The Reinforced Liquid State Machine: A New Training Architecture for Spiking Neural Networks

Dominik Krenzer[1] and Martin Bogdan[1] *

1- Leipzig University - Neuromorphic Information Processing
Augustusplatz 10, Leipzig - Germany

Abstract. This work presents a novel Spiking Neural Network training architecture based on a deepened Liquid State Machine integrating Winner-Takes-All computation and Reward-Modulated Synaptic Plasticity. The networks performance is evaluated on the Heidelberg Dataset for spoken digit recognition. A two-layer liquid configuration improves classification accuracy by 5% over a single-layer baseline, while incorporating feedback between liquid layers. This architecture demonstrates that deep liquid models, combined with feedback and reward-driven learning, can effectively capture complex spatio-temporal patterns, offering significant advantages in terms of accuracy over traditional Liquid State Machines.

1 Introduction

Spiking Neural Networks (SNNs) have emerged as a promising alternative to traditional deep learning architectures, motivated by their biological plausibility and energy-efficiency[1]. Unlike conventional Artificial Neural Networks, which rely on continuous activation functions, SNNs operate using discrete, time-dependent, spike-based communication, closely mimicking the way neurons interact in the brain. This temporal coding of information allows SNNs to process data like neurological networks, making them a compelling solution for tasks requiring low latency and energy-efficient inference.

Despite the remarkable success of Deep Learning (DL), its reliance on large amounts of labeled data results in huge computational expense, as numerous floating-point operations are required during training and inference. In contrast, SNNs leverage sparse event-driven computation, where neurons only activate when a threshold is crossed. This dynamic nature allows to reduce energy consumption, especially in combination with neuromorphic hardware[2].

However, one of the major challenges in adopting SNNs is their complex training process. Traditional DL relies on gradient-based methods, such as backpropagation, to optimize the model. These methods are not directly applicable to SNNs because the computation of gradients is difficult due to the non-differentiable and discrete nature of spikes. As a result, several approaches have been proposed, such as surrogate gradient methods[3], reformulation of backpropagation through time[4] or biologically-inspired local learning rules[5].

*This work was partially funded by the German Federal Ministry of Education and Research (BMBF) within the project (ScaDS.AI) Dresden/Leipzig (BMBF grant 01IS18026B).

While these techniques have shown promise, training SNNs remains a non-trivial task due to the intricacies of temporal dynamics, precise spike-timing, and the need to balance accuracy with energy efficiency[6].

Our approach aims to bridge the gap between biologically plausible neural models and the high-performance requirements of modern Machine Learning. It builds upon the Deep Liquid State Machine[7], a type of model that consists of multiple liquid layers with intermediate unsupervised Winner-Takes-All (WTA) layers. A liquid layer is a recurrent Spiking Neural Network that processes input signals by transforming them into high-dimensional spike patterns, which can be separated linearly[8]. Subsequent to each liquid layer, a WTA layer with unsupervised learning is extracting spatio-temporal features that are passed to the next liquid.

This paper introduces a new way to train architectures like the Deep Liquid State Machine by utilizing the idea of predictive coding[9]. Predictive coding suggests that the brain continuously generates predictions about sensory input and minimizes the error between its predictions and actual sensory input. This process happens hierarchically, with predictions sent downward through the system, and prediction errors, or mismatches, propagated upward to refine future predictions. Therefore, a new presented Reward-System plays a similar role in reducing discrepancies between predictions and outcomes. When the network makes an incorrect prediction, the reward signal informs the system that an error has occurred, just as a prediction error would indicate a mismatch between the networks internal model and the actual data.

In a first step, the new architecture is explained, detailing its components and mechanisms, focusing on the training method including Reward-Modulated Spike-Time Dependent Plasticity (R-STDP)[10]. This sets the foundation for understanding how the model processes spatio-temporal data and adapts through synaptic updates. Afterwards, the architecture is evaluated using the Spiking Heidelberg Dataset (SHD)[11] for spoken word digits, allowing its performance to be assessed in a real-world scenario involving biological plausible spike-based input for multiple classes. This two-stage process demonstrates both the theoretical underpinnings and practical effectiveness of the proposed Reinforced Liquid State Machine.

2 Method

The architecture, depicted in Figure 1, features a deepened liquid structure, where two interconnected liquid layers are dynamically modulated by an intermediate WTA layer, driven by a novel introduced Reward-System including R-STDP synapses to optimize temporal pattern recognition.

