NEURAL INFORMATION PROCESSING SYSTEMS

Distributed-Order Fractional Spiking Neural Network

Chengjie Ge, Yufeng Peng, Qiyu Kang, Xueyang Fu, Zheng-Jun Zha*

University of Science and Technology of China

github.com/cjge-ustc/DF-SNN-official

1. Motivation: heterogeneous temporal memory

Conventional LIF neurons have exponential memory. Single-order f-LIF adds long-range history with one power-law decay. Biological neurons integrate information across heterogeneous time scales.

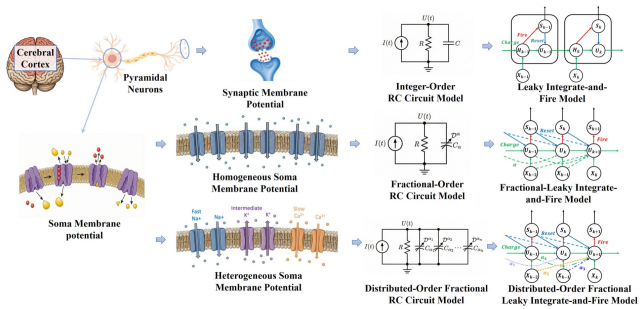


Figure 1. Biological motivation, circuit abstractions and spiking neurons.

2. Distributed-order neuron model

Core idea: integrate multiple fractional orders through a learnable non-negative measure.

DF-LIF $$\tau \int_a^b D^\alpha U(t)\,\mathrm{d}\mu(\alpha) = -U(t) + R\,I_{\mathrm{in}}(t).$$

A point mass at order 1 recovers LIF. A point mass at a single fractional order recovers f-LIF.

Same input, different membrane histories

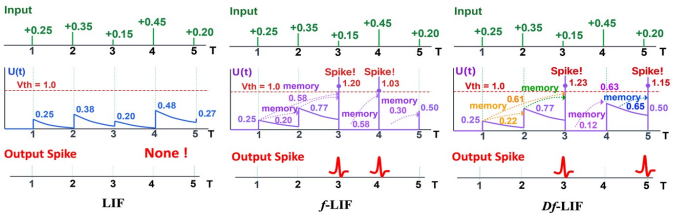


Figure 2. Illustration of LIF, *f*-LIF and *DF*-LIF.

3. Trainable GL implementation

Choose discrete fractional orders and learn their mixing weights. The Grünwald–Letnikov (GL) scheme aggregates all orders into one history-dependent kernel.

GL kernel $$\tilde{c}_m = \sum_{j=0}^{n} w_j h^{-\alpha_j} c_m^{(\alpha_j)}, \qquad c_m^{(\alpha_j)} = (-1)^m \binom{\alpha_j}{m}$$

Softplus learns non-negative weights. Softmax also normalizes their sum to one. The kernel includes the time-step factor h to the power −α for each order.

Charge update $$U_k = \frac{1}{\tilde{c}_0}\left(\frac{-V_{k-1} + X_k}{\tau} - \sum_{m=1}^{k-1} \tilde{c}_m V_{k-m}\right),$$

U is the charged potential and V is the post-reset potential. Each update combines past V states. All experiments use full history, with standard spikes, resets and surrogate gradients.

4. Theoretical analysis

Proposition 1. Multi-scale asymptotic decay
For h > 0 and positive weights on at least two distinct orders in (0, 1), the leading asymptotic kernel has a lag-dependent log-slope. The lowest order dominates its tail.

Proposition 2. Irreducibility
A mixture of at least two distinct fractional orders with non-zero weights cannot reduce to a scaled single-order kernel.

5. Experiments

Neuromorphic classification (Table 1)

Accuracy (%), mean ± std over three seeds All neuron models use the same training settings

CNN128 (DVSNet, 128 channels)

Dataset	T	LIF (SJ)	LIF (snnTorch)	f-LIF	DF-LIF
N-MNIST	16	99.09 ± 0.16	98.58 ± 0.47	99.34 ± 0.21	**99.62 ± 0.09**
DVS-Lip	16	41.58 ± 0.81	32.85 ± 1.75	43.23 ± 0.36	**45.28 ± 1.01**
DVS128Gesture	16	93.40 ± 0.35	90.31 ± 1.34	94.22 ± 0.53	**95.14 ± 0.70**
N-Caltech101	16	66.38 ± 0.53	65.63 ± 1.20	70.76 ± 0.49	**71.66 ± 0.15**
HarDVS	8	45.80 ± 0.44	46.86 ± 0.62	47.37 ± 0.48	**48.44 ± 0.37**

Spikformer256 (embedding dimension 256)

Dataset	T	LIF (SJ)	LIF (snnTorch)	f-LIF	DF-LIF
DVS128Gesture	16	94.10 ± 0.35	88.81 ± 1.57	94.56 ± 0.53	**95.95 ± 0.53**
N-Caltech101	16	71.86 ± 0.75	65.77 ± 0.38	74.31 ± 1.05	**77.82 ± 0.56**
HarDVS	8	45.59 ± 0.47	46.44 ± 0.28	47.18 ± 0.28	**48.49 ± 0.26**

DF-LIF achieves the best accuracy in all eight settings.

Accuracy across model scales (Table 2)

N-Caltech101, Spikformer, T = 16
Accuracy (%), mean ± sample std over five seeds

Neuron	256 (2.6M)	384 (5.8M)	512 (10.3M)
LIF	68.10 ± 3.00	72.12 ± 1.08	73.14 ± 1.33
f-LIF	73.52 ± 1.89	74.38 ± 1.17	75.78 ± 0.84
DF (Softplus)	**76.88 ± 0.30**	78.41 ± 0.53	**80.34 ± 0.68**
DF (Softmax)	76.86 ± 0.36	**78.69 ± 0.44**	79.77 ± 0.25

+3.36 to +4.56 accuracy points

Event-based object detection (Table 4)

Gen1, mAP@0.5:0.95 (%), single runs, T = 5

Backbone	LIF	f-LIF	DF-LIF
EMS-Res10	26.7	27.6	**29.4**
EMS-Res34	31.0	32.8	**34.7**

+2.7 / +3.7 points over LIF with the same backbones.

Transfer to richer neurons (Table 7)

N-Caltech101, Spikformer-256, T = 16, five seeds
Accuracy (%) and gain over the fractional variant

Base neuron	Fractional	DF (Softplus)	Gain
LIF	73.52 ± 1.89	**76.88 ± 0.30**	+3.36
P-LIF	72.65 ± 0.56	**76.42 ± 0.49**	+3.77
TC-LIF	57.23 ± 0.65	**66.99 ± 0.38**	+9.76
TS-LIF	67.38 ± 1.63	**71.25 ± 0.61**	+3.87

DF gains 3.36–9.76 points across the four base neurons. Test throughput drops by 0.7–4.3% vs. fractional variants. Learning weights adds training compute and memory (Tables 6–7).