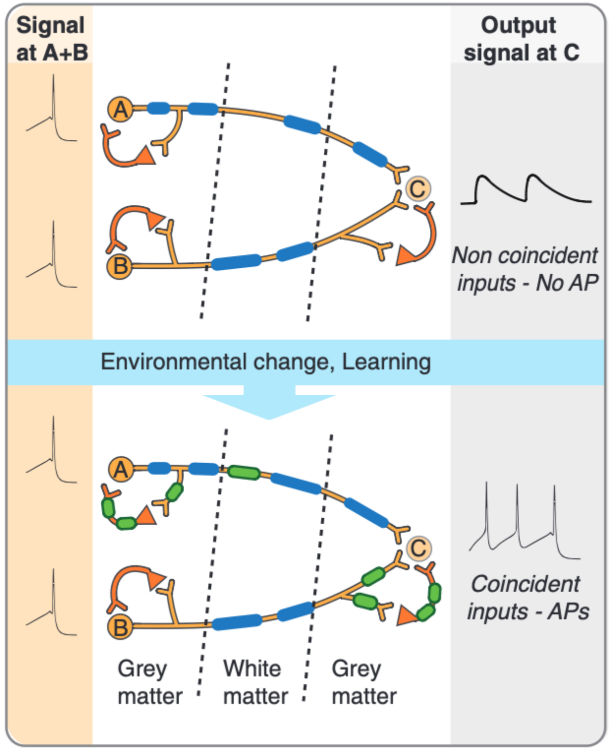
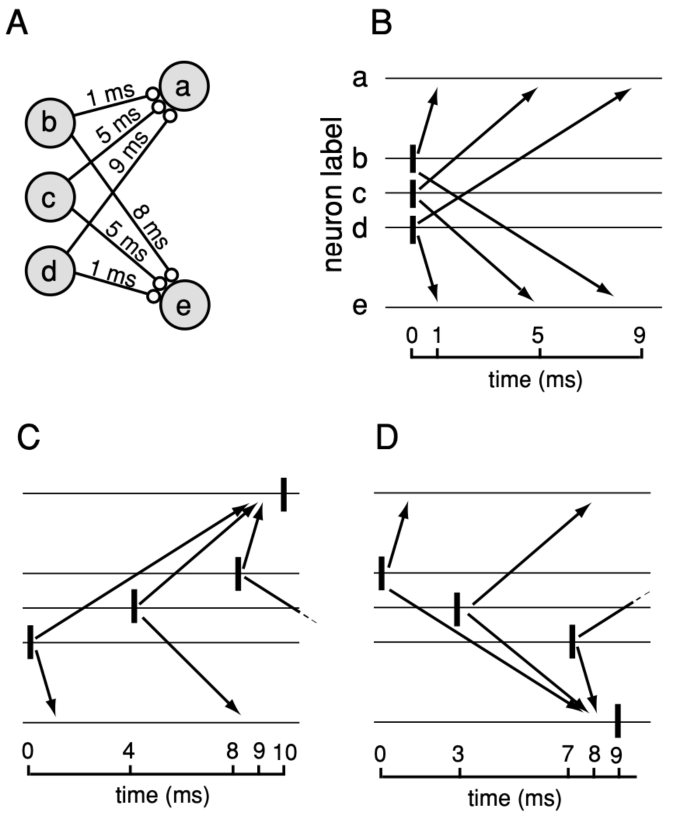
Efficient Event-based Delay Learning in Spiking Neural Networks

Balázs Mészáros, James Knight and Thomas Nowotny

Neuro-Inspired Computational Elements NICE
GeNN
The Alan Turing Institute
LEVERHULME TRUST
Sussex AI
US UNIVERSITY OF SUSSEX



