Grid Search Optimization of Novel SNN-ESN Classifier on a Supercomputer Platform

Dimitar Penkov, Petia Koprinkova-Hristova(✉), Nikola Kasabov, Simona Nedelcheva, Sofiya Ivanovska, and Svetlozar Yordanov

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia, Bulgaria
{dimitar.penkov,petia.koprinkova,nikola.kasabov, simona.nedelcheva}@iict.bas.bg, {sofia,svetlozar}@parallel.bas.bg

Abstract. This work is demonstrating the use of a supercomputer platform to optimise hyper-parameters of a proposed by the team novel SNN-ESN computational model, that combines a brain template of spiking neurons in a spiking neural network (SNN) for feature extraction and an Echo State Network (ESN) for dynamic data series classification. A case study problem and data are used to illustrate the functionalities of the SNN-ESN. The overall SNN-ESN classifier has several hyper-parameters that are subject to refinement, such as: spiking threshold, duration of the refractory period and STDP learning rate for the SNN part; reservoir size, spectral radius of the connectivity matrix and leaking rate for the ESN part. In order to find the optimal hyper-parameter values exhaustive search over all possible combinations within reasonable intervals was performed using supercomputer Avitohol. The resulted optimal parameters led to improved classification accuracy. This work demonstrates the importance of model parameter optimisation using a supercomputer platform, which improves the usability of the proposed SNN-ESN for real-time applications on complex spatio-temporal data.

Keywords: Spikking Neural Network · Echo State Network · Classification

1 Introduction

Electroencephalography (EEG) is a method to record the electrical activity of the brain. It is typically non-invasive, with the EEG electrodes placed along the

---

Supported by the HORIZON-EIC action under the project "Auto-adaptive Neuromorphic Brain Machine Interface: toward fully embedded neuroprosthetics (NEMO-BMI)", No 101070891/01.10.2022. S. Nedelcheva was partially supported by the Bulgarian Ministry of Education and Science under the National Research Programme "Young scientists and postdoctoral students-2" approved by DCM 206/07.04.2022. S. Ivanovska was partially supported by a grant by CAF America. We acknowledge the provided access to the e-infrastructure of the NCHDC - part of the Bulgarian National Roadmap on RIs, with the financial support by the Grant No D01-168/28.07.2022.

I. Lirkov and S. Margenov (Eds.): LSSC 2023, LNCS 13952, pp. 435–443, 2024.
https://doi.org/10.1007/978-3-031-56208-2_45

scalp. The measured in this way signals represent the postsynaptic potentials of pyramidal neurons in the cortex. Since the electrical activity in the brain surface originates in the deeper brain areas that do not contribute directly to an EEG recording, their influence could be assessed accounting for the electrodes orientation and distance to the source of the activity. By far EEG has been applied to numerous domains from brain-computer interface [1–3], emotion recognition [4–8], control of movements [9,10], diagnostic of brain diseases [11] etc. Through the years the EEG data processing methodology has evolved from simple methods, such as mean and amplitude comparison to complicated methods, such as connectivity topology and deep learning [12,13]. In particular, deep learning exhibits better performance in EEG classification in comparison with the conventional methods. Nevertheless, accurate on-line classification and explanation of dynamic spatio-temporal brain data, such as EEG is still an open problem. While there are many methods introduced for brain data classification, most of them lack explainability in relation to the measured brain functions as spatio-temporal patterns.

In order to exploit EEG data for analysis the first step traditionally is to "decode" them [10,14] or to extract a range of signal properties referred to as "features" which are then utilized for detection or classification purposes [15]. The analysis outcome is largely influenced by the quality of extracted features. A recent trend in EEG features extraction and processing exploits recurrent neural networks [11] and especially a member of reservoir family - fast on-line trainable Echo state networks (ESN) [6–9,16,17]. The reservoir computing advantage is in its training algorithms. They usually need a single training epoch by Least Squares (LS) algorithm or on-line training via the Recursive Least Squares (RLS) method. However, their randomly initialized pool of neurons does not account for brain structure, so reservoir computing lacks explainability. Accounting for spatial brain structure in the design of RNN decoders or feature extractor became a natural direction of work nowadays [4]. Recently developed brain-inspired spiking neural network (SNN) models, such as NeuCube [1–3,5,18], demonstrated a good classification accuracy and excellent explanation of the spatio-temporal patterns learned from spatio-temporal brain data, such as EEG and fMRI [19].

