A Reservoir Computing Approach for Human Gesture Recognition from Kinect Data

Claudio Gallicchio (✉) and Alessio Micheli

Department of Computer Science, University of Pisa,
Largo B. Pontecorvo 3, Pisa, Italy
gallicch@di.unipi.it, micheli@di.unipi.it

Abstract. This paper describes a novel approach for human gesture recognition from motion data captured by a Kinect camera. The proposed method is based on encoding the temporal history of input data using bidirectional Echo State Networks, whereas the output is computed by means of a multi-layer perceptron with softmax. Results achieved at the time-series classification challenge organized within the 2016 ECML PKDD Workshop on Advanced Analytics and Learning on Temporal Data show the potentiality of the approach.

Keywords: Human Gesture Recognition, Reservoir Computing, Echo State Networks, Learning with Temporal Data

1 Introduction

A major aspect of Ambient Assisted Living (AAL) and Ambient Intelligence (AmI) applications regards the development of human-centric computer interfaces [12, 28]. Indeed, interfaces that are easy and natural to use allow for a simplification of the interaction between the user and the intelligent environment, ultimately leading to an overall improvement of acceptance and usability of the developed systems. In particular, in the area of human-machine interaction a relevant task is represented by the the automatic recognition of human gestures [32], where the challenge consists in interpreting the sensed data in order to recognize patterns of human body motion in a robust fashion. In recent years, the availability of relatively cheap cameras and motion sensor devices, such as Microsoft Kinect, allowed for a broader diffusion of methods based on captured data in the form of a set of 3-dimensional trajectories of human skeleton joints. In this context, literature approaches mainly consist in systems that exploit the extraction of relevant features from such 3-dimensional trajectories, along with classification methods based e.g. on multi-layer perceptrons (MLPs) [34], support vector machines [34, 8, 7], dynamic time warping [31, 23, 16], decision trees [34, 7], or hidden Markov models [23].

In this paper we propose a novel approach to the problem of human gesture recognition from noisy 3-dimensional motion data from a Kinect device, based on the efficient Reservoir Computing (RC) [30] paradigm for modeling Recurrent Neural Networks (RNNs) [29]. In particular, we describe the application

of this RC-based method to the time-series classification challenge organized in the context of the 2nd ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data (AALTD2016) [1]. A major characterization of the proposed approach consists in addressing the problem of learning with temporal data through a direct processing of time-series signals without further steps of feature extraction with respect to the provided challenge datasets.

2 RC for Human Gesture Recognition

RC represents a framework for modeling RNNs based on a conceptual and practical separation between an untrained recurrent component that encodes the input history within the "reservoir" state of the network, and a readout component, which is trained to compute the output based on the information in the reservoir state space. In particular, within this context, the Echo State Network (ESN) model [26, 25, 17] is considered as a state-of-the-art approach for efficiently learning in sequential/temporal domains, with outstanding results in various real-world domains (possibly involving heterogeneous and noisy input information) such as time-series prediction [26], financial forecasting [13], sentiment analysis [19], speech processing [38, 39] health care monitoring [20, 21], human activity recognition (e.g. [33, 3, 2, 14]) and robotics (e.g. [15, 5, 4, 36]).

In our approach to human gesture recognition, we used Leaky Integrator ESNs (LI-ESNs) [27], in which the reservoir part of the network is implemented by means of leaky integrator units. We modeled the evolution of the network state dynamics by splitting the reservoir into two parts that receive the input elements in opposite orders, i.e. from left to right and from right to left, respectively. Such strategy is rooted in the *bidirectional* approaches for RNN dynamics [37, 6], also explored in the context of RC models (e.g. [39, 35]). Note that the use of this bidirectional strategy allows the network to develop a state dynamics that at each time step is able to include information coming from both the left side and right side of the sequence, thereby resulting in a richer state representation of the input than adopting a standard uni-directional strategy. Moreover, in order to process the input information in the two directions at the same time, it is required that each input sequence is entirely available during the encoding process, an assumption that is practically fulfilled in our application context and in the AALTD2016 challenge, in which each sequence corresponds to an isolated human gesture that should be classified as soon as its execution has been completed.

