S-TLLR: STDP-inspired Temporal Local Learning Rule for Spiking Neural Networks

Marco Apolinario, Kaushik Roy

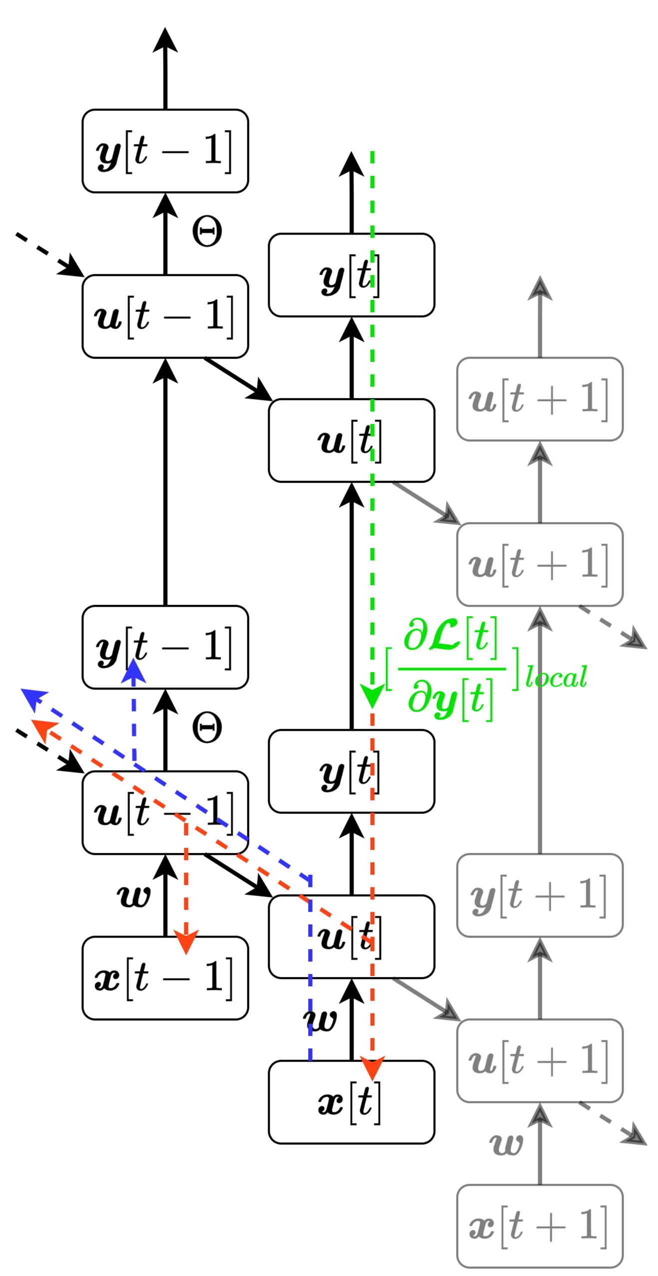
PURDUE UNIVERSITY

ICCV HONOLULU HAWAII OCT 19-23, 2025

Motivation

- Backpropagation-through-time (BPTT) assigns temporal credit but scales with timesteps T, inflating memory and compute in SNNs. We want a **temporal-local** rule that preserves rich spike-timing cues while avoiding BPTT's unfolding.

Method Overview



$$\Delta \boldsymbol{w}[t] = \left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{y}[t]}\right] \alpha_{pre} \boldsymbol{\Psi}(\boldsymbol{u}[t]) \sum_{t'=0}^{t} \lambda_{pre}^{t-t'} \boldsymbol{x}[t'] + \alpha_{post} \boldsymbol{x}[t] \sum_{t'=0}^{t-1} \lambda_{post}^{t-t'} \boldsymbol{\Psi}(\boldsymbol{u}[t'])$$

S-TLLR is a three-factor update: an instantaneous eligibility trace that blends *causal* (pre→post) and *non-causal* (post→pre) terms, modulated by a global learning signal $\delta[t]$ (via BP-through-layers or DFA). No time-unrolling or state per-synapse is kept.

Key Ideas

- **Temporal locality:** forward-only computation; time-constant memory O(n).
- **Balanced timing:** α_{pre} (causal) + α_{post} (non-causal) improves generalization.
- **Flexible δ:** BP/DFA at layer level; late δ reduces MACs with minor accuracy impact.

Experiments

S-TLLR matches or surpasses BPTT on event-vision and audio benchmarks while using *far less* memory; late δ further reduces compute.

