2022 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2022, Virtual, December 12-15, 2022

Transfer learning based on photonic reservoir computing using semiconductor laser with optical feedback

Rie Sakamaki[1], Kazutaka Kanno[1], Masanobu Inubushi[2], and Atsushi Uchida[1]

[1]Department of Information and Computer Sciences, Saitama University
255 Shimo-Okubo, Sakura-ku, Saitama City, Saitama, 338-8570, Japan
[2]Department of Applied Mathematics, Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
Email: r.sakamaki.543@saitama-u.ac.jp, auchida@mail.saitama-u.ac.jp

Abstract– Transfer learning has been studied to reduce the learning cost when the system parameters are changed. We numerically investigate transfer learning based on photonic reservoir computing using a semiconductor laser with optical feedback modulation. We succeed in inferring the dynamics of one variable in the Lorenz model whose parameters are changed from the learning scheme.

1. Introduction

Neural networks are crucial for machine learning that imitate the structure of the human brain. One type of the neural networks is the recurrent neural networks, which are the networks with self-feedback connections of neurons and retain past information in the network as memory. Recurrent neural networks need to learn the weights among connections in the input, network, and output layers, and large computational cost is required. Recently, reservoir computing has been proposed, which is derived from the recurrent neural networks and used for performing speech recognition and time series prediction [1]. The advantage of reservoir computing is easy learning of the weights because the weights in the network are fixed randomly and only the output weights are trained.

Recently, physical implementations of delay-based reservoir computing have been proposed [2]. In this method, the network is constructed by defining virtual nodes using temporal waveforms of laser output in a time-delayed feedback loop. Several methods of photonic delay-based reservoir computing have been reported using semiconductor lasers [3-6].

It is important to satisfy two properties for the implementation of reservoir computing: high dimensionality and consistency [7]. High dimensionality is the property of reservoir computing to convert low dimensional input signals into high dimensional signals. Consistency is the property of exhibiting the same output signal from the same input signal. It has been shown that semiconductor lasers with optical feedback exhibit consistency, although they generate chaotic signals by time-delayed optical feedback. Therefore, semiconductor lasers are considered as suitable photonic sources for implementing delay-based reservoir computing.

Transfer learning has been proposed [8,9], which is a learning method that transfers the knowledge (weights) from one system with large training data to another system with small training data. Transfer learning is effective under the condition where the property of training data is changed frequently. Transfer learning is useful for inferring unmeasured variables when the amount of training data is insufficient. However, transfer learning has not been applied to photonic delay-based reservoir computing.

In this study, we numerically investigate transfer learning based on photonic reservoir computing using a semiconductor laser with optical feedback modulation.

2. Numerical model of semiconductor laser with optical feedback modulation.

We use a semiconductor laser with optical feedback modulation for reservoir computing. The light emitted from the laser is injected into a phase modulator, reflected by an external mirror, and reinjected into the laser. The numerical model of the Lang-Kobayashi equations can be expressed as follows [10].

$$\frac{dE(t)}{dt}=\frac{1+i\alpha}{2}\left\{\frac{G_N(N(t)-N_0)}{1+\epsilon|E(t)|^2}-\frac{1}{\tau_p}\right\}E(t) \quad (1)$$
$$+\kappa E(t-\tau)exp[i(-\omega\tau+\phi_0+S(t))]+\xi(t)$$

$$\frac{dN(t)}{dt}=J-\frac{N(t)}{\tau_s}-\frac{G_N(N(t)-N_0)}{1+\epsilon|E(t)|^2}|E(t)|^2 \quad (2)$$

Where $E(t)$ is the complex electric-field amplitude and $N(t)$ is the carrier density. In Eq. (1), the term including G_N represents spontaneous emission, and gain saturation is taken into account. Numerical simulations can reproduce values close to the laser output intensity obtained in the experiments. $-1/\tau_p$ is the decay rate of the electric-field amplitude, τ is the feedback delay time, and ϕ_0 is the bias of the optical feedback phase. In Eq. (2), $-1/\tau_s$ represents the decay rate of the carrier density.

3. Concept of delay-based reservoir computing

Figure 1 shows a schematic diagram of delay-based reservoir computing. Reservoir computing consists of an input layer, a reservoir layer, and an output layer. The input layer generates an input modulation signal for reservoir computing. First, the input data are expanded for a mask period. The mask period is set to be equal to the feedback delay time τ. Next, a binary random mask signal with an interval θ is generated and multiplied to the input data. The input scaling factor is also multiplied, which is a parameter that determines the magnitude of the modulation signal. The modulation signal is expressed as follows.

$$S(t) = u(t) \times m(t) \times \gamma \tag{3}$$

Where $S(t)$ is the modulation signal, $u(t)$ is the input signal, $m(t)$ is the mask signal, and γ is the input scaling factor. The generated modulation signal is injected into the reservoir layer.

