Machine learning at the speed of light

Piotr Antonik

April 7, 2022
The 2D Fourier transform

Li Y. et al., Journal of Biomedical Optics 7 (2002)
Summary

1. Photonics for computing
2. Reservoir computing
3. Time-delay RC
4. Parallel RC
5. Conclusion
<table>
<thead>
<tr>
<th>Feature</th>
<th>Electron</th>
<th>Photon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spin</td>
<td>Fermion</td>
<td>Boson</td>
</tr>
<tr>
<td>Charge</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Interaction</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Nonlinear transformations</td>
<td>✓✓</td>
<td>X</td>
</tr>
<tr>
<td>Information transport</td>
<td>✓</td>
<td>✓✓</td>
</tr>
</tbody>
</table>

**Photonics for computing**

- Reservoir computing
- Time-delay RC
- Parallel RC
- Conclusion
Optical Fourier transform

Image source: PhysWiki, York University
Optical matrix multiplication

J. H. Hong et al., Applied Optics 29 (1990)
Digital vs Analogue computing

Image source: towardsdatascience.com
Summary

1. Photonics for computing
2. Reservoir computing
3. Time-delay RC
4. Parallel RC
5. Conclusion
An abstract model of computation

Photonics for computing

Reservoir computing

Time-delay RC

Parallel RC

Conclusion

Photonic Reservoir Computing, De Gruyter (2019)
Reservoir computing with a reservoir

C. Fernando et al., European conference on artificial life (2003)
An abstract model of computation


Photonic Reservoir Computing, De Gruyter (2019)
Echo-state networks

$$x(n) = \tanh \left[ W^{\text{in}} u(n) + W x(n - 1) \right]$$

$$y(n) = W^{\text{out}} x(n)$$

M. Lukoševičius, Neural Netw.: Tricks of the trade (2012)
Linear regression layer:

\[
\text{MSE}(\mathbf{y}, \mathbf{y}^{\text{target}}) = \frac{1}{T} \sum_{n=1}^{T} (\mathbf{y}(n) - \mathbf{y}^{\text{target}}(n))^2
\]

Minimise the MSE with:

\[
\mathbf{W}^{\text{out}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}
\]

Ridge regression (regularisation):

\[
\text{MSE}_{\text{ridge}} = \text{MSE} + \lambda (\mathbf{W}^{\text{out}})^T \mathbf{W}^{\text{out}}
\]
Physical reservoir computing

(a) Conventional RC

(b) Physical RC

G. Tanaka et al., Neural Netw. 115 (2019)
Summary

1. Photonics for computing
2. Reservoir computing
3. Time-delay RC
4. Parallel RC
5. Conclusion
Delay systems in real life

P. Antonik et al., Photoniques 104 (2020)
Dynamics of a simple delay system

\[ \dot{x}(t) = -\alpha x(t - \tau) \]

Photonic Reservoir Computing, De Gruyter (2019)
Delay-based reservoir computing

P. Antonik et al., Photoniques 104 (2020)
Interconnection through desynchronisation

(a) random topology with diagonal element

(b) random topology without diagonal element

(c) ring topology
Interconnection through the inherent dynamics

Photonic Reservoir Computing, De Gruyter (2019)
Mach-Zehnder intensity modulator

Jing Gao et al., OpEx 24 (2016)
Pioneer opto-electronic reservoir computer

P. Antonik et al., Photoniques 104 (2020)
SOA-based setup

P. Antonik et al., Photoniques 104 (2020)
Laser-based setup

Injection laser

Semiconductor laser

Att

Delay

Detector

Input

Readout

P. Antonik et al., Photoniques 104 (2020)
High-Speed Photonic Reservoir Computing Using a Time-Delay-Based Architecture:
Million Words per Second Classification

Laurent Larger, 1 Antonio Baylón-Fuentes, 1 Romain Martinenghi, 1 Vladimir S. Udaltsov, 1,2
Yanne K. Chembo, 1 and Maxime Jacquot 1

1FEMTO-ST Institute/Optics Department, CNRS & University Bourgogne Franche-Comté,
15B avenue des Montboucons, 25030 Besançon Cedex, France
2Vavilov Optical State Institute, Saint-Petersburg, Russia

(Received 30 January 2015; revised manuscript received 13 November 2016; published 6 February 2017)

Reservoir computing, originally referred to as an echo state network or a liquid state machine, is a brain-inspired paradigm for processing temporal information. It involves learning a “read-out” interpretation for nonlinear transients developed by high-dimensional dynamics when the latter is excited by the information signal to be processed. This novel computational paradigm is derived from recurrent neural network and machine learning techniques. It has recently been implemented in photonic hardware for a dynamical system, which opens the path to ultrafast brain-inspired computing. We report on a novel implementation involving an electro-optic phase-delay dynamics designed with off-the-shelf optoelectronic telecom devices, thus providing the targeted wide bandwidth. Computational efficiency is demonstrated experimentally with speech-recognition tasks. State-of-the-art speed performances reach one million words per second, with very low word error rate. Additionally, to record speed processing, our investigations have revealed computing-efficiency improvements through yet-unexplored temporal-information-processing techniques, such as simultaneous multisample injection and pitched sampling at the read-out compared to information “write-in”.