2.1 Liquid Configuration

The Liquid in the following experiments consists of 135 spiking neurons with recurrent synaptic connections. The neurons are modeled as Leaky Integrate-and-Fire (LIF) neurons, as in [8], ensuring a rich spatio-temporal response to input

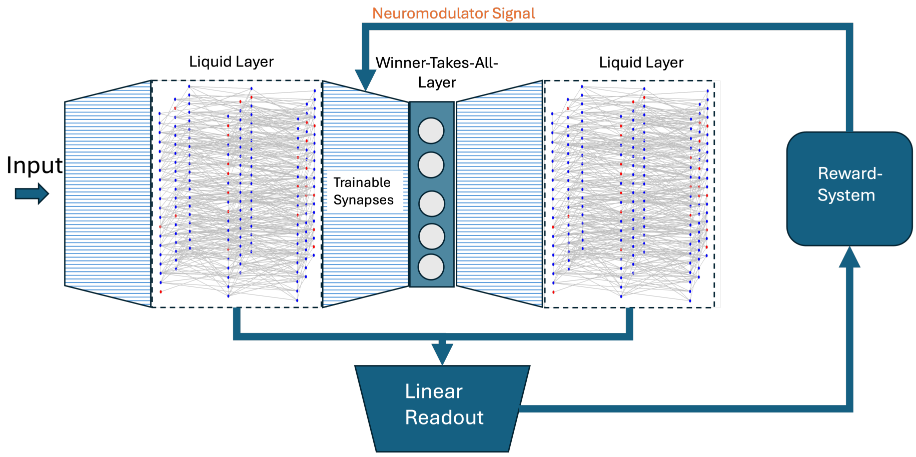

Fig. 1: The architecture of the Reinforced Liquid State Machine.

signals. The synapses in the liquid are represented by the modified stochastic synapse model (MSSM)[12] to reflect the probabilistic nature of synaptic transmission and plasticity. This probabilistic synaptic structure is allowing the liquid to explore a wide range of states and maintain robust temporal dynamics[13].

2.2 Winner-Takes-All Configuration and Reward-Modulated STDP

The WTA layer in the architecture is composed of both excitatory and inhibitory LIF neurons, structured to enforce competition among neurons and promote sparse activity. The excitatory neurons in the WTA layer receive input from the preceding liquid layer, selectively amplifying the most relevant signals. In contrast, the lateral inhibitory neurons serve to suppress less active excitatory neurons, ensuring that only a subset of the most active neurons wins and contributes to downstream processing. The synaptic connections between the liquid and the excitatory neurons of the WTA are governed by R-STDP synapses. These synapses are the central components of the architecture that are subject to training, highlighted in Figure 1 as *Trainable Synapses*. In this mechanism, synaptic weights between the liquid and excitatory neurons in the WTA layer are updated based on both the correlation between the spikes of pre- and post-synaptic neurons and a reward signal that reflects the networks performance on the task. In order to implement the local part of the learning rule in an online way, each arrival of a presynaptic spike leaves a presynaptic trace x_{pre} described by equation 1. Input spikes f are represented by the Dirac function δ.

$$\frac{dx_{pre}}{dt} = \frac{x_{pre}}{\tau_{pre}} + \sum_{f} \delta(t - t^f). \tag{1}$$

This trace reflects the activity of the presynaptic site of the synapse on a timescale set by τ_{pre}. If the postsynaptic neuron is firing, the local weight is updated by comparing x_{pre} to an activity target value x_{tar}, which determines whether potentiation or depression occurs, as defined in (2). Furthermore, a soft bound and a homeostatic rule like in [7] is added to the local learning rule to keep the synaptic weights in a certain range with w_{max} as the maximum weight, μ as control on how much the update relies on the previous weight and η as the learning rate.

$$\Delta w = \eta(x_{pre} - x_{tar})(w_{max} * w)^{\mu}. \tag{2}$$

The eligibility trace e_{trace} serves as a mechanism that links the local synaptic updates of equation 2 to the global delayed reward signals where τ_{trace} is determining the time scale of the delay. This connection allows the influence of a global reward to be distributed over time and across relevant local synapses, ensuring more efficient salient activity selection in the WTA layer. It reads

$$\frac{de_{trace}}{dt} = \frac{e_{trace}}{\tau_{trace}} + \Delta w. \tag{3}$$

Finally, the weight update, including feedback, is performed after the entire sample is processed. It is calculated as the product of the Reward-Signal r and the eligibility trace for each synapse. The Reward-Signal is generated in the Reward-System by comparing the prediction of the readout to the actual label.

$$\Delta w_{feedback} = r * e_{trace}. \tag{4}$$

These learning rules form the core of the algorithm, driving the underlying process of adaptation and optimization.

2.3 Network Dynamics

The network dynamics involve a reciprocal interaction between the WTA layer, the liquid layers, and the reward system, operating in a feed-forward and feedback loop to facilitate learning and adaptation. After processing the inputs from the initial liquid layer, the WTA layer selects the most active neurons, representing the most salient features of the input spike pattern. These selected neurons in the WTA layer then project their activity to the next liquid layer. The next liquid layer, in turn, processes this input, transforming it through its recurrent dynamics, thereby deepening the temporal representation. Once the second liquid layer has processed the signal, a readout layer interprets the spiking activity of both liquid layers. This readout layer produces the networks prediction or decision, which could be a classification or a specific output value, depending on the task. The reward system evaluates the output of the readout layer by comparing it to the correct label or target. Based on this comparison, a reward signal is generated, which is used to update the synapses in the WTA layer. This process ensures that the WTA layer improves its selection of neurons over time, optimizing the input representation fed into the next liquid layer for future predictions.