In [24] we proposed another combination between fast trainable reservoir computing and brain-inspired architectures - the integration of the SNN module with an ESN classifier, aiming at improved classification accuracy in an on-line learning mode. The overall SNN-ESN classifier has multiple hyper-parameters that are subject to refinement, such as spiking threshold, duration of the refractory period and STDP learning rate for the SNN part; reservoir size, spectral radius of the connectivity matrix, leaking rate and scalling parameters for the ESN part. Since all possible combinations within reasonable intervals of these parameters are too much, their refinement using the exhaustive search on a desktop computer would take enormous time. That is why it was performed on the supercomputer Avitohol.

Further the paper is organized as follows: Sect. 2 presents briefly NeuCube and ESN structures and the proposed hierarchical architecture SNN-ESN for

EEG classification; Sect. 3 presents the optimization procedure and the obtained best hyper-parameters values; finally the concluding remarks summarize the main achievements in the presented work and points the directions for future work.

2 SNN-ESN Architecture

2.1 SNN Structure

The NeuCube architecture is an open one, allowing for new algorithms to be explored for encoding, learning, classification, regression [21]. It consists of three parts [19]: data encoding part, where input streaming data is encoded into spike sequences using a suitable algorithm [22]; a 3D Cube structure of spiking neurons, where every neuron has a 3D spatial coordinates defined through the use of brain-template, such as Talairach or MNI. Initial connections were generated randomly based on the distances between each two neurons; SNN classifiers of evolving spiking neuron networks (eSNN) or dynamic evolving spiking neuron networks (deSNN) [23] are used to separate the outputs of NeuCube into classes. After encoding of the spatio-temporal EEG data into sequences of spikes (spike-time information), the Cube, structured according to a brain template, receives as input EEG recordings at neurons corresponding to the electrodes' positions on the scull. As a result all neurons in the Cube generate spike trains whose dynamics depends on the input signal as well as on both the connectivity within the Cube (small world connectivity at the beginning) and on the spike time dependent plasticity (STDP) of the connections (synapses). Finally, the output classifier takes the Cube spike trains as classification features. The hyper-parameters of the SNN are the parameters of neurons such as threshold of the membrane potential V_{th}, refractory period t_{ref} during with the membrane potential returns to its base value as well as STDP learning rate λ.

2.2 ESN Structure

Echo state networks (ESN) belong to a novel and rapidly developing family of reservoir computing approaches [25–27] whose aim was development of fast trainable recurrent neural network (RNN) architectures able to approximate nonlinear time series dependencies. It incorporates a pool of neurons with sigmoid activation function (usually the hyperbolic tangent) that has randomly generated recurrent connection weights. The only trainable parameters of ESN are the output weights. In case of identity output function the least squares method is applied to train the ESN in a single iteration. For the purpose of online training the recursive version of least squares (RLS) can be applied too [25]. The ESN hyper parameters that are subject of manual tuning, usually via grid search, are the reservoir size (number of neurons), reservoir connection matrix sparsity and spectral radius and leaking rate. Additionally input and output scaling could be included.

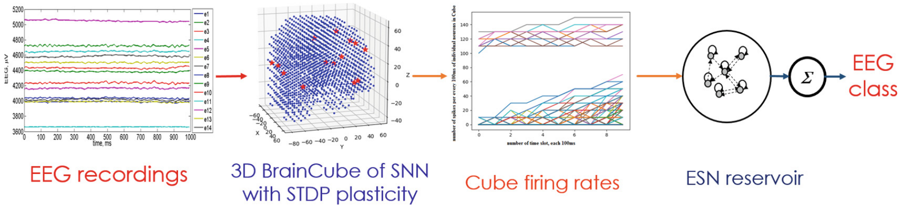

Fig. 1. Proposed NeuCube - ESN structure.

2.3 SNN-ESN Brain Data Classifier

The proposed novel brain data classifier called SNN-ESN is a hierarchical RNN composed by two recurrent architectures - a brain inspired NeuCube, that is a spatio-temporal structure of SNN neurons and a fast trainable ESN as a nonlinear time series classifier. The overall structure is shown on Fig. 1.