The reservoir layer consists of a semiconductor laser and a time-delayed feedback loop. Complex transient dynamics is obtained from the laser, and the temporal waveforms of the laser output are generated. Virtual node states are obtained by measuring the temporal waveforms. The number of virtual nodes N is determined as follows.

$$N = \frac{\tau}{\theta} \tag{4}$$

Where τ is the delay time and θ is the node interval.

In the output layer, the weighted linear sum of the virtual node states is calculated as the output signal as follows.

$$Y(n) = \sum_{i=1}^{N} w_i x_i(n) \tag{5}$$

where $Y(n)$ is the output for the n-th input and x_i is the i-th virtual node state. The weight w_i of the i-th virtual node states is calculated using the least squares method as follows.

$$\frac{1}{L}\sum_{n=1}^{L}\left(y(n) - Y(n)\right)^2 \rightarrow min \tag{6}$$

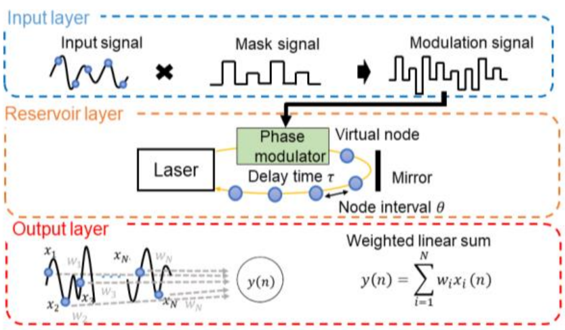

Fig. 1: Schematic diagram of delay-based reservoir computing.

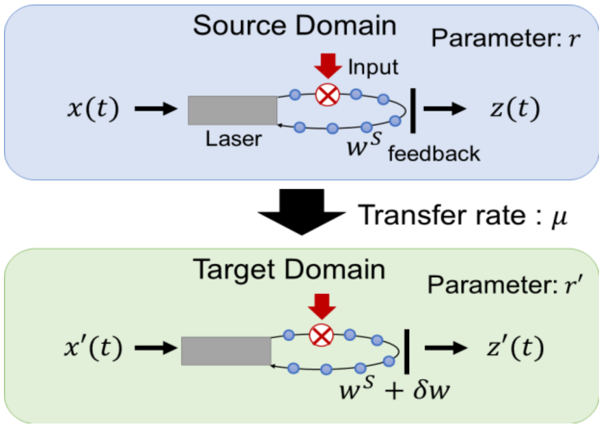

Fig. 2: Schematic diagram of transfer learning based on reservoir computing.

4. Transfer learning based on reservoir computing

A schematic diagram of the transfer learning used in this study is shown in Fig. 2. Transfer learning consists of two domains: source and target domain. In the source domain, we train the weights w^S of reservoir computing using 30000 training data. In the training process, reservoir computing infers one variable from another variable of the dynamical model with a parameter value.

In the target domain, the correction weights δw are calculated using a fewer training data (500 data) for the dynamical model with a slightly different parameter value. The correction weights δw are calculated to minimize the following equation [8].

$$\varepsilon(\delta w) = \langle\left(v'(t) - \sum_{i=1}^{N}(w_i^S + \delta w_i) r_i'(t)\right)^2\rangle_{T_L'} + \mu\|\delta w\|_2^2 \tag{7}$$

where $v'(t)$ is target value, $r_i'(t)$ is the i-th virtual node state, $\langle a\rangle_T := \frac{1}{T}\Sigma_{k=1}^{T} a(k)$ and $\|\delta w\|_2^2 = \Sigma_{i=1}^{N}\delta w_i^2$. We denote the weights calculated in the source domain by w^S. The amount of knowledge transfer from the source to target domain is controlled by the transfer rate μ.

5. Inference task

As a task for reservoir computing and transfer learning, we infer the dynamics of one variable from another variable in the Lorenz model with different parameters. The Lorenz model is described as follows.

$$\frac{dx}{dt} = \sigma(y - x) \tag{8}$$

$$\frac{dy}{dt} = x(r - z) - y \tag{9}$$

$$\frac{dz}{dt} = xy - bz \tag{10}$$

Where σ is set to 10 and b is set to 8/3. r is a variable.

The performance of reservoir computing is evaluated by using the normalized mean square error (NMSE). NMSE is defined by the following equation.

$$NMSE = \frac{1}{L}\frac{\sum_{n=1}^{L}\left(\bar{y}(n) - y(n)\right)^2}{var(\bar{y})} \qquad (11)$$

where L is the total number of data, n is the time step of the input data, $y(n)$ is the output of reservoir computing, $\bar{y}(\mathrm{n})$ is the ideal target, and $var(\bar{y})$ is the variance of $\bar{y}(\mathrm{n})$. A smaller value for NMSE indicates higher performance.

