Fibre Optic Nonlinear Technologies [FONTE - A European Industrial Doctorate [GA766115]

Document Details

Title	Deliverable 3.1 Survey of machine learning algorithms for optical performance monitoring
Deliverable number	D3.1
Deliverable Type	Report (public)
Work Package	WP3 - Machine learning techniques for fibre-optic channel
Deliverable due date	31/01/2019
Actual date of submission	21/01/2019
Lead beneficiary	DTU
Version number	V1.0
Status	FINAL

Dissemination level

PU	Public	X
CO	Confidential, only for members of the consortium (including Commission Services	

Project Details

Grant Agreement	766115
Project Acronym	FONTE
Project Title	Fibre Optic Nonlinear TEchnologies
Call Identifier	H2020-MSCA-ITN-2017
Project Website	fonte.astonphotonics.uk
Start of the Project	1 June 2018
Project Duration	48 months

Consortium

EC Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 766115

Executive Summary

D3.1 Survey of machine learning algorithms for optical performance monitoring

An incredible 4.8 zettabytes of total yearly IP traffic data has been predicted by Cisco for 2022 [1]. New applications of internet of things, the increasing number of smartphones, Ultra High Definition (4K) streaming videos and many other applications are driving the necessity for higher transmission rates to support this increasing need for data. Since fiber-optic communications are the backbone of the telecommunications systems, new solutions to cope with this ever-increasing need for transmission speed are critical.

Nowadays, nonlinear effects in optical fibers are one of the major limiting factors for optical communications. Different technologies have attempted to address these impairments for optical transmission. Nonlinearity mitigation through digital signal processing (DSP)[2], optical phase conjugation (OPC) [3] and nonlinear frequency division multiplexing (NFDM) [4] are the most known topics in this area. Although great results show the increase of the transmission data rate and reach, the implementation costs of any of them are still prohibitive. It is also worth to mention that all the mentioned techniques require full knowledge of the fiber transmission parameters to work properly.

Optical performance monitoring (OPM) techniques estimates the parameter of the optical fiber channel, which is required at multiples point along the link. The multiple uses of this technology imply the necessity of simpler and cheaper solutions [5].

Machine learning (ML) techniques may help solve the fiber nonlinearities for optical transmission and OPM, potentially reducing implementation costs. ML tools has a broad area of application and are very well known for being extremely effective for classification problems, typically for image classification and speech detection. Nonetheless, ML implementations with central processing units (CPU) are suboptimal in terms of speed and power efficiency [6]. Considering the optical communication field, it also implies the necessity to convert the signal from optical domain to electrical/digital domain. In other words, an expensive solution. Alternatively, several researches using hybrid optical-electronic systems with reservoir computing (RC) [7, 8] and full optical neural networks (ONN) with programmable nanophotonic processor (PNP) have been demonstrated [6, 9]. Here, we review these algorithms and point out potential advantages and disadvantages.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

AiPT	Aston Institute Of Photonic Technologies
ANN	Artificial Neural Network
CD	Chromatic Dispersion
CPU	Central Processing Units
DSP	Digital Signal Processing
EC	European Commission
EID	European Industrial Doctorates
ESR	Early Stage Researcher
FONTE	Fibre Optic Nonlinear Technologies
ML	Machine Learning
MZI	Mach-Zehnder interferometer
NFDM	Nonlinear Frequency Division Multiplexing
ONN	Optical Neural Network
OPC	Optical Phase Conjugation
OPM	Optical Performance Monitoring
PNP	Programmable Nanophotonic Processor
RC	Reservoir Computing
RNN	Recurrent Neural Network
SVD	Singular value decomposition

1 Reservoir Computing

Optical fibre communications impairments such as chromatic dispersion (CD) and nonlinearities induced by Kerr effect are time-dependent. Whereas this is a challenging task to overcome in a common ML architecture, reservoir computing (RC) allows mitigating time-dependent impairments with the aid of recurrent connections.

Recurrent neural networks (RNNs) have been shown to be universal approximators, considering a finite time trajectory of a given n-dimensional dynamical systems with appropriate initial conditions [10]. In other words, they may mitigate nonlinear optical impairments. Though highly attractive, RNN are very difficult to train. In order to simplify the training process, Jaeger [7] and Maass [8] developed, independently and at the same time, the idea of using a random distribution for the input and hidden layer of the RNN, leaving it untrained.