Considering an input sequence $\mathbf{s}$ of length $L_{\mathbf{s}}$, i.e. $\mathbf{s} = [\mathbf{u}(1) \dots \mathbf{u}(L_{\mathbf{s}})]$, where $\mathbf{u}(t) \in \mathbb{R}^{N_U}$ for each $t = 1 \dots L_{\mathbf{s}}$, the state of the bidirectional LI-ESN is computed by applying the following state update equations:

$$\begin{aligned} \mathbf{x}_F(t) &= (1-a)\mathbf{x}_F(t-1) + a\tanh(\mathbf{W}_{in}\mathbf{u}(t) + \hat{\mathbf{W}}\mathbf{x}_F(t-1)) \\ \mathbf{x}_B(t) &= (1-a)\mathbf{x}_B(t+1) + a\tanh(\mathbf{W}_{in}\mathbf{u}(t) + \hat{\mathbf{W}}\mathbf{x}_B(t+1)) \end{aligned} \tag{1}$$

where $\mathbf{x}_F(t) \in \mathbb{R}^{N_R}$ and $\mathbf{x}_B(t) \in \mathbb{R}^{N_R}$ denote the states computed by the reservoir receiving the input elements respectively in the forward direction (i.e. from the oldest to the most recent one) and in the backward direction (i.e. from the most recent to the oldest one), starting from null initial states $\mathbf{x}_F(0) = \mathbf{0}$ and $\mathbf{x}_B(L_\mathbf{s}) = \mathbf{0}$. Moreover, in equation 1, $a \in [0, 1]$ denotes the leaking rate, $\mathbf{W}_{in} \in \mathbb{R}^{N_R \times N_U + 1}$ is the input-to-reservoir weight matrix (including a bias term), $\hat{\mathbf{W}} \in \mathbb{R}^{N_R \times N_R}$ is the recurrent reservoir weight matrix[1]. The overall state of the bidirectional LI-ESN at time t, i.e. $\mathbf{x}(t) \in \mathbb{R}^{2N_R}$, is then computed as:

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_F(t) \\ \mathbf{x}_B(t) \end{bmatrix} \tag{2}$$

By applying equations 1 and 2, each input sequence $\mathbf{s} = [\mathbf{u}(1) \dots \mathbf{u}(L_\mathbf{s})]$ is encoded into a sequence of states $[\mathbf{x}(1) \dots \mathbf{x}(L_\mathbf{s})]$. In tasks requiring one single output element in correspondence of each entire input sequence, a *state mapping function* [18] can be used to map the variable-size state representation $[\mathbf{x}(1) \dots \mathbf{x}(L_\mathbf{s})]$ into a fixed-size reservoir state $\chi(\mathbf{s}) \in \mathbb{R}^{2N_R}$. In particular, the use of a mean state mapping has proved to be effective in different application contexts, as reported in [18, 20, 19]. In this case, the states computed at each time step are averaged and $\chi(\mathbf{s})$ is computed as follows:

$$\chi(\mathbf{s}) = \frac{1}{L_\mathbf{s}} \sum_{t=1}^{L_\mathbf{s}} \mathbf{x}(t). \tag{3}$$

The state encoding process operated by the bidirectional LI-ESN and the computation of the mean state mapping function are graphically illustrated in Figure 1.

The output of the state mapping function is then used as input for the readout component, implemented by means of a MLP with N_H hidden units and in which the last layer is a softmax layer. In correspondence of each input sequence $\mathbf{s}$, the output is therefore a 6-class probability distribution $\mathbf{y}(\mathbf{s}) \in \mathbb{R}^{N_Y}$, where N_Y denotes the number of possible class labels, i.e. the number of possible human gestures to detect (we used $N_Y = 6$ as detailed in the following).