3 Results

The experimental evaluation of the Reinforced Liquid State Machine is conducted on the SHD for spoken digit recognition, comparing three different configurations within each experiment: a standard single-layer liquid, a two-layer liquid, and a two-layer liquid with reward modulation. The goal is to assess how increasing depth and incorporating the novel feedback-system affected the networks ability to capture temporal dynamics and improve classification accuracy. A total of five experiments are carried out, each of them having a different structure of connectivity in the liquids and also new initial weights. The results in Figure 2 of the experiments demonstrate a clear improvement

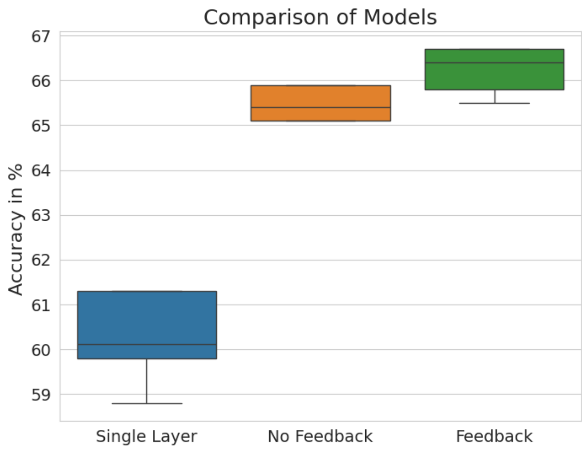

Fig. 2: Comparison of classification accuracy on the Heidelberg Dataset across three different configurations: a single-layer liquid (baseline), a two-layer liquid with unsupervised learning (No Feedback), and a two-layer liquid with the novel introduced feedback learning architecture (Feedback).

in classification accuracy with the introduction of additional liquid layers and feedback mechanisms. When a second liquid layer was added to the architecture, the unsupervised learning showed an average accuracy improvement of 5% compared to the standard single-layer configuration. This increase highlights the benefit of deeper architectures in capturing more complex temporal dependencies within the spiking patterns of the SHD. In addition, the introduction of feedback from the second liquid layer back to the WTA layer further improved network performance. This feedback mechanism enabled the network to refine ongoing dynamics within the first liquid layer in real time and reinforcing relevant spiking patterns. As a result, the new proposed architecture achieves the highest accuracy across all tested setups, surpassing the two-layer configuration without feedback. The incorporation of the new reward-modulated plasticity enhances the models ability to generalize and adapt, leading to an additional improvement in accuracy beyond the gains achieved with the second layer alone.

4 Discussion and Future Work

The results confirm that adding depth to Liquid State Machines enhances their capacity to process complex spatio-temporal patterns. The introduction of a second liquid layer yielded a 5% improvement in accuracy, underscoring the advantage of deeper architectures for capturing long-range dependencies. Furthermore, incorporating feedback between liquid layers in the proposed Reinforced Liquid State Machine is leading to the highest overall performance, highlighting the importance of feedback in refining neural representations in real time. These findings suggest that deep liquid architectures with adaptive learning mechanisms are effective for tasks requiring precise temporal recognition, offering a robust alternative to traditional spiking network models. Future work should explore the impact of adding additional liquid layers to deepen the architecture further, compared to simply increasing the size of a single liquid layer.

References

[1] Michael Pfeiffer and Thomas Pfeil. Deep learning with spiking neurons: Opportunities and challenges. *Frontiers in Neuroscience*, 12, 2018.

[2] Aaron R. Young, Mark E. Dean, James S. Plank, and Garrett S. Rose. A review of spiking neuromorphic hardware communication systems. *IEEE Access*, 7:135606–135620, 2019.

[3] Emre O. Neftci, Hesham Mostafa, and Friedemann Zenke. Surrogate gradient learning in spiking neural networks. *IEEE Signal Processing Magazine*, 2019.

[4] Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature Communications*, 11, 12 2020.

[5] Peter U. Diehl and Matthew Cook. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience*, 9, 8 2015.

[6] Martin Bogdan. *Learning algorithms for spiking neural networks: should one use learning algorithms from ANN/DL or neurological plausible learning? - A thought-provoking impulse*, pages 201–207. Servizo de Publicacions da UDC, 9 2022.

[7] Nicholas Soures and Dhireesha Kudithipudi. Deep liquid state machines with neural plasticity for video activity recognition. *Frontiers in Neuroscience*, 13, 2019.

[8] Wolfgang Maass, Thomas Natschlaeger, and Henry Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations, 2002.

[9] Rajesh Rao and Dana Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2:79–87, 02 1999.

[10] Nicolas Fremaux and Wulfram Gerstner. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits*, 9, 2015.

[11] Benjamin Cramer, Yannik Stradmann, Johannes Schemmel, and Friedemann Zenke. The heidelberg spiking data sets for the systematic evaluation of spiking neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33:2744–2757, 7 2022.

[12] Karim Ellatihy and Martin Bogdan. Enhancements on the modified stochastic synaptic model: The functional heterogeneity. *ICANN 2017*, 10613 LNCS:389–396, 2017.

[13] Karim El-Laithy and Martin Bogdan. Temporal finite-state machines: A novel framework for the general class of dynamic networks. *ICONIP 2012*.