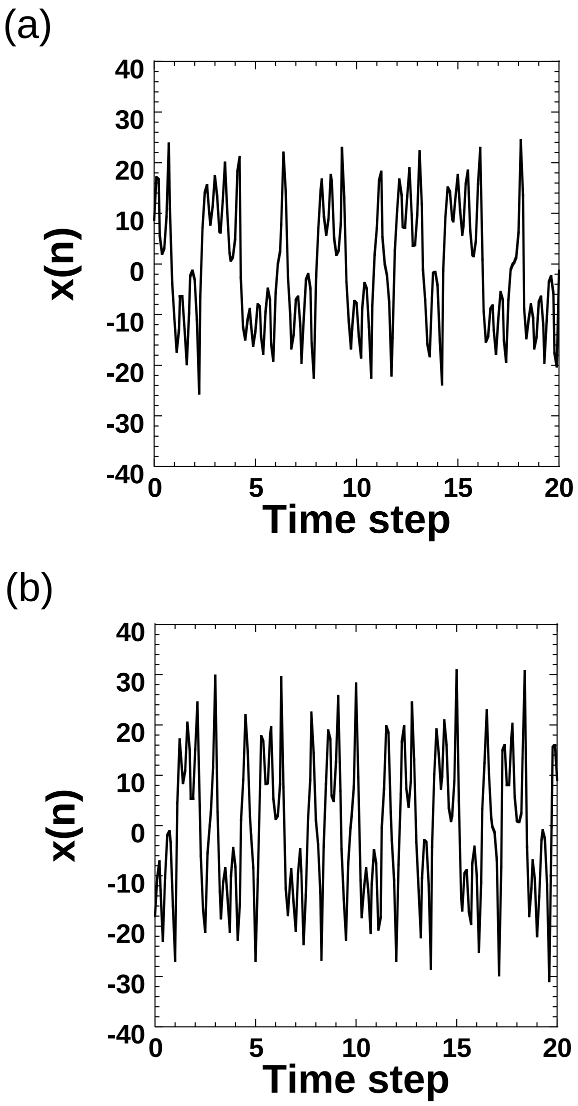

Fig. 3: Temporal waveforms of the Lorenz model with the parameter value of (a) $r = 60$ (c) and (b) $r' = 80$.

6. Results

We numerically investigate the performance of inference of transfer learning based on reservoir computing using a semiconductor laser. We set the parameter values of r = 60 in the source domain, and r' = 80 in the target domain. We use 30000 and 500 training data in the source and target domains, respectively. Figure 3 shows the temporal waveforms of the source and target domains. Chaotic dynamics are changed at different parameter values.

Figure 4(a) shows the result of inferred output $z(t)$ from the input $x(t)$ in the Lorenz model without transfer learning. The black line indicates the temporal waveform of the target, the red line indicates the inferred waveform, and the blue line indicates the error between them. The input and inferred signals are very different, and large errors are obtained.

Figure 4(b) shows the result of inferred output $z(t)$ from the input $x(t)$ in the Lorenz model with transfer learning. The input and inferred signals look very similar. The value of NMSE is 0.0398 with transfer learning. Therefore, we succeed in performing transfer learning when the parameter value of the Lorenz model is changed.

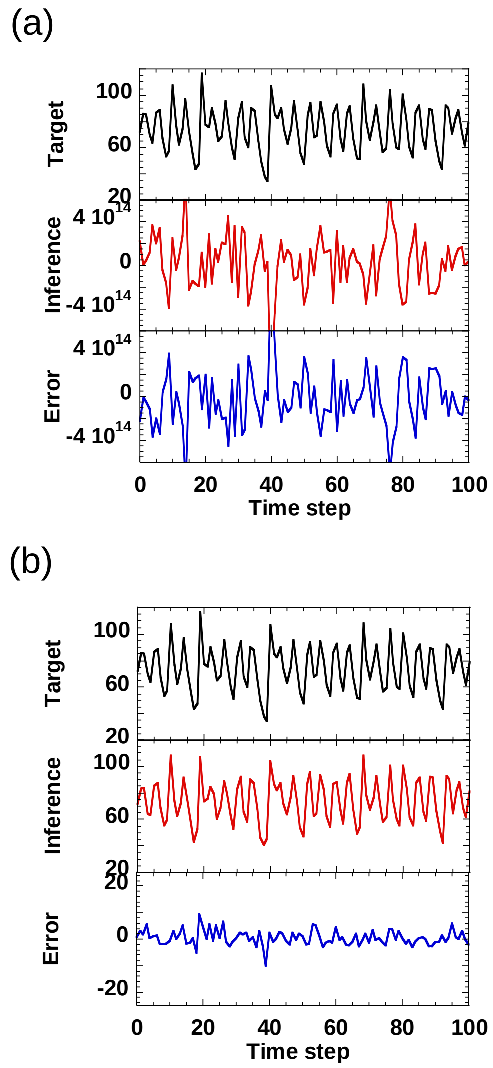

Fig. 4: Temporal waveforms of inference of $z(t)$ by using $x(t)$ (a) without and (b) with transfer learning. The target (black), inferred (red) and error (blue) signals are shown.