Following the RC approach, the output part of the system, i.e. the MLP in our case, is the only component of the network architecture that undergoes a training process, in this case implemented by means of scaled conjugate gradient backpropagation. The parameters of the reservoir are left untrained after being initialized basing on the necessary condition for the *echo state property* [25, 26], related to the stability of network state dynamics and involving the scaling of the spectral radius of the matrix $(1 - a)\mathbf{I} + a\hat{\mathbf{W}}$, denoted by ρ [2]. Although the standard ESN recipe prescribes that $\rho < 1$, stability of the network dynamics can

[1] Note that, although in general the parameters of the reservoirs spanning the input in the two opposite directions could be different, in our approach for the sake of simplicity the same reservoir is used to compute both $\mathbf{x}_F(t)$ and $\mathbf{x}_B(t)$.

[2] The spectral radius ρ is defined as the maximum among the magnitudes of the eigenvalues of the matrix $(1 - a)\mathbf{I} + a\hat{\mathbf{W}}$.

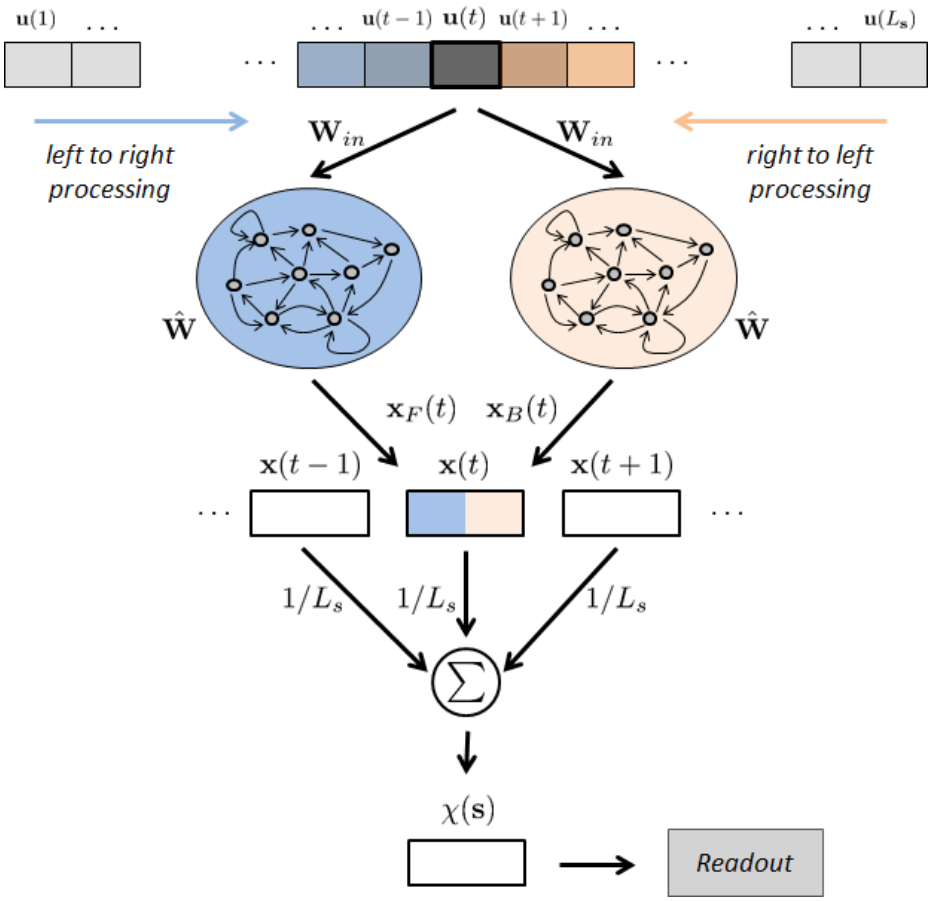

Fig. 1. Bidirectional LI-ESN: the encoding process and the mean state mapping.