In contrast to NeuCube approach, here the EEG data is scaled and fed into the Cube as generating currents into neurons corresponding to the electrodes positions. A 3D SNN Cube is initialised by defining the size of the Cube of neurons and their 3D locations, including positions of EEG electordes. The Cube is designed according to the scalable Talairach atlas [28]. Small-world connectivity method is used to derive the initial connections in the Cube, where the closer two neurons are in the 3D space, the higher the probability of them to be connected is. All synapses are plastic, i.e. they adapt their weights via Spike Timing Dependent Plasticity (STDP) rule that is a kind of unsupervised learning. Thus the achieved after feeding the input signal to the Cube connectivity will reflect the EEG data characteristics. Spiking activity of all neurons was recorded and for a given time window the spiking frequency of all neurons in the Cube is calculated. Thus a sample EEG record a new time series of Cube firing rates is extracted as feature vector as shown on Fig. 1. The output classifier is an ESN reservoir. It receives generated by Cube time series feature per given EEG sample. The achieved reservoir state after presentation of each EEG sample feature vector was send to its readout and the output weights were adjusted to predict the correct EEG class. The training was done via RLS in on-line mode.

3 Grid Search Optimization

The SNN Cube is implemented in NEST Simulator, version 3.3 [29], using leaky integrate-and-fire neuron model. The ESN was implemented in Python.

The benchmark data used here is taken from [20]. The EEG data of 14 channels *Emotiv* measuring device were collected for 1000ms with sampling frequency of 128 Hz. The test subject is asked to perform three different types of wrist movement - up, down and straight - that are separated into three EEG classes and 20 examples per class are collected, making all number of samples

60. The location of the input neurons, corresponding to the used in this case 14 EEG channels, is defined following the 10–20 EEG location system. The Cube consists of 1471 neurons with initial randomly generated connectivity having 80% positive and 20% negative values.

The parallel implementation of the ESN module in Python was done using mpi4py library[1]. The NEST Simulator has its own MPI that distributes the SNN module simulation among the specified number of threads. However it does not allow to run it as a part of another parallel simulation. That is why we run both modules separately.

Two optimization experiments were performed:

1. Complete SNN-ESN model exploration running consecutively each one of them
2. Exploration of the ESN module that receives as input directly EEG data

Hyper-parameters of both modules that has to be refined and their values are given in Table 1. The number of all possible combinations of the SNN module parameters is $3 \times 6 \times 3 = 54$. The ESN module parameters yield $5 \times 5 \times 5 \times 6 \times 6 = 4500$ combinations. Thus the grid search should evaluate accuracy of totally 243000 variants of SNN-ESN model and 4500 for only ESN module respectively.

Table 1. Hyper-parameters subject to optimization

Parameter	Values
SNN	
threshold of membrane potential V_{th}	$-60, -55, -50$
refractory time t_{ref}	$0, 1, 2, 3, 4, 5$
STDP learning rate λ	$0.1, 0.01, 0.001$
ESN	
leaking rate	$0.4, 0.5, 0.6, 0.7, 0.8$
reservoir size	$3000, 3500, 4000, 4500, 5000$
reservoir sparsity	$0.4, 0.5, 0.6, 0.7, 0.8$
scale in	$0.00001, 0.0001, 0.001, 0.01, 0.1, 1$
scale out	$0.00001, 0.0001, 0.001, 0.01, 0.1, 1$

We compare time needed for grid search on an HPC facility of our institute - the supercomputer Avitohol[2] - with estimated time needed to run the same task on a desktop computer architecture. The desktop configuration has 2.60 GHz Intel(R) Core(TM) i7-6500U CPU with 2 cores and 16.0 GB RAM. The HPC System Avitohol consists of 150 servers ProLiant SL250s Gen8 each with dual Xeon CPU E5-2650 v2 at 2.60 GHz and dualXeon Phi 7120P accelerator cards.

[1] https://mpi4py.readthedocs.io/en/stable/index.html.
[2] http://www.hpc.acad.bg/system-1/.

In total it has 9600 GB RAM accessible by the regular CPUs and 4800 GB RAM on the accelerator cards. The operating system on the servers is Red Hat Enterprise Linux. The exact version on the servers that were deployed during our tests was 6.7. Table 2 summarizes the estimated simulation time of both optimization tasks on the desktop and supercomputer architectures.