DOI: 10.1103/PhysRevX.7.011015

Subject Areas: Complex Systems, Nonlinear Dynamics, Photonics
Final thoughts on time-delay systems

Time-delay photonic reservoir computers:

• were the first experimental demonstrations

• are bulky $\Rightarrow$ integration

• are difficult to scale up $\Rightarrow$ parallel systems
Summary

1. Photonics for computing
2. Reservoir computing
3. Time-delay RC
4. Parallel RC
5. Conclusion
Spatial Light Modulator

Jullien A., Photoniques 101 (2020)
Spatial Light Modulator

Jullien A., Photoniques 101 (2020)
SLM-based reservoir computer

Image and video processing


### Table 1 | Performance of various state-of-the-art digital approaches compared to our best experimental result

<table>
<thead>
<tr>
<th>Authors</th>
<th>Method</th>
<th>Database split</th>
<th>Training time</th>
<th>Processing speed (f.p.s.)</th>
<th>s1 scenario (%)</th>
<th>Full database (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yadav et al.(^{38})</td>
<td>IP + SVM</td>
<td>80%–20%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>98.20</td>
</tr>
<tr>
<td>Shi et al.(^{39})</td>
<td>DTD, DNN</td>
<td>9–16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>95.6</td>
</tr>
<tr>
<td>Kovashka et al.(^{40})</td>
<td>BoW + SVM</td>
<td>8–8–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.53</td>
</tr>
<tr>
<td>Gilbert et al.(^{33})</td>
<td>HCF + SVM</td>
<td>LOOCV</td>
<td>-5.6 h</td>
<td>24</td>
<td>-</td>
<td>94.5</td>
</tr>
<tr>
<td>Baccouche et al.(^{31})</td>
<td>CNN &amp; RNN</td>
<td>16–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.39</td>
</tr>
<tr>
<td>Ali and Wang(^{42})</td>
<td>DBN &amp; SVM</td>
<td>50%–20%–30%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.3</td>
</tr>
<tr>
<td>Wang et al.(^{43})</td>
<td>DT + SVM</td>
<td>16–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.2</td>
</tr>
<tr>
<td>Liu et al.(^{44})</td>
<td>MMI + SVM</td>
<td>LOOCV</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.15</td>
</tr>
<tr>
<td>Sun et al.(^{45})</td>
<td>FT + SVM</td>
<td>Auto</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.0</td>
</tr>
<tr>
<td>Veeriah et al.(^{46})</td>
<td>Differential RNN</td>
<td>16–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>93.96</td>
</tr>
<tr>
<td>Shu et al.(^{47})</td>
<td>SNN</td>
<td>9–16</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>92.3</td>
</tr>
<tr>
<td>Laptev et al.(^{48})</td>
<td>FT + SVM</td>
<td>8–8–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>91.8</td>
</tr>
<tr>
<td>Jhuang(^{31})</td>
<td>StC(_2)+ SVM</td>
<td>16–9</td>
<td>0.4</td>
<td>96.0</td>
<td>91.6</td>
<td></td>
</tr>
<tr>
<td>Klaeser et al.(^{49})</td>
<td>3D Grad + SVM</td>
<td>8–8–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>91.4</td>
</tr>
<tr>
<td><strong>This work</strong></td>
<td><strong>Photonic RC</strong></td>
<td><strong>75%–25%</strong></td>
<td><strong>1.6–5.5 h</strong></td>
<td><strong>2–7</strong></td>
<td><strong>91.3</strong></td>
<td>-</td>
</tr>
<tr>
<td>Grushin et al.(^{52})</td>
<td>LSTM</td>
<td>16–9</td>
<td>1 day</td>
<td>12–15</td>
<td>-</td>
<td>90.7</td>
</tr>
<tr>
<td>Ji et al.(^{50})</td>
<td>3DCNN</td>
<td>8–8–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>90.02</td>
</tr>
<tr>
<td>Escobar et al.(^{51})</td>
<td>MT cells</td>
<td>16–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>74.63</td>
</tr>
<tr>
<td>Schuldte et al.(^{23})</td>
<td>FT + SVM</td>
<td>8–8–9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>71.83</td>
</tr>
</tbody>
</table>

‘Database split’ indicates how the KTH database was split for training and testing of the system. Most studies choose to split by the number of subjects into either two groups (training and test, for example 16 subjects for training; 9 for the test) or three groups (training, validation and test, for example 8–8–9). LOOCV corresponds to ‘leave-one-out cross validation’; the system is trained on 24 subjects and tested on the remaining one. Training times and processing speeds are not discussed in most of the works, focusing on the classification performance. Some studies report specific results on the s1 scenario, as considered in this work. BoW, bag of words; CNN, convolutional neural network; DBN, deep belief network; DNN, deep neural network; DT, dense trajectories; DTD, deep trajectory descriptor; FT, features; HCF, hierarchical compound features; IP, interest points; LSTM, long short-term memory neural network; MMI, maximization of mutual information; MT, middle temporal area of the visual cortex; RNN, recurrent neural network; SNN, spiking neural network; StC\(_2\), space-time oriented C\(_2\) features; SVM, support vector machine.
More advanced setup

Photonic Reservoir Computing, De Gruyter (2019)
Networks of vertically emitting lasers

Photonic Reservoir Computing, De Gruyter (2019)
Summary

1. Photonics for computing
2. Reservoir computing
3. Time-delay RC
4. Parallel RC
5. Conclusion
The results were obtained with the help of these awesome people:

Serge Massar  Daniel Brunner  Damien Rontani  Nicolas Marsal

And these generous organisations:
For future reference

Photonics for computing

Reservoir computing

Time-delay RC

Parallel RC

Conclusion
Why change the substrate?

- Photonics are getting better
- Moore’s law doomed to fail
- New models for computation

P. Antonik et al., Photoniques 104 (2020)  
P. Antonik et al., JSTQE 26 (2019)
Final thoughts

Why change the substrate?

- Photonics are getting better
- Moore’s law doomed to fail
- New models for computation

P. Antonik et al., Photoniques 104 (2020)  
P. Antonik et al., JSTQE 26 (2019)

Thank you!