be achieved also if this condition is not satisfied, depending on the actual data that is fed in input to the network (see e.g. [40, 9]). Thereby, in our experiments we also explored values of ρ slightly larger than 1. The input weights in matrix $\mathbf{W}_{in}$ were randomly chosen from a uniform distribution in $[-scale_{in}, scale_{in}]$. In addition to this, in our implementation we considered a pattern of connectivity among the reservoir units described by a permutation matrix, i.e. $\hat{\mathbf{W}} = \mathbf{DP}$, where $\mathbf{D} \in \mathbb{R}^{N_R \times N_R}$ is a diagonal matrix (containing the non-zero elements of $\hat{\mathbf{W}}$) and $\mathbf{P} \in \mathbb{R}^{N_R \times N_R}$ is obtained by a (column) permutation of the identity matrix. Reservoirs initialized in such a way are related to a critical regime of network dynamics, which has been shown to have a beneficial effect on the predictive performance in different ESN applications (see e.g. [11, 10, 22]).

3 Experiments

The RC-based system for human gesture recognition described in Section 2 has been assessed within the time-series classification challenge organized in the context of AALTD2016 [1]. In particular, the proposed method has been applied to *task 1* of the challenge, in correspondence of team name "CIML" [3] and method name "RC".

The task consisted in a multi-class classification of data recorded by a Kinect system during the execution of isolated gestures by different users. The input was collected as a multivariate time-series of 3-dimensional data gathered in correspondence of sensors located at 8 body positions: left hand tip, right hand tip, left elbow, right elbow, left wrist, right wrist, left thumb and right thumb. Target data consisted in the type of gesture performed, within a set of 6 possible values. Overall, this setting resulted in an input of size $N_U = 24$, whereas the target was represented by means of a 1-of-6 hot encoding of the class label, i.e. $N_Y = 6$. In our approach, a preliminary pre-processing step has been individually applied to the input sequences in order to re-scale the input components to the [-1,1] range.

Data for the challenge[4] has been provided by means of a labeled (balanced) training set (with target information) and a blind test set (without target information), each containing 180 sequences. Data in the training set have been used for a *development phase*, consisting in experimental assessment and model selection according to a stratified 6-fold cross-validation scheme. After that, the RC setting leading to the best accuracy on the validation set was selected and successively adopted for training on the whole training set and assessment on the blind test set, using an ensemble approach as described in the following.

In our experiments we considered network hyper-parametrizations varying the values of: reservoir size $N_R \in \{200, 500\}$, spectral radius $\rho \in \{0.9, 1, 1.1\}$ and leaking rate $a \in \{0.1, 0.5, 1\}$. The input scaling and the number of MLP hidden units were fixed based on preliminary experiments to the values $scale_{in} = 1$ and $N_H = 10$, respectively. For each reservoir hyper-parametrization we independently and randomly generated 10 reservoir guesses, and averaged the performance on such guesses. As for the challenge, during the model development phase the predictive performance of the RC networks has been assessed by computing the multi-class (6-class) accuracy (i.e. the rate of correctly classified sequences).

The values of the hyper-parameters selected by the model selection process during the development phase and the correspondingly obtained validation accuracy are reported in Table 1. As it can be seen, the accuracy achieved on the validation set in the 6-fold cross-validation scheme is 95.7% ($\pm$5.6). This result indicates that the proposed approach was able to tackle the human gesture recognition task achieving a very good performance with a validation accuracy that is in line with the best results reported in literature (see e.g. [24]). It is

[3] Computational Intelligence & Machine Learning (CIML) group, Department of Computer Science, University of Pisa. Website http://www.di.unipi.it/groups/ciml/.

[4] Datasets are property of IRISA, research team EXPRESSION, see [1].

Hyper-parameter	Selected Value
reservoir size	500
spectral radius	1
leaking rate	0.1

Validation Set Accuracy
95.7% (±5.6)

Table 1. Values of RC hyper-parameters selected in the development phase and corresponding validation accuracy.

also worth reporting that the selected RC configuration in Table 1 led to a 100% training accuracy, whereas configurations with a lower training performance also generally resulted in a worse validation performance.