The techniques that consists on leaving the recurrent network layer untrained received the name of RC, coined by Verstraenten et al. [11]. Figure 1 shows a schematic of a general RC topology. The idea is applying a nonlinear transformation which will map the input space to a higher-dimensional space, resulting in a dimensionality expansion that might be linear separable.

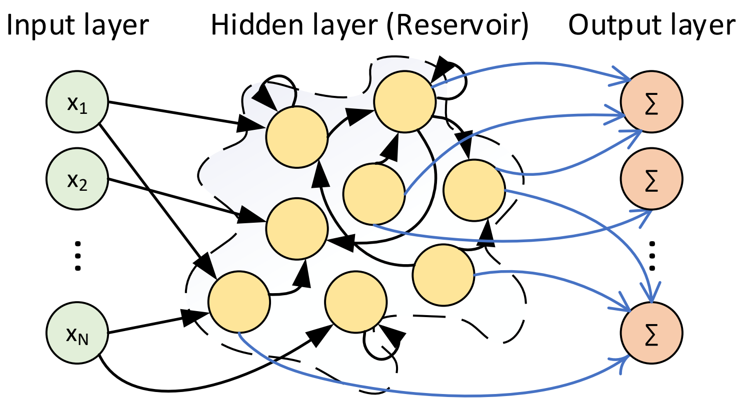

Figure 1 – Schematic of a general RC. The blue arrows are the only weights trained.

Considering the linear case, which can be obtained if there are enough independent projections from the reservoir, the solution of the optimal weights is the Moore-Penrose pseudoinverse $A^{\dagger}$ of the matrix state A (obtained from the reservoir) [12]:

$$A^{\dagger} = (A^T A)^{-1} A^T \quad (1)$$

$$W_{out} = \left(\hat{Y} A^{\dagger}\right)^T \quad (2)$$

Where the desired output is represented by $\hat{Y}$ and the weights of the output layer by W_{out}.

To be able to apply RC in order to solve a time-dependent problem, one of the principal requirements is that the weights of the hidden layer need to be scaled to acquire a dynamical regime. In addition, considering that, there is no external input, the system should return to a quiescent state. If such requirements are satisfied, the reservoir experiences fading memory. In other words, an old information has less influence over time [12].

RC has the beneficious of a RNN without the challenge task of training all the weights of the reservoir. The feature of leaving the initialization of the reservoir untrained are an enormous benefit for a hardware implementation.

2 OPTICAL NEURAL NETWORK

Artificial Neural Network (ANN) is traditionally organized in layers with units called neurons. In each layer, a linear transformation (or a matrix multiplication) is applied followed by a nonlinear activation function. Figure 2a shows a general ANN topology. It is well known that ANN with one hidden layer and linear output units can approximate arbitrary well any continuous function with sufficient number of neurons in the hidden layer.

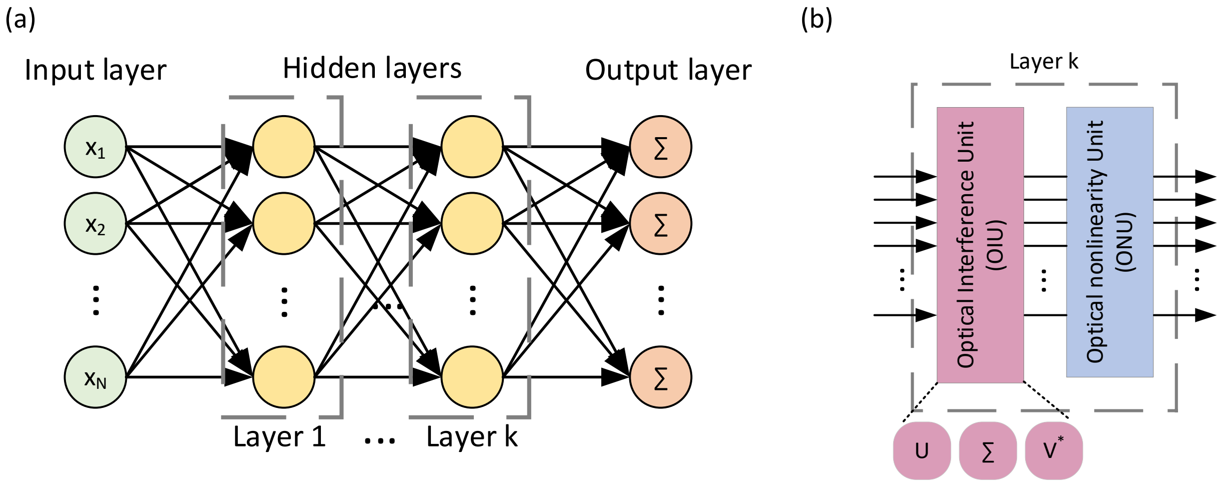

Figure 2 – (a) General schematic of an artificial neural network. (b) Building blocks for an optical neural network implementation.