Eligibility trace function

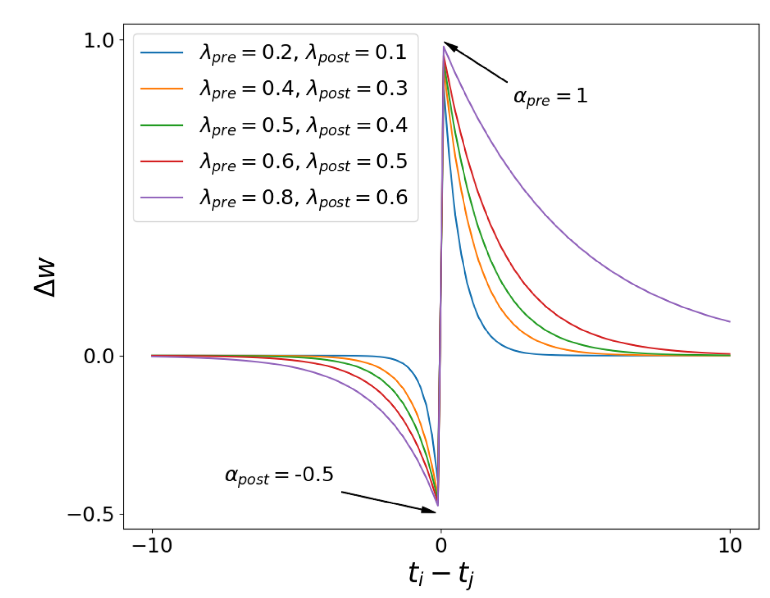


Table 1. Effects of including the non-causal terms in the eligibility traces

Dataset	Model	T	T_l	$\alpha_{post} = 0$	$\alpha_{post} = +1$	$\alpha_{post} = -1$
DVS Gesture	VGG9	20	15	$94.61 \pm 0.73\%$	$94.01 \pm 1.10\%$	$\mathbf{95.07 \pm 0.48\%}$
DVS CIFAR10	VGG9	10	5	$72.93 \pm 0.94\%$	$73.42 \pm 0.50\%$	$\mathbf{73.93 \pm 0.62\%}$
N-CALTECH101	VGG9	10	5	$62.24 \pm 1.22\%$	$53.42 \pm 1.50\%$	$\mathbf{66.33 \pm 0.86\%}$
SHD	RSNN	100	10	$77.09 \pm 0.33\%$	$\mathbf{78.23 \pm 1.84\%}$	$74.69 \pm 0.47\%$