The selected configuration was then considered for a further experimental phase in which training has been performed on the whole training set of the challenge and the final classification for each sequence was computed by an ensemble of 30 RC networks, all with the same hyper-parametrization as in Table 1 (but differing for the RC initialization values). The achieved accuracy is reported in Table 2, showing the performance on the validation set (according to the same 6-folds splitting considered in the development phase), as well as the result on the blind test set as provided by the challenge organizers after the submission deadline. As can be seen in Table 2, the proposed RC-based approach achieved a 94.4% accuracy on the test set of the AALTD2016 challenge, ranking 5-th overall in the competition, with an accuracy in the top 4 best performances on a total number of 22 submissions (spanning test set accuracy values in the range 78.9% - 96.1%). The official leaderboard of the challenge, showing the final results of all the submissions can be found at [1]. Moreover, a comparison between the values of the validation accuracy obtained by the proposed approach in the development phase (Table 1) and in the final setting (Table 2), points out the significant improvement obtained by the ensemble method, resulting in a performance gain of more than 2% (on the same data).

Validation Set	Test Set
97.8%	94.4%

Table 2. Accuracy achieved by the ensemble of 30 RC networks on the validation and test sets. The performance on the validation set corresponds to the same 6-fold cross-validation scheme used in the development phase, whereas the performance on the test set corresponds to the accuracy obtained on the blind test set of the AALTD 2016 challenge task, as reported by the organizers.

4 Conclusions

We have presented a novel approach for human gesture recognition from Kinect data, based on an ensemble of bidirectional RC networks using MLP readouts with softmax. The proposed method has been experimentally assessed during the AALTD2016 time-series classification challenge, achieving a classification accuracy of 97.8% in the development phase, and of 94.4% on the blind test set of the challenge. The outcome of the challenge showed that our approach compared well with the heterogeneity of methods used by the challenge participants (see details at [1]), with a classification accuracy within the best 4 values over 22 submissions.

Overall, the experimental analysis described in this paper has put in evidence the potentiality of the proposed RC-based approach, especially in light of its suitability for direct processing of time-series data, its general applicability (it has not been specifically tailored for this type of application) and of its training efficiency (typical of all RC models). Such characterizations allow us to envisage possible developments within integrated activity monitoring systems for AAL and AmI applications, able to jointly exploit both a good recognition rate of human gestures and a fast re-training e.g. in presence of concept drifts.

Acknowledgments

The authors would like to thank the organizers of the AALTD2016 challenge.

References

1. Time Series Classification Challenge, 2nd ECML PKDD Workshop on Advanced Analytics and Learning on Temporal Data. https://aaltd16.irisa.fr/challenge/ (2016)
2. Amato, G., Bacciu, D., Broxvall, M., Chessa, S., Coleman, S., Di Rocco, M., Dragone, M., Gallicchio, C., Gennaro, C., Lozano, H., McGinnity, M., Micheli, A., Ray, A.K., Renteria, A., Saffiotti, A., Swords, D., Vairo, C., Vance, P.: Robotic ubiquitous cognitive ecology for smart homes. Journal of Intelligent & Robotic Systems 80(1), 57–81 (2015)
3. Amato, G., Bacciu, D., Chessa, S., Dragone, M., Gallicchio, C., Gennaro, C., Lozano, H., Micheli, A., O'Hare, G., Renteria, A., Vairo, C.: A benchmark dataset for human activity recognition and ambient assisted living. In: Ambient Intelligence-Software and Applications–7th International Symposium on Ambient Intelligence (ISAmI 2016). pp. 1–9. Springer (2016)
4. Antonelo, E., Schrauwen, B., Stroobandt, D.: Modeling multiple autonomous robot behaviors and behavior switching with a single reservoir computing network. In: IEEE International Conference on Systems, Man and Cybernetics, 2008 (SMC 2008). pp. 1843–1848. IEEE (2008)
5. Bacciu, D., Gallicchio, C., Micheli, A., Di Rocco, M., Saffiotti, A.: Learning context-aware mobile robot navigation in home environments. In: The 5th International Conference on Information, Intelligence, Systems and Applications, IISA 2014. pp. 57–62. IEEE (2014)

