Review-Reservoir Computing Trend on Software and Hardware Implementation

Yongbo Liao[1]

[1] University of Electronic Science and Technology of China

Received: 8 December 2016 Accepted: 3 January 2017 Published: 15 January 2017

Abstract

Since Reservoir Computing proposed, it has progressed in two directions, software and hardware implementation, both sharing the same goal of better performance. While applying on the former, the chosen task is increasingly complex and practical, even blending noise to close to physical situation. Meanwhile, the latter, evaluated by benchmark tasks, is proposed as a compensation of software implementation, which will be utilized for complex and practical tasks in the future when it matures. Here will give a brief introduction of conception, methodology, benchmark tasks, developments and some applications of RC.

Index terms— reservoir computing, software implementation, hardware implementation.

1 Introduction

eservoir Computing(RC) has been a preferred research topic when the conception raised for training Recurrent Neural Network (RNN)in the first time. The approach that reinvigorates RNN is a collection of Echo State Network (ESN) [1] and Liquid State Machine (LSM) [2], the former tending to engineering applications while the latter to neurophysiology. RC retains input weights and reservoir weights initialized and only supervised readout weights are trained to obtain excellent performance in many tasks, provided RC possesses certain generic properties. Notably, tasks based on RC model should be time-dependent.

RC is renowned for less time consumption and resource occupancy. Therefore, many researchers attempt to improve existing research results in their arena by applying RC. After survey comprehensively, the research of RC falls into two general categories: software and hardware implementation. In software implementation, the mainstream is exploring a practicability of applying RC in a new arena or an improvement of existing results by RC, while in hardware implementation, the research is more challenge, and so far the aim is building a RC structure on hardware platform and performing simply tasks successfully. With the dedication of researchers, delay-based RC [4] and mask conception [5] emerged and improved hardware implementation.

The emphasis of RC research switches from software to hardware implementation, therefore, this paper will focus on hardware implementation, and will also introduce emerging works based on software implementation as well as some classical benchmark tasks.

2 II.

3 RC Methodology

Before introduce implementation, the basic knowledge of RC should be mentioned. In consideration of that every works related to RC involving basic RC idea, here will briefly and concisely introduce basic knowledge of RC and list out key equations.

The first key equation is reservoir state update equation, which is slightly alterable due to researcher's thoughts and the actual task. The classic reservoir state update equation [1] is as simply as:in target (n 1) f((n 1) (n) (n)) back + = + + + x W u Wx W y**(1)**

For the sake of a better dynamic behavior, a complex reservoir state update equation with a leaky integration [1,6] is proposed as:in target (n 1) (1 a) (n) f((n 1) (n) (n)) back + = ? + + + + x x W u Wx W y**(2)**

where n is discrete time, (n 1) u N + ? u R is the input signals, x N (n) ? x R isN u N in × ? W R , x N x N × ? W R and x N y N back × ? W R

are the input weight matrix, reservoir weight matrix and the output feedback weight matrix respectively, indicating the connections among nodes. And parameter a in equation (2) is a leaky integration constant confined in the range of 0 to 1.

The second key equation is readout output equation [1]:out out (n 1) f (((n 1), (n 1), (n))) + = + + y W u x y**(3)**

where(n) y N ? y R is the readout output signals, f out (?)is a linear continuous function, typically identity function, u x (N N) out y N × + ? W R

is output weight matrix, ((n 1), (n 1), (n))+ + u x y

is the concatenation of the input signals, reservoir state and previous readout output vectors.

The third key equation is training output weight matrix [1], which calculated by equation (1) via ridge regression as (there assumes y equals to y target , i.e. error is zero) [1]:out T 2 1 arg ( ) T t et ? ? = + W Y X XX I (4)

where Y target is the collection of y target(n+1), X is the collection of (u(n+1),x(n+1)), I is identity matrix and ? is regularization parameter which is applied for the optimization of RC structure. (?) -1 is matrix inversion. The final equation is a normalized root-meansquare error (NRMSE) [7], which should be minimized during the training, also as a typical standard measures results:2 target target 2 arg target || (n) (n) || (,) || (n) (n) || t et E < ? > = < ? < > > y y y y y y**(5)**

where ||?|| stands for the Euclidean distance, and <?> for mean function.

III.

4 Preprocessing

Before feeding data into RC system, raw data should be preprocessed to meet some index for a higher performance expectation. To our knowledge, the present existing preprocessing approaches are normalization, masking and Fourier transformation of images.

Normalization preprocessing. Normalization is an approach preferred in preprocessing, benefitted from its simplicity. There are several advantages in adopting normalization preprocessing, which can be exemplified by conveniently processing data, accelerating network training, improving network stability and making the average of input data close to zero or possessing a smaller mean square error. The basic concept of this approach is limiting the amplitude of the input data to the range of 0 to 1 or -1 to 1. After training and obtaining the normalization output data, anti-normalization should be applied on them to obtain genuine output data.