Next, we investigate the dependence of the performance of transfer learning on the amount of training data. Figure 5 shows the change in NMSE by varying the amount of training data in the target domain. The black line represents NMSEs in the target domain with transfer learning at the transfer rate $\mu = 10$, and the red line represents the NMSEs at the transfer rate $\mu = 0$. The blue dashed line represents the NMSEs without transfer learning. NMSEs are matched in the cases of the transfer rate $\mu = 0$ and no transfer learning. In the case of transfer rate $\mu = 0$, trained weights in the source domain are not used, and only the trained weights in the target domain are used. The difference between the black and red lines indicates that transfer learning is effective to reduce NMSEs when the amount of training data is small.

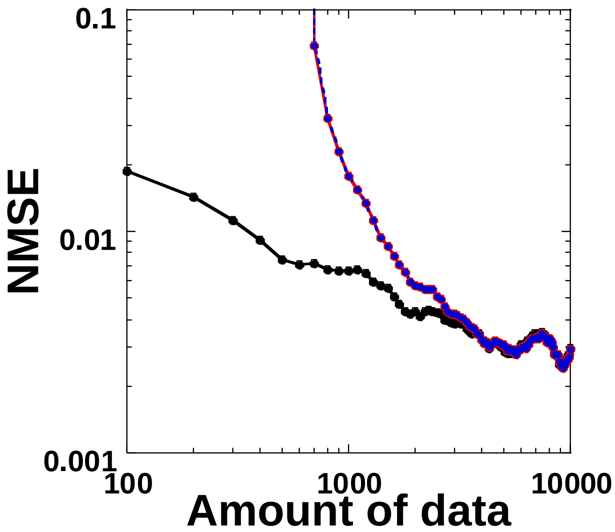

Fig. 5: Inference error (NMSEs) as the amount of training data in target domain is changed. The transfer rate is $\mu = 10$ (black) and $\mu = 0$ (red). No transfer learning is used (blue).

7. Conclusions

In our study, we numerically investigated transfer learning based on photonic reservoir computing using a semiconductor laser with optical feedback. We succeeded in inferring one variable from another variable in the Lorenz model with different parameter values in the source and target domains. We found that transfer learning is effective when the amount of data is small in the target domain.

Acknowledgements

This study was supported in part by JSPS KAKENHI (JP19H00868, JP20K15185, JP20H00233, JP22H05195), JST CREST (JPMJCR17N2), and the Telecommunications Advancement Foundation.

References

[1] H. Jaeger and H. Hass, “Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication,” Science, Vol. 304, pp. 78-80 (2004).

[2] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, “Information processing using a single dynamical node as complex system,” Nature Communications, Vol. 2, Article No. 468 (2011).

[3] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, “Parallel photonic information processing gigabyte per second data rates transient states,” Nature Communications, Vol. 4, No. 1, pp.1-7 (2013).

[4] K. Hicke, M. A. Escalona-Moran, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, “Information processing using transient dynamics of semiconductor lasers subject to delayed feedback,” IEEE Journal of Selected Topics in Quantum Electronics, Vol. 19, No. 4, pp. 1501610 (2013).

[5] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, “Optoelectronic reservoir computing,” Scientific Reports, Vol. 2, Article No. 287, pp. 1-15 (2012).

[6] R. M. Nguimdo, E. Lacot, O.Jacquin, O. Hugon, G. V. der Sande, and H. G. de Chatellus, “Prediction performance of reservoir computing systems based on a diode-pumped erbium-doped microtip laser subject to optical feedback,” Optics Letters, Vol. 42, No. 3, pp. 375-378 (2017).

[7] A. Uchida, R. McAllister, and R. Roy, “Consistency of nonlinear system response to complex drive signals,” Physical Review Letters, Vol. 93, pp. 244102 (2004).

[8] M. Inubushi and S. Goto, “Transfer learning for nonlinear dynamics and its application to fluid turbulence,” Physical Review E, Vol. 102, pp. 043301 (2020).

[9] Y. Guo, H. Zhang, L. Wang, H. Fan, J. Xiao, and X. Wang, “Transfer learning of chaotic systems,” Chaos, Vol. 31, No. 1, pp. 011104 (2021).

[10] R. Lang and K. Kobayashi, “External optical feedback effects on semiconductor injection laser properties,” IEEE Journal of Quantum Electronics, Vol. 16, No. 3, pp. 347-355 (1980).