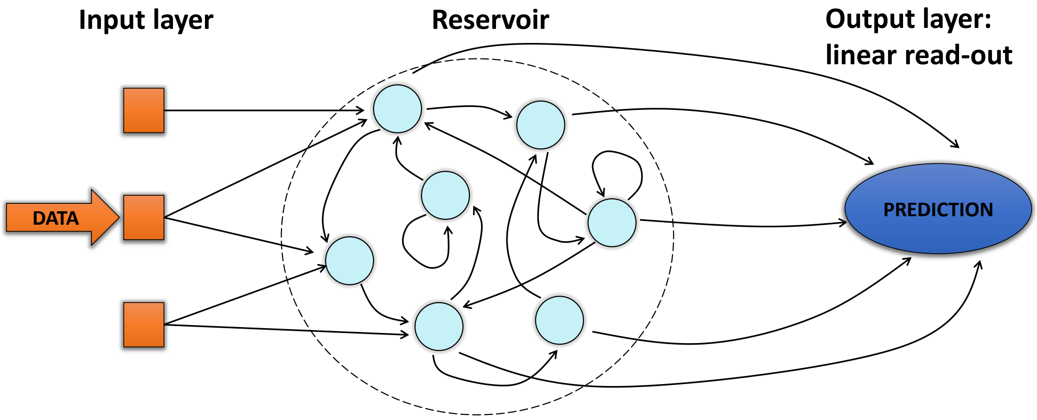

Figure 1. Illustration of a general RC system, with input layer, reservoir and output layer.

In this work, we study a photonic delay-based reservoir which uses a semiconductor laser as nonlinear node.[11–13] We combine this photonic reservoir with two different output layers: firstly, when paired with a linear read-out layer, used in conventional reservoir computing, and secondly, when integrated with a DNN. In the combined reservoir and DNN setup, we pass the output of the photonic reservoir as input to the DNN, effectively utilizing the reservoir as a preprocessor for the DNN.

2. NUMERICAL IMPLEMENTATIONS

2.1 Reservoir computing

We use a photonic delay-based RC system employing a semiconductor laser with delayed feedback, and which uses a time-multiplexing method for data injection.[14] The intensity is measured at virtual nodes along the delay line to serve as the reservoir output, and the state of these nodes is recorded in the state matrix $\mathbf{A}$. In our study, we use 200 nodes for the reservoir. The training of linear weights w in this RC approach is accomplished by computing $w = \mathbf{A}^{\dagger} y_{train}$, where † denotes the Moore-Penrose pseudoinverse. Input data u_k is optically injected into the reservoir through a Mach-Zehnder modulator.[15] Due to the necessary time-multiplexing technique, a mask $m(t)$ must be implemented to ensure proper coupling between nodes. An illustration of this RC system is shown in Fig. 2. The dynamics of the single-mode laser are governed using the rate equations from Ref.[16]

2.2 Deep neural network

In our study, the DNN takes a vector of 200 features as input data, corresponding to the number of reservoir nodes, and outputs a single value prediction. This DNN architecture comprises three fully-connected layers, with a decreasing number of nodes as we progress through the network. The initial hidden layer consists of a fully-connected layer with 100 nodes, each employing a ReLU activation function. The subsequent hidden layer also utilizes a ReLU activation function and contains 50 nodes. The resulting signal is then fed into the final layer, which comprises a single node responsible for making predictions, denoted as $\hat{y}_k$. We train the DNN using

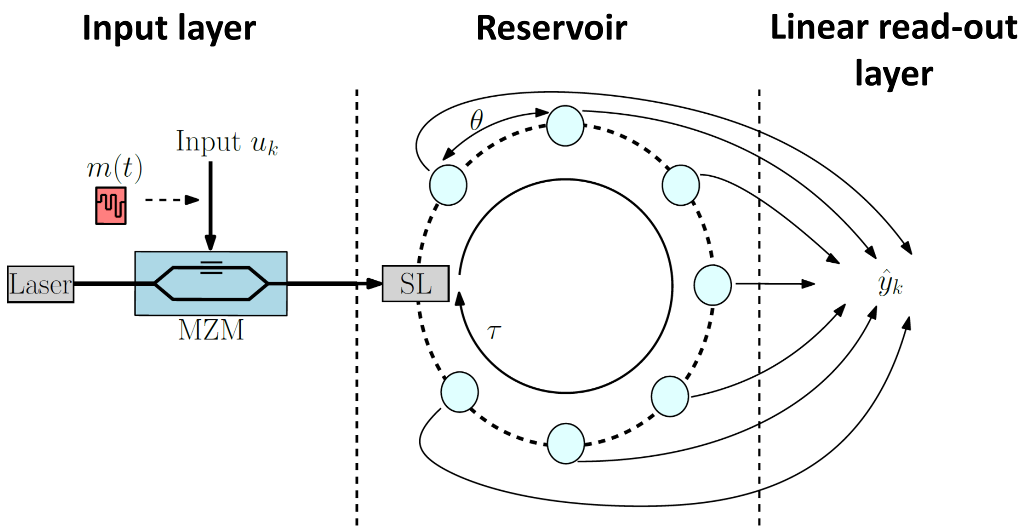

Figure 2. Delay-based RC system using a semiconductor laser (SL), with input data u_k, mask $m(t)$, node separation θ and delay time τ. The light blue circles represent the virtual nodes and the output layer is defined to make predictions $\hat{y}_k$.

the Adam optimizer, setting a learning rate of 10^{-3} for 10^4 epochs. To prevent overfitting, we implement early stopping based on a validation set distinct from the training and test sets. This configuration of a photonic preprocessor network combined with a DNN is illustrated in Fig. 3.