Masking preprocessing. There are two masking approaches including digital binary mask [5,8] and analog chaos mask [5], both are periodic mask and emerge for delay-based RC based on semiconductor laser subjected to optical feedback and injection. Masking preprocessing can be expressed as [5]:(t) mask(t) u( ) M n ? = × ×**(7)**

where M(t) stands for masked signal, mask(t) for mask signal with a periodicity T, and ? for scaling factor.

In [8], Schneider explained how to conduct the preprocessing of digital binary mask in detail, holding input signal a sample time by a zero-order holder which produces a piecewise-constant signal, subsequently doing periodic amplitude modulation on piecewiseconstant signal by a mask signal whose periodicity equals to sample time. In [9], Tezuka simplified the digital binary mask by utilizing a step waveform as a mask signal at the cost of losing variability. In [5], Nakayama introduced analog chaos mask which differs to digital binary mask approach in mask signal. Indeed, he compared this two mask approaches in prediction task and the results are similar to original signals in both cases but response in analog case is more complex and error is smaller.

5 IV.

6 Training

The essence of the training is computing the weight matrix W out by given input driven signals and target readout output signals while other weight matrixes W in , W, W back keep fixed. Note that, regularization parameter ? is also fixed after it is determined by optimal strategy. After W out is determined, feeding the input signals, harvesting the corresponding output readout signals, and calculating the NRMSE.

The values of fixed weight matrix W in , W and W back are arbitrarily initialized by random number generator or Gaussian probability distribution [10] depending on specific circumstances. Once determined, they keep unchanged in the whole training and testing duration.

While the number of the reservoir internal nodes is arbitrarily determined, it should satisfy two constraints, i.e. the structure should be as simple as possible which means the number of reservoir nodes should be small, the another is the NRMSE should be sufficiently small which needs reservoir to be complex and large contrary to the aforementioned constraint [11]. Therefore, a compromise scheme indicating the optimal number of reservoir nodes is considered in experiment.

Training algorithm varies with different applications, but the variation merely is function selection, staying the basic concept unchanged. In electrocardiogram classification application [11], logistic regression algorithm was applied to cope readout process. In multiclass prediction [10], support vector machine was utilized to replace the traditional linear readout layer. In visual contents detection [12], traditional algorithm hyperbolic tangent function was applied. There is another typical function frequently used, namely sigmoid function also called S function.

V.

7 Benchmark Tasks

Benchmark tasks are frequently harnessed to evaluate upgraded reservoir computing structure in many masterpiece works, indicating benchmark tasks possess a vital place in RC research. Therefore, subsequent contents within this section are center on benchmark tasks.

The extremely popular benchmark task is time series prediction due to the time-dependent property, including Mackey Glass chaotic time series prediction [13] and Santa-Fe time series prediction [14,9]. The other long-standing benchmark task is classification covering simple signal classification. Recognition benchmark task is also noticeable, such as isolated spoken digit recognition and digital handwriting recognition [8]. Meanwhile, there are other frequently harnessed benchmark tasks such as the 10th order nonlinear auto regressive moving average (NARMA) system [13,15,4], 5-Bit Parity [4] and nonlinear channel equalization [14,15,16].

8 VI.

Application RC approaches has been widely applied in academic arena. These application spans from bioscience field (in this field RC named LSM) to engineering domain (in this field RC named ESN) that are amenable to supervised model. Meanwhile, the platform constructing the neural network model has transited from software to hardware.

What follows is the application of RC in the last three years.

Classification. Since the classification application will not confine to simple task, the development direction is achieving complex real-time task. One of the examples is electrocardiogram classification [11], where the novel reservoir computing with logistic regression was applied, instead of standard reservoir computing with linear regression. The methodology requires a computationally inexpensive preprocessing of the electrocardiographic signals, which leads to a fast algorithm, approaching a real-time classification solution. Another is classifying three types of wind power ramp events (WPRE) [17] by ESN with support vector machine (SVM), to be specific, ESN for WPRE recognition and SVM for WPRE classification and training this novel model for multiple WPRE prediction.

Recognition. With the passage of time, the recognition contents are increasingly widespread. In [18], a visual contents obtained by a surveillance camera was for detecting the status of a door by a leaky integration RC. In [19], High speed recognition of dispersive Fourier images was achieved by a photonic RC system. In [20], activity recognition, which is vital in activity assistance and smart homes, was achieved by appropriate configuration of RC with low cost and good accuracy.