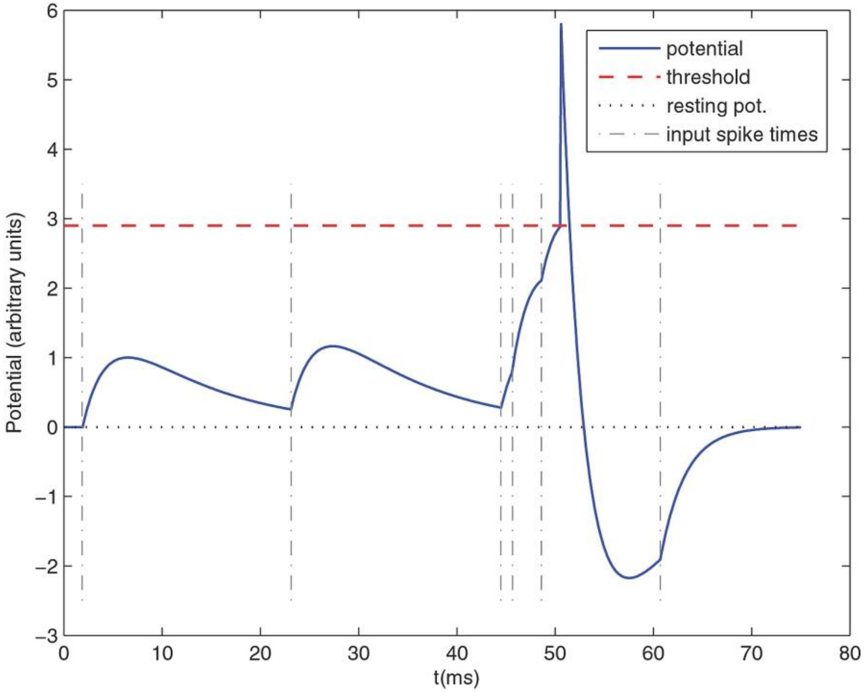


Dataset	Method	Rec.	Delays	#Params	Top1 Acc.
SHD	EventProp-GeNN (Nowotny et al., 2022)	✓	✗	N/a	84.80±1.5%
	Cuba-LIF (Dampfhoffer et al., 2022)	✓	✗	0.14M	87.80±1.1%
	Adaptive SRNN (Yin et al., 2021)	✓	✗	N/a	90.40%
	SNN+Delays (Patiño-Saucedo et al., 2023)	✗	✓	0.1M	90.43%
	TA-SNN (Yao et al., 2021)	✗	✗	N/a	91.08%
	STSC-SNN (Yu et al., 2022)	✗	✗	2.1M	92.36%
	Adaptive Delays (Sun et al., 2023b)	✗	✓	0.1M	92.45%
	DL128-SNN-Dloss (Sun et al., 2023a)	✗	✓	0.14M	92.56%
	Dense Conv Delays (ours)	✗	✓	2.7M	93.44%
	RadLIF (Bittar & Garner, 2022)	✓	✗	3.9M	94.62%
	DCLS-Delays (2L-1KC)	✗	✓	**0.2M**	**95.07±0.24%**
SSC	Recurrent SNN (Cramer et al., 2022)	✓	✗	N/a	50.90 ± 1.1%
	Heter. RSNN (Perez-Nieves et al., 2021)	✓	✗	N/a	57.30%
	SNN-CNN (Sadovsky et al., 2023)	✗	✓	N/a	72.03%
	Adaptive SRNN (Yin et al., 2021)	✓	✗	N/a	74.20%
	SpikGRU (Dampfhoffer et al., 2022)	✓	✗	0.28M	77.00±0.4%
	RadLIF (Bittar & Garner, 2022)	✓	✗	3.9M	77.40%
	Dense Conv Delays 2L (ours)	✗	✓	10.9M	77.86%
	Dense Conv Delays 3L (ours)	✗	✓	19M	78.44%
	DCLS-Delays (2L-1KC)	✗	✓	**0.7M**	**79.77±0.09%**
	DCLS-Delays (2L-2KC)	✗	✓	**1.4M**	**80.16±0.09%**
	DCLS-Delays (3L-1KC)	✗	✓	**1.2M**	**80.29±0.06%**
	DCLS-Delays (3L-2KC)	✗	✓	**2.5M**	**80.69±0.21%**
GSC-35	MSAT (He et al., 2023)	✗	✗	N/a	87.33%
	Dense Conv Delays 2L (ours)	✗	✓	10.9M	92.97%
	Dense Conv Delays 3L (ours)	✗	✓	19M	93.19%
	RadLIF (Bittar & Garner, 2022)	✓	✗	1.2M	94.51%
	DCLS-Delays (2L-1KC)	✗	✓	**0.7M**	**94.91±0.09%**
	DCLS-Delays (2L-2KC)	✗	✓	**1.4M**	**95.00±0.06%**
	DCLS-Delays (3L-1KC)	✗	✓	**1.2M**	**95.29±0.11%**
	DCLS-Delays (3L-2KC)	✗	✓	**2.5M**	**95.35±0.04%**

Bonetto, G., Belin, D., & Káradóttir, R. T. (2021). Myelin: A gatekeeper of activity-dependent circuit plasticity?. *Science*, *374*(6569), eaba6905.
Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural computation*, *18*(2), 245-282.
Hammouamri, I., Khalfaoui-Hassani, I., & Masquelier, T. Learning Delays in Spiking Neural Networks using Dilated Convolutions with Learnable Spacings. In *The Twelfth International Conference on Learning Representations*.