Table 2. Estimated optimization times on desktop and HPC architectures in days.

Task	Module	Time	Processes	Nodes (Cores)
Desktop				
1	SNN	0.868	1	1(2)
1	ESN	112.808	2	1(2)
2	ESN	1.981	2	1(2)
Supercomputer				
1	SNN	0.622	32	1 (16)
1	ESN	8	54	4 (64)
2	ESN	0.124	30	2 (32)

It is obvious that HPC allows to decrease significantly the time needed to solve the optimization time, especially with respect to the Python module. The best SNN parameters are: threshold of membrane potential $V_{th} = -60$, refractory time $t_{ref} = 0$, STDP learning rate $\lambda = 0.001$. The obtained optimal ESN hyper-parameter values for all considered reservoir sizes and corresponding test error of the classifier are given in Table 3 for task 1 and in Table 4 for task 2 respectively.

Table 3. Minimum test error for each size of the ESN reservoir and SNN parameters $V_{th} = -60$, $t_{ref} = 0$, $\lambda = 0.001$ (task 1)

Reservoir size	Reservoir sparsity	Leaking rate	Scale in	Scale out	Test error
3000	0.7	0.4	1e−5	1e−5	1.589e−05
3500	0.7	0.6	1e−3	1e−2	1.358e−05
4000	0.6	0.4	1e−4	1e−2	1.590e−05
4500	0.7	0.4	1e−3	1e−5	1.360e−05
5000	0.6	0.4	1e−5	1e−2	1.292e−05

While for task 1 the smallest test error was obtained using the highest tested ESN reservoir size (5000 neurons), for task 2 the best results were obtained with ESN having smaller neurons number (4000). However, the test error in that case is much higher for all considered reservoir sizes. So the features extracted by the SNN module improved significantly the classification accuracy.

Table 4. Minimum test error for each size of the ESN reservoir without the SNN pre-processing (task 2)

Reservoir size	Reservoir sparsity	Leaking rate	Scale in	Scale out	Test error
3000	0.4	0.7	1e−5	1e−2	0.013
3500	0.4	0.5	1e−5	1e−1	0.015
4000	0.4	0.5	1e−5	1e−5	0.011
4500	0.7	0.5	1e−5	1e−2	0.013
5000	0.7	0.4	1e−3	1e−2	0.018

4 Conclusions

The carried out simulation investigation confirmed the significance of HPC facility in solving hard optimization tasks such as presented here grid search of hierarchical RNN model parameters. The decrease of simulation time was about 14–16 times in comparison with a given desktop configuration.

References

1. Padfield, N., Camilleri, K., Camilleri, T., Fabri, S., Bugeja, M.: A comprehensive review of endogenous EEG-based BCIs for dynamic device control. Sensors **22**(15), Article no. 5802 (2022)
2. Ieracitano, C., Mammone, N., Hussain, A., Morabito, F.C.: A novel explainable machine learning approach for EEG-based brain-computer interface systems. Neural Comput. Appl. **34**(14), 11347–11360 (2022)
3. Singanamalla, S.K.R., Lin, C.-T.: Spike-representation of EEG signals for performance enhancement of brain-computer interfaces. Front. Neurosci. **16**, Article no. 792318 (2022)
4. Zhou, J., Zhao, T., Xie, Y., Xiao, F., Sun, L.: Emotion recognition based on brain connectivity reservoir and valence lateralization for cyber-physical-social systems. Pattern Recogn. Lett. **161**, 154–160 (2022)
5. Luo, Y., et al.: EEG-based emotion classification using spiking neural networks. IEEE Access **8**, 46007–46016 (2020)
6. Fourati, R., Ammar, B., Sanchez-Medina, J., Alimi, A.M.: Unsupervised learning in reservoir computing for EEG-based emotion recognition. IEEE Trans. Affect. Comput. **13**(2), 972–984 (2022)
7. Bozhkov, L., Koprinkova-Hristova, P., Georgieva, P.: Learning to decode human emotions with Echo State Networks. Neural Netw. **78**, 112–119 (2016)
8. Bozhkov, L., Koprinkova-Hristova, P., Georgieva, P.: Reservoir computing for emotion valence discrimination from EEG signals. Neurocomputing **231**, 28–40 (2017)
9. Khan, Z.H., Hussain, N., Tiwana, M.I.: Classification of EEG signals for wrist and grip movements using echo state network. Biomed. Res. (India) **28**(3), 1095–1102 (2017)
10. Kim, H., Kim, J.S., Chung, C.K.: Identification of cerebral cortices processing acceleration, velocity, and position during directional reaching movement with deep neural network and explainable AI. NeuroImage **266**, Article no. 119783 (2023)