6. Baldi, P., Brunak, S., Frasconi, P., Soda, G., Pollastri, G.: Exploiting the past and the future in protein secondary structure prediction. Bioinformatics 15(11), 937–946 (1999)
7. Bhattacharya, S., Czejdo, B., Perez, N.: Gesture classification with machine learning using kinect sensor data. In: Third International Conference on Emerging Applications of Information Technology (EAIT). pp. 348–351. IEEE (2012)
8. Biswas, K., Basu, S.K.: Gesture recognition using microsoft kinect®. In: 5th International Conference on Automation, Robotics and Applications (ICARA). pp. 100–103. IEEE (2011)
9. Boedecker, J., Obst, O., Lizier, J., Mayer, N., Asada, M.: Information processing in echo state networks at the edge of chaos. Theory in Biosciences 131(3), 205–213 (2012)
10. Boedecker, J., Obst, O., Mayer, N., Asada, M.: Initialization and self-organized optimization of recurrent neural network connectivity. HFSP journal 3(5), 340–349 (2009)
11. Boedecker, J., Obst, O., Mayer, N., Asada, M.: Studies on reservoir initialization and dynamics shaping in echo state networks. In: Proceedings of the 18th European Symposium on Artificial Neural Networks (ESANN). pp. 227–232. d-side (2009)
12. Cook, D., Augusto, J., Jakkula, V.: Ambient intelligence: Technologies, applications, and opportunities. Pervasive and Mobile Computing 5(4), 277–298 (2009)
13. Crisostomi, E., Gallicchio, C., Micheli, A., Raugi, M., Tucci, M.: Prediction of the italian electricity price for smart grid applications. Neurocomputing 170, 286–295 (2015)
14. Dragone, M., Amato, G., Bacciu, D., Chessa, S., Coleman, S., Rocco, M.D., Gallicchio, C., Gennaro, C., Lozano, H., Maguire, L., McGinnity, M., Micheli, A., O'Hare, G., Renteria, A., Saffiotti, A., Vairo, C., Vance, P.: A cognitive robotic ecology approach to self-configuring and evolving AAL systems. Engineering Applications of Artificial Intelligence 45, 269–280 (2015)
15. Dragone, M., Gallicchio, C., Guzman, R., Micheli, A.: Rss-based robot localization in critical environments using reservoir computing. In: Proceedings of the 24th European Symposium on Artificial Neural Networks (ESANN). pp. 71–76. i6doc.com (2016)
16. Dupont, M., Marteau, P.F.: Coarse-dtw for sparse time series alignment. In: Douzal-Chouakria, A., Vilar, J., Marteau, P.F. (eds.) Advanced Analysis and Learning on Temporal Data: First ECML PKDD Workshop, AALTD 2015, Porto, Portugal, September 11, 2015, Revised Selected Papers. Lecture Notes in Computer Science, vol. 9785, pp. 157–172. Springer International Publishing (2016)
17. Gallicchio, C., Micheli, A.: Architectural and markovian factors of echo state networks. Neural Networks 24(5), 440–456 (2011)
18. Gallicchio, C., Micheli, A.: Tree echo state networks. Neurocomputing 101, 319–337 (2013)
19. Gallicchio, C., Micheli, A.: A preliminary application of echo state networks to emotion recognition. In: Proceedings of EVALITA 2014. pp. 116–119 (2014)
20. Gallicchio, C., Micheli, A., Pedrelli, L., Fortunati, L., Vozzi, F., Parodi, O.: A reservoir computing approach for balance assessment. In: Douzal-Chouakria, A., Vilar, J., Marteau, P.F. (eds.) Advanced Analysis and Learning on Temporal Data: First ECML PKDD Workshop, AALTD 2015, Porto, Portugal, September 11, 2015, Revised Selected Papers. Lecture Notes in Computer Science, vol. 9785, pp. 65–77. Springer International Publishing (2016)