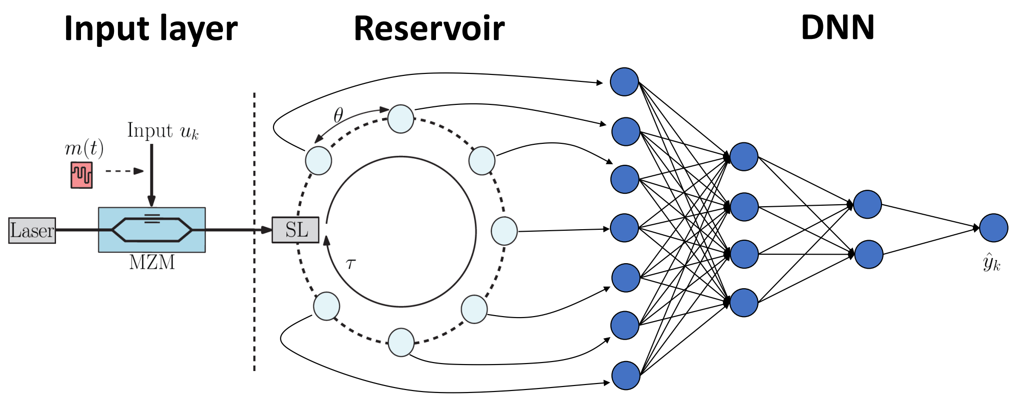

Figure 3. Illustration of the photonic reservoir combined with the DNN.

3. NUMERICAL RESULTS

We assess the performance of various networks through a task involving one-step ahead time-series prediction on the Santa Fe dataset.[17] This dataset consists of more than 9000 data points sampled from a far-IR laser operating in a chaotic regime. For the RC system, we utilize 3000 data samples from the discrete Santa Fe dataset, labeled as u_k^{train}, for training. Additionally, we use 1000 distinct data samples, denoted as u_k^{test}, for testing purposes. Our assessment relies on the normalized mean squared error (NMSE) as a performance metric. In Fig. 4, we illustrate the performance of both the RC system and the photonic reservoir combined with the DNN as a preprocessor, alongside the corresponding relative performance enhancements. This analysis is conducted across a 2D parameter scan of the feedback rate to the semiconductor laser and the normalized excess pump current, which constitutes the reservoir.

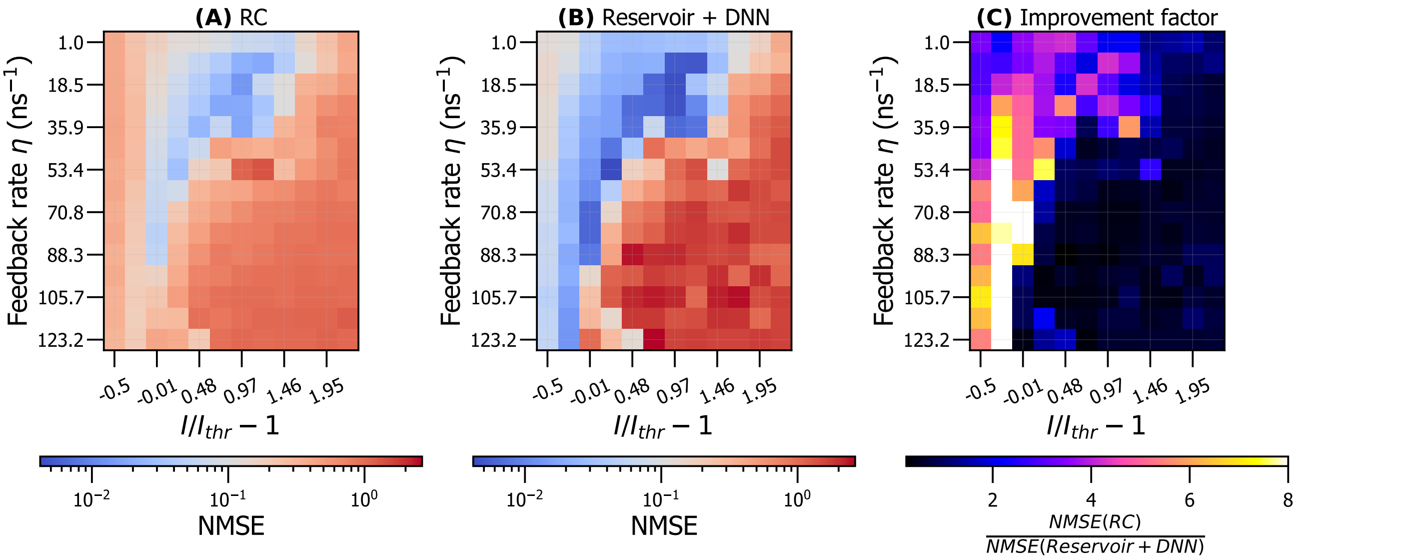

Figure 4. Performance after a grid search: NMSE results after scanning the feedback rate and pump current, at constant injection rate, for conventional RC (a), for the photonic reservoir coupled with the DNN (b) and the improvement factor of the reservoir combined with the DNN compared to the RC system (c).