Prediction. The most research about prediction is temporal task. As the RC technology matures, forecasting short-term stream flow [21], wind power ramp events [17] and BBS score via balance assessment [22] become increasingly accurate. In line with minimizing downstream flood damage and maximizing the generated power with low costs, Bezerra [21] devoted to water source study and attempt to short-term stream flow forecasting with the aid of RC model, which do achieve better performance than traditional approaches. In [17], Manuel investigated alleviating the impact of wind power ramp events by forecasting it via RC model that was modified slightly for better performance. Reference [22] devoted to improving health monitoring in the elderly, Gallicchio conducted measurement campaign on elderly volunteers and obtained Balance dataset including weight value recorded by Wii Balance Board and target BBS score. The whole experiment was conducted on a variant of RC model, a leaky integration Echo State Network model suitable for dealing with the characteristics of the temporal data generated from sensors.

Optical implementation. Semiconductor laser based on optical feedback and injection due to its potential high speed processing capacity repeatedly serves for perfecting RC model. Respecting to how to apply laser, idea differs to everyone. Harnessing only one drive laser converting input signals into optical signals [23,24, ??5], or two lasers including a drive laser and a response laser [5], or nanophotonic crystal cavities [26] diminishing model size, or a ring laser simultaneously computing two unrelated input signals ??27], or a mutually coupled optoelectronic system compensating slowly masking preprocess [9], is the response of aforementioned different applying ideas, which will increase with researchers' unremitting efforts. While adding optical implementation in RC model reaps great performance, classic readout layer may need necessary alters for specific optical methods. Coupling to an analogelectronic readout eliminates offline postprocessing [23].In ??26], the approach about design full optical readout layer is a problem urgently solved. [1]

y

t ? y R
et is the
target
N

output signals applied for producing error feedback by comparing with signals calculated by readout output equation, f(?)is a nonlinear continuous function, typically sigmoid or hyperbolic tangent function,

[Note: x]

Figure 1: the reservoir states, arg (n)

[Lun et al. ()] 'A novel model of leaky integrator echo state network for time-series prediction'. S X Lun , X S Yao , H Y Qi , H F Hu . *Neurocomputing* 2015. 159 (1) p. .

[Gallicchio et al. ()] *A Reservoir Computing Approach for Balance Assessment. Advanced Analysis and Learning on Temporal Data*, C Gallicchio , A Micheli , L Pedrelli , L Fortunati , F Vozzi , O Parodi . 2016. Springer International Publishing.

[Schumacher et al. ()] *An Introduction to Delay-Coupled Reservoir Computing. Artificial Neural Networks*, J Schumacher , H Toutounji , G Pipa . 2015. Springer International Publishing.

[Vinckier et al. ()] 'Autonomous bio-inspired photonic processor based on reservoir computing paradigm'. Q Vinckier , F Duport , A Smerieri , S Massar , M Haelterman . *IEEE 2016 Summer Topicals Meeting Series*, 2016. IEEE.

[Vinckier et al. ()] 'Autonomous bio-inspired photonic processor based on reservoir computing paradigm'. Q Vinckier , F Duport , A Smerieri , S Massar , M Haelterman . *IEEE 2016 Summer Topicals Meeting Series*, 2016. IEEE.

[Bezerra et al. ()] S G T A Bezerra , C B D Andrade , M J S Valença . *Using Reservoir Computing and Trend Information for Short-Term Stream flow Forecasting. Artificial Neural Networks and Machine Learning - ICANN*, 2016. 2016.

[Scardapane et al. ()] 'Distributed reservoir computing with sparse readouts'. S Scardapane , M Panella , D Comminiello , A Hussain , A Uncini . *IEEE Computational Intelligence Magazine* 2016. 11 (4) p. . (research frontier)

[Escalona-Morán et al. ()] 'Electrocardiogram classification using reservoir computing with logistic regression'. M A Escalona-Morán , M C Soriano , I Fischer , C R Mirasso . 10.1109/JBHI.2014.2332001. `https://doi.org/10.1109/JBHI.2014.2332001` *IEEE Journal of Biomedical and Health Informatics* 2015. 19 (3) p. .

[Yi et al. ()] 'FPGA based spike-time dependent encoder and reservoir design in neuromorphic computing processors'. Y Yi , Y Liao , B Wang , X Fu , F Shen , H Hou , L Liu . 10.1016/j.micpro.2016.03.009. `https://doi.org/10.1016/j.micpro.2016.03.009` *Microprocessors and Microsystems* 2015. 46 p. .

[Mesaritakis et al. ()] 'High-speed all-optical pattern recognition of dispersive fourier images through a photonic reservoir computing subsystem'. C Mesaritakis , A Bogris , A Kapsalis , D Syvridis . *Optics Letters* 2015. 40 (14) p. .