21. Gallicchio, C., Micheli, A., Pedrelli, L., Vozzi, F., Parodi, O.: Preliminary experimental analysis of reservoir computing approach for balance assessment. In: Douzal-Chouakria, A., et al. (eds.) Proceedings of the 1st International Workshop on Advanced Analytics and Learning on Temporal Data (AALTD). pp. 57–62. No. 1425 in CEUR Workshop Proceedings (Sept 2015)
22. Hajnal, M.A., Lőrincz, A.: Critical Echo State Networks, pp. 658–667. Springer Berlin Heidelberg, Berlin, Heidelberg (2006)
23. Ibañez, R., Soria, A., Teyseyre, A., Campo, M.: Easy gesture recognition for kinect. Advances in Engineering Software 76, 171–180 (2014)
24. Ibañez, R., Soria, A., Teyseyre, A., Berdun, L., Campo, M.: A comparative study of machine learning techniques for gesture recognition using kinect. In: Handbook of Research on Human-Computer Interfaces, Developments, and Applications, pp. 1–22. IGI Global (2016)
25. Jaeger, H.: The "echo state" approach to analysing and training recurrent neural networks. Tech. rep., GMD - German National Research Institute for Computer Science, Tech. Rep. (2001)
26. Jaeger, H., Haas, H.: Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science 304(5667), 78–80 (2004)
27. Jaeger, H., Lukoševičius, M., Popovici, D., Siewert, U.: Optimization and applications of echo state networks with leaky-integrator neurons. Neural Networks 20(3), 335–352 (2007)
28. Kleinberger, T., Becker, M., Ras, E., Holzinger, A., Müller, P.: Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In: Douzal-Chouakria, A., Vilar, J., Marteau, P.F. (eds.) Universal Access in Human-Computer Interaction. Ambient Interaction. Lecture Notes in Computer Science, vol. 4555, pp. 103–112. Springer Berlin Heidelberg (2007)
29. Kolen, J.F., Kremer, S.C.: A field guide to dynamical recurrent networks. IEEE Press (2001)
30. Lukoševičius, M., Jaeger, H.: Reservoir computing approaches to recurrent neural network training. Computer Science Review 3(3), 127–149 (2009)
31. Marteau, P.F., Gibet, S., Reverdy, C.: Down-sampling coupled to elastic kernel machines for efficient recognition of isolated gestures. In: Proceedings of the 22nd International Conference on Pattern Recognition. pp. 363–368. IEEE Computer Society (2014)
32. Mitra, S., Acharya, T.: Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37(3), 311–324 (2007)
33. Palumbo, F., Gallicchio, C., Pucci, R., Micheli, A.: Human activity recognition using multisensor data fusion based on reservoir computing. Journal of Ambient Intelligence and Smart Environments 8(2), 87–107 (2016)
34. Patsadu, O., Nukoolkit, C., Watanapa, B.: Human gesture recognition using kinect camera. In: International Joint Conference on Computer Science and Software Engineering (JCSSE). pp. 28–32. IEEE (2012)
35. Rodan, A., Sheta, A., Faris, H.: Bidirectional reservoir networks trained using svm+ privileged information for manufacturing process modeling. Soft Computing pp. 1–14 (2016)
36. Salmen, M., Ploger, P.: Echo state networks used for motor control. In: Proceedings of the 2005 IEEE international conference on robotics and automation. pp. 1953–1958. IEEE (2005)
37. Schuster, M., Paliwal, K.: Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing 45(11), 2673–2681 (1997)

38. Skowronski, M., Harris, J.: Automatic speech recognition using a predictive echo state network classifier. Neural networks 20(3), 414–423 (2007)
39. Triefenbach, F., Jalalvand, A., Demuynck, K., Martens, J.P.: Acoustic modeling with hierarchical reservoirs. IEEE Transactions on Audio, Speech, and Language Processing 21(11), 2439–2450 (2013)
40. Verstraeten, D., Schrauwen, B., Stroobandt, D.: Reservoir-based techniques for speech recognition. In: The 2006 IEEE International Joint Conference on Neural Network Proceedings. pp. 1050–1053. IEEE (2006)