We observe that coupling the photonic reservoir with the DNN yields enhanced performance compared to the conventional RC system with a linear read-out. Moreover, this combined approach demonstrates a broad parameter range wherein it achieves an improved performance, suggesting it does not require careful fine-tuning to achieve good results.

4. CONCLUSION

We have numerically investigated the impact of integrating a photonic reservoir with a deep neural network. Using such a photonic reservoir leads to enhanced performance and offers the benefit of serving as an energy-efficient preprocessor to DNNs. Additionally, the parameters of this photonic reservoir do not require careful adjustment to achieve improved performance.

REFERENCES

[1] Bauwens, I., Van der Sande, G., Bienstman, P., and Verschaffelt, G., "Using photonic reservoirs as preprocessors for deep neural networks," *Frontiers in Physics* **10**, 1051941 (2022).

[2] Van der Sande, G., Brunner, D., and Soriano, M. C., "Advances in photonic reservoir computing," *Nanophotonics* **6**(3), 561–576 (2017).

[3] De Lima, T. F., Shastri, B. J., Tait, A. N., Nahmias, M. A., and Prucnal, P. R., "Progress in neuromorphic photonics," *Nanophotonics* **6**(3), 577–599 (2017).

[4] Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., and Hirose, A., "Recent advances in physical reservoir computing: A review," *Neural Networks* **115**, 100–123 (2019).

[5] Verstraeten, D., Schrauwen, B., Stroobandt, D., and Van Campenhout, J., "Isolated word recognition with the liquid state machine: a case study," *Information Processing Letters* **95**(6), 521–528 (2005).

[6] Verstraeten, D., Schrauwen, B., and Stroobandt, D., "Reservoir-based techniques for speech recognition," in [*The 2006 IEEE International Joint Conference on Neural Network Proceedings*], 1050–1053, IEEE (2006).

[7] Jaeger, H. and Haas, H., "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science* **304**(5667), 78–80 (2004).

[8] Skibinsky-Gitlin, E. S., Alomar, M. L., Isern, E., Roca, M., Canals, V., and Rossello, J. L., "Reservoir computing hardware for time series forecasting," in [*2018 28th International Symposium on Power and Timing Modeling, Optimization and Simulation (PATMOS)*], 133–139, IEEE (2018).

[9] Canaday, D., Griffith, A., and Gauthier, D. J., "Rapid time series prediction with a hardware-based reservoir computer," *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28**(12), 123119 (2018).

[10] Paquot, Y., Duport, F., Smerieri, A., Dambre, J., Schrauwen, B., Haelterman, M., and Massar, S., "Optoelectronic reservoir computing," *Scientific reports* **2**(1), 1–6 (2012).

[11] Appeltant, L., Soriano, M. C., Van der Sande, G., Danckaert, J., Massar, S., Dambre, J., Schrauwen, B., Mirasso, C. R., and Fischer, I., "Information processing using a single dynamical node as complex system," *Nature communications* **2**(1), 1–6 (2011).

[12] Harkhoe, K., Verschaffelt, G., Katumba, A., Bienstman, P., and Van der Sande, G., "Demonstrating delay-based reservoir computing using a compact photonic integrated chip," *Optics express* **28**(3), 3086–3096 (2020).

[13] Harkhoe, K. and Van der Sande, G., "Delay-based reservoir computing using multimode semiconductor lasers: Exploiting the rich carrier dynamics," *IEEE Journal of Selected Topics in Quantum Electronics* **25**(6), 1–9 (2019).

[14] Harkhoe, K., Verschaffelt, G., Katumba, A., Bienstman, P., and Van der Sande, G., "Demonstrating delay-based reservoir computing using a compact photonic integrated chip," *Optics express* **28**(3), 3086–3096 (2020).

[15] Bauwens, I., Harkhoe, K., Bienstman, P., Verschaffelt, G., and Van der Sande, G., "Influence of the input signal's phase modulation on the performance of optical delay-based reservoir computing using semiconductor lasers," *Optics Express* **30**(8), 13434–13446 (2022).

[16] Lenstra, D. and Yousefi, M., "Rate-equation model for multi-mode semiconductor lasers with spatial hole burning," *Optics express* **22**(7), 8143–8149 (2014).

[17] Weigend, A. S. and Gershenfeld, N. A., "Results of the time series prediction competition at the Santa Fe Institute," in [*IEEE international conference on neural networks*], 1786–1793, IEEE (1993).