Table 2. Accuracy performance of methods on vision and audio datasets

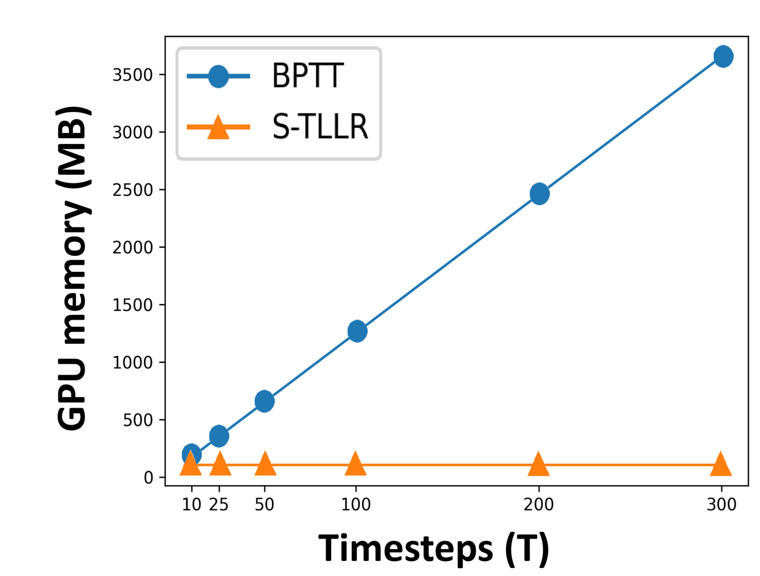
Method	Model	Accuracy (mean±std)	# MAC[1] ($\times 10^9$)	Memory[1] (MB)
DVS CIFAR10				
TET (Deng et al., 2022)	VGG-11	$83.17 \pm 0.15\%$	-	-
DSR (Meng et al., 2022)	VGG-11	$77.27 \pm 0.24\%$	-	-
BPTT (Li et al., 2021)	ResNet-18	$75.4 \pm 0.05\%$	-	-
OTTT_A(Xiao et al., 2022)	VGG-9	$76.27 \pm 0.05\%$	-	-
BPTT (baseline)	VGG-9	$75.44 \pm 0.76\%$	6.82	18.12
S-TLLR (Ours, $T_l = 5$, $\alpha_{post} = -1$)	VGG-9	$73.93 \pm 0.62\%$	5.12	3.62
S-TLLR (Ours, $T_l = 0$, $\alpha_{post} = -1$)	VGG-9	$75.6 \pm 0.10\%$	10.26	3.62
S-TLLR (Ours, $T_l = 0$, $\alpha_{post} = 0$)	VGG-9	$74.8 \pm 0.15\%$	6.82	3.62
BPTT (baseline)	ResNet18	$72.68 \pm 0.87\%$	7.13	28.14
S-TLLR (Ours, $T_l = 5$)	ResNet18	$71.94 \pm 0.75\%$	5.12	5.62
S-TLLR (Ours, $T_l = 0$)	ResNet18	$74.5 \pm 0.64\%$	10.24	5.62
DVS Gesture				
SLAYER (Shrestha & Orchard, 2018)	SNN (8 layers)	$93.64 \pm 0.49\%$	-	-
DECOLLE(Kaiser et al., 2020)	SNN (4 layers)	$95.54 \pm 0.16\%$	-	-
OTTT_A(Xiao et al., 2022)	VGG-9	96.88%	-	-
BPTT (baseline)	VGG-9	$95.58 \pm 1.08\%$	6.06	16.13
S-TLLR (Ours)	VGG-9	$97.72 \pm 0.38\%$	2.27	1.61
BPTT (baseline)	ResNet18	$94.92 \pm 0.38\%$	6.34	25.03
S-TLLR (Ours)	ResNet18	$94.92 \pm 0.61\%$	2.27	2.50
N-CALTECH101				
BPTT (She et al., 2022)	SNN (12 layers)	71.2%	-	-
BPTT (Kim et al., 2023)	VGG-16	64.40%	-	-
BPTT (baseline)	VGG-9	$65.92 \pm 0.82\%$	22.81	20.15
S-TLLR (Ours)	VGG-9	$66.058 \pm 0.92\%$	17.22	4.03
BPTT (baseline)	ResNet18	$60.89 \pm 0.89\%$	6.34	31.31
S-TLLR (Ours)	ResNet18	$61.65 \pm 0.99\%$	4.27	6.26
SHD				
ETLP (Quintana et al., 2023)	ALIF-RSNN	$74.59 \pm 0.44\%$	-	-
OSTTP (Ortner et al., 2023)	LIF-RSNN	$77.33 \pm 0.8\%$	-	-
BPTT (Bouanane et al., 2022)	LIF-RSNN	83.41	-	-
BPTT (Cramer et al., 2022)	LIF-RSNN	83.2 ± 1.3	-	-
BPTT (baseline)	LIF-RSNN	70.57 ± 0.96	0.054	0.961
S-TLLR_BP (Ours)	LIF-RSNN	$78.24 \pm 1.84\%$	0.096	0.019
S-TLLR_DFA (Ours)	LIF-RSNN	$74.60 \pm 0.52\%$	0.096	0.019

Advantages Over Existing Techniques

- **Time-local updates**: no gradient backprop through time.
- **Lower Memory**: O(n) vs O(T·n) activations for BPTT.
- **Task-adaptive timing**: tune α_{post} sign for spatial vs temporal tasks.

Table 3. Comparison with learning methods

Method	Memory Complexity	Time Complexity	Temporal Local	Leverage Non-Causality
BPTT	Tn	Tn^2	X	X
RTRL	n^3	n^4	✓	X
e-prop	n^2	n^2	✓	X
OSTL	n^2	n^2	✓	X
ETLP	n^2	n^2	✓	X
OSTTP	n^2	n^2	✓	X
OTTT	n	n^2	✓	X
S-TLLR	n	n^2	✓	✓



Temporal-local rule → constant memory footprint, enabling long-sequence training on resource-limited hardware.

Takeaways

- Competitive accuracy with markedly reduced memory (up to 10 ×).
- Late learning signal yields 1.3– 6.6 × fewer MACs.
- Simple to implement; works for feedforward, CNN, and recurrent SNNs.

Acknowledgment

This work was supported in part by the Center for CoDesign of Cognitive Systems (CoCoSys), one of seven centers in JUMP 2.0, funded by the Semiconductor Research Corporation (SRC) and DARPA, in part by the MicroE4AI program of IARPA, the DoE Microelectronics program, the National Science Foundation, and Intel Corporation.