[Appeltant et al. ()] 'Information processing using a single dynamical node as complex system'. L Appeltant , M C Soriano , G V D Sande , J Danckaert , S Massar , J Dambre . *Nature Communications* 2011. 2 (9) p. 468.

[Nakayama et al. ()] 'Laser dynamical reservoir computing with consistency: an approach of a chaos mask signal'. J Nakayama , K Kanno , A Uchida . 10.1364/OE.24.008679. `https://doi.org/10.1364/OE.24.008679` *Optics Express* 2016. 24 (8) p. 8679.

[Dorado-Moreno et al. ()] 'Multiclass Prediction of Wind Power Ramp Events Combining Reservoir Computing and Support Vector Machines'. Manuel Dorado-Moreno , Antonio Manuel Dur´an-Rosal , David Guijo-Rubio , Pedro Antonio Guti´errez1 , LP , Sancho Salcedo-Sanz , C As-Mart´ , H . 10.1007/978-3-319-44636-3_28. `https://doi.org/10.1007/978-3-319-44636-3_28` *Canadian Conference on AI* 2016. 9868 p. .

[Dorado-Moreno et al. ()] 'Multiclass Prediction of Wind Power Ramp Events Combining Reservoir Computing and Support Vector Machines'. Manuel Dorado-Moreno , Antonio Manuel Dur´an-Rosal , David Guijo-Rubio , Pedro Antonio Guti´errez1 , LP , Sancho Salcedo-Sanz , C As-Mart´ , H . 10.1007/978-3-319-44636-3_28. `https://doi.org/10.1007/978-3-319-44636-3_28` *Canadian Conference on AI* 2016. 9868 p. .

[Palumbo et al. ()] F Palumbo , C Gallicchio , R Pucci , A Micheli . *of Ambient Intelligence & Smart Environments*, 2015. 8 p. .

[Maass et al. ()] 'Real-time computing without stable states: a new framework for neural computation based on perturbations'. W Maass , Natschl , T Ger , H Markram . *Neural Computation* 2002. 14 (11) p. .

[Jalalvand et al. ()] 'Real-Time Reservoir Computing Network-Based Systems for Detection Tasks on Visual Contents'. A Jalalvand , G Wallendael , Van , R Walle , Van De . 10.1109/CICSyN.2015.35. `https://doi.org/10.1109/CICSyN.2015.35` *Proceedings -7th International Conference on Computational Intelligence, Communication Systems and Networks*, (-7th International Conference on Computational Intelligence, Communication Systems and NetworksCICSyN) 2015. 2015. p. .

[Jalalvand et al. ()] 'Real-Time Reservoir Computing Network-Based Systems for Detection Tasks on Visual Contents'. A Jalalvand , G Wallendael , Van , R Walle , Van De . 10.1109/CICSyN.2015.35. `https://doi.org/10.1109/CICSyN.2015.35` *Proceedings -7th International Conference on Computational Intelligence, Communication Systems and Networks*, (-7th International Conference on Computational Intelligence, Communication Systems and NetworksCICSyN) 2015. 2015. p. .

[Tezuka et al. ()] *Reservoir computing with a slowly modulated mask signal for preprocessing using a mutually coupled optoelectronic system*, M Tezuka , K Kanno , M Bunsen . 2016. 55 p. .

[Nguimdo et al. ()] 'Simultaneous computation of two independent tasks using reservoir computing based on a single photonic nonlinear node with optical feedback'. Modeste Nguimdo , R Verschaffelt , G Danckaert , J Van Der , G Sande . 10.1109/TNNLS.2015.2404346. https://doi.org/10.1109/TNNLS.2015.2404346 *IEEE Transactions on Neural Networks and Learning Systems* 2015. 26 (12) p. .

[Luko?evi?ius and Jaeger ()] 'Survey: reservoir computing approaches to recurrent neural network training'. Mantas Luko?evi?ius , H Jaeger . *Computer Science Review* 2009. 3 (3) p. .

[Takeda et al. ()] S Takeda , D Nakano , T Yamane , G Tanaka , R Nakane , A Hirose . *Photonic Reservoir Computing Based on Laser Dynamics*, 2016.

[Jaeger ()] *The " echo state " approach to analysing and training recurrent neural networks*, H Jaeger . https://doi.org/citeulike-article-id:9635932 2010. p. . (with an Erratum note 1. GMD Report)

[Schneider et al. ()] 'Using digital masks to enhance the bandwidth tolerance and improve the performance of on-chip reservoir computing systems'. B Schneider , J Dambre , P Bienstman . *IEEE Transactions on Neural Networks & Learning Systems* 2016. 27 (12) p. .