In order to implement any ANN with optical components, two different components are needed. One to implement the linear transformation, called optical interference unit (OIU), and other that applies a nonlinear function, called optical nonlinear unit (ONU) [6]. Figure 2b shows a schematic of the building blocks for an optical neural network (ONN).

An ONU can be implemented using common optical nonlinearities such as saturable absorption [13]. An implementation of OIU needs to execute a matrix multiplication in the optical domain. The work of Michael Reck et al. [9] experimentally demonstrates an algorithm that implements any $N\ x\ N$ unitary matrix by using optical beam splitters with variable reflectivity. It is also possible to replace it by a Mach-Zehnder interferometer (MZI), which can be implemented using a programmable nanophotonic processor (PNP) – a silicon photonic integrated circuit fabricated in the OPSIS foundry [6]. Bringing these results together with the theory of singular-value decomposition (SVD) it is possible to build a cascade of MZI to implement any real or complex matrix multiplication operator.

The SVD factorizes any rectangular matrix A as [14]:

$$A = U\Sigma V^* \tag{3}$$

Where U and V are unitary matrices, Σ is a rectangular diagonal matrix with non-negative numbers and V^* is the conjugate transpose of V. U and V matrices can be optically implemented by using the algorithm demonstrated by Michael Reck et al. and Σ can be implemented using optical attenuators [6].

By implementing an ONN instead of the classical ANN with CPUs, speed and power consumption can be greatly improved. The linear transformations can be done at the speed of light and detected at rates of 100 GHz (photodetection rate) with minimal power consumption [6].

4 REFERENCES

[1] "White paper: Cisco Visual Networking Index: Forecast and Trends, 2017–2022," 2018.

[2] E. Ip, "Nonlinear Compensation Using Backpropagation for Polarization-Multiplexed Transmission," *J. Light. Technol.*, vol. 28, no. 6, pp. 939–951, 2010.

[3] A. D. Ellis, M. E. McCarthy, M. A. Z. Al Khateeb, M. Sorokina, and N. J. Doran, "Performance limits in optical communications due to fiber nonlinearity," *Adv. Opt. Photonics*, vol. 9, no. 3, p. 429, 2017.

[4] S. K. Turitsyn *et al.*, "Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives," *Optica*, vol. 4, no. 3, p. 307, 2017.

[5] F. Musumeci *et al.*, "A Survey on Application of Machine Learning Techniques in Optical Networks," pp. 1–21, 2018.

[6] Y. Shen, N. C. Harris, D. Englund, and M. Soljacic, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics*, pp. 1–2, 2018.

[7] H. Jaeger, "The 'echo state' approach to analysing and training recurrent neural networks - with an Erratum note," *GMD - Ger. Natl. Res. Cent. Inf. Technol.*, no. 148, pp. 1–47, 2001.

[8] W. Maass, T. Natschlager, and H. Markram, "Real-time computing without stable states: a new framework for neural computation based on perturbations," *Neural Comput.*, vol. 14, no. 11, pp. 2531–2560, 2002.

[9] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, "Experimental Realization of Any Discrete Unitary Operator," *Phys. Rev. Lett.*, vol. 73, no. 1, pp. 58–61, 1994.

[10] K. Funahashi and Y. Nakamura, "Approximation of Dynamical Systems by Continuous Time Recurrent Neural Networks," *Neural Networks*, vol. 6, pp. 801–806, 1993.

[11] B. Schrauwen, D. Verstraeten, and J. Van Campenhout, "An overview of reservoir computing: theory, applications and implementations," *Proc. 15th Eur. Symp. Artif. Neural Networks*, no. April, pp. 471–82, 2007.

[12] G. Van Der Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," *De Gruyter*, vol. 6, no. 3, pp. 561–576, 2017.

[13] Y. Shen, S. Skirlo, N. C. Harris, D. Englund, and M. Soljaˇ, "On-Chip Optical Neuromorphic Computing," pp. 3–4, 2016.

[14] G. Golub and W. Kahan, "Calculating the Singular Values and Pseudo-inverse of a Matrix," vol. 2, no. 2, 1964.