11. Ruffini, G., Ibañez, D., Castellano, M., Dunne, S., Soria-Frisch, A.: EEG-driven RNN classification for prognosis of neurodegeneration in at-risk patients. In: Villa, A.E.P., Masulli, P., Pons Rivero, A.J. (eds.) ICANN 2016. LNCS, vol. 9886, pp. 306–313. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44778-0_36
12. Gong, S., Xing, K., Cichocki, A., Li, J.: Deep learning in EEG: advance of the last ten-year critical period. IEEE Trans. Cogn. Dev. Syst. **14**(2), 348–365 (2022)
13. Nakagome, S., Craik, A., Ravindran, A.S., He, Y., Cruz-Garza, J.G., Contreras-Vidal, J.L.: Deep learning methods for EEG neural classification. In: Thakor, N.V. (ed.) Handbook of Neuroengineering, pp. 1–39. Springer, Singapore (2023). https://doi.org/10.1007/978-981-15-2848-4_78-1
14. Phadikar, S., Sinha, N., Ghosh, R.: Unsupervised feature extraction with autoencoders for EEG based multiclass motor imagery BCI. Expert Syst. Appl. **213**, Article no. 118901 (2023)
15. Yuvaraj, R., Thagavel, P., Thomas, J., Fogarty, J., Ali, F.: Comprehensive analysis of feature extraction methods for emotion recognition from multichannel EEG recordings. Sensors **23**(2), Article no. 915 (2023)
16. Jeong, D.-H., Jeong, J.: In-ear EEG based attention state classification using echo state network. Brain Sci. **10**(6), Article no. 321 (2020)
17. Sun, L., Jin, B., Yang, H., Tong, J., Liu, C., Xiong, H.: Unsupervised EEG feature extraction based on echo state network. Inf. Sci. **475**, 1–17 (2019)
18. Kasabov, N.K.: NeuCube: a spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. Neural Netw. **52**, 62–76 (2014)
19. Kasabov, N.K.: Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence. Springer, Cham (2019)
20. Hu, J., Hou, Z.-G., Chen, Y.-X., Kasabov, N., Scott, N.: EEG-based classification of upper-limb ADL using SNN for active robotic rehabilitation. In: 5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics, Sao Paulo, Brazil, pp. 409–414 (2014)
21. NeuCube development environment. https://kedri.aut.ac.nz/neucube
22. Petro, B., Kasabov, N., Kiss, R.: Selection and optimisation of spike encoding methods for spiking neural networks, algorithms. IEEE Trans. Neural Netw. Learn. Syst. **31**(2), 358–370 (2019)
23. Kasabov, N.K., Dhoble, K., Nuntalid, N., Indiveri, G.: Dynamic evolving spiking neural networks for online spatio- and spectro-temporal pattern recognition. Neural Netw. **41**, 188–201 (2013)
24. Koprinkova-Hristova, P., Kasabov, N., Nedelcheva, S., Yordanov, S., Penkov, D.: On-line learning, classification and interpretation of brain signals using 3D SNN and ESN. IJCNN 2023 (accepted)
25. Jaeger, H.: Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach. GMD Report 159, German National Research Center for Information Technology (2002)
26. Gallicchio, C., Lukosevicius, M., Scardapane, S.: Frontiers in reservoir computing. In: Proceedings of 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2020, Belgium, pp. 559–566 (2020)
27. Lukosevicius, M., Jaeger, H.: Reservoir computing approaches to recurrent neural network training. Comput. Sci. Rev. **3**, 127–149 (2009)
28. Talairach daemon. http://www.talairach.org
29. Spreizer, S., et al.: NEST 3.3 (3.3) (2022). Zenodo. https://doi.org/10.5281/zenodo.6368024