Leaky Integrate-and-Fire neuron

Free dynamics	Transition condition	Jumps at transition
$\tau_{\text{mem}} \frac{\mathrm{d}}{\mathrm{d}t} V = -V + I$ $\tau_{\text{syn}} \frac{\mathrm{d}}{\mathrm{d}t} I = -I$	$(V)_n - \vartheta = 0$ $(\dot{V})_n \neq 0$ for any n	$(V^+)_n = 0$ $I^+ = I^- + W e_n$



Backward propagation

Free dynamics	Transition condition	Jump at transition
$\tau_{\mathrm{mem}} \lambda'_V = -\lambda_V - \frac{\partial l_V}{\partial V}$ $\tau_{\mathrm{syn}} \lambda'_I = -\lambda_I + \lambda_V$	$t - t_k^{\mathrm{post}} = 0$ for any k	$(\lambda_V^-)_{n(k)} = (\lambda_V^+)_{n(k)} + \frac{1}{\tau_{\mathrm{mem}} (\dot{V}^-)_{n(k)}} \Big[\vartheta (\lambda_V^+)_{n(k)} + \big(W^\top (\lambda_V^+ - \lambda_I)\big)_{n(k)} + \frac{\partial l_{\mathrm{p}}}{\partial t_k^{\mathrm{post}}} + l_V^- - l_V^+ \Big]$

Gradient updates: $$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}w_{ji}} = -\tau_{\mathrm{syn}} \sum_{\text{spikes from } i} (\lambda_I)_j$$

Wunderlich, T. C., & Pehle, C. (2021). Event-based backpropagation can compute exact gradients for spiking neural networks. *Scientific Reports*, *11*(1), 12829.

Synaptic delays

Free dynamics	Transition condition	Jumps at transition
$\tau_{\mathrm{mem}} \frac{\mathrm{d}}{\mathrm{d}t} V = -V + I$ $\tau_{\mathrm{syn}} \frac{\mathrm{d}}{\mathrm{d}t} I = -I$	$(V)_n - \vartheta = 0$ $(\dot{V})_n \neq 0$ for any n	$(V^+)_n = 0$ $I^+ = I^- + We_n$ ← Delay here

Fixed and learnable synaptic delays

- EventProp updates weights only at spike time:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}w_{ji}} = -\tau_{\mathrm{syn}} \sum_{\text{spikes from } i} (\lambda_I)_j$$

- That doesn't change when delays are included:

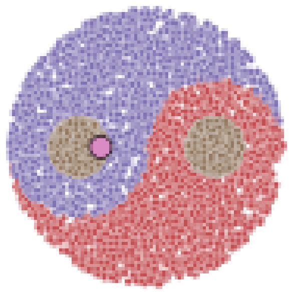
$$\frac{d\mathcal{L}}{dw_{ji}} = -\tau_{syn} \sum_{t_s \in \substack{\text{spikes} \\ \text{from } i}} (\lambda_I)_j \Big|_{t_s + d_{ji}}$$

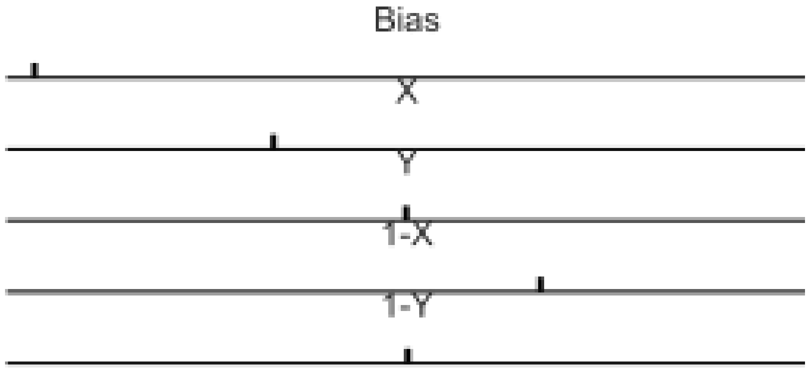
- And delay updates also only happen at spike times

$$\frac{d\mathcal{L}}{dd_{ji}} = -w_{ji} \sum_{t_s \in \substack{\text{spikes} \\ \text{from } i}} (\lambda_I - \lambda_V)_j \Big|_{t_k + d_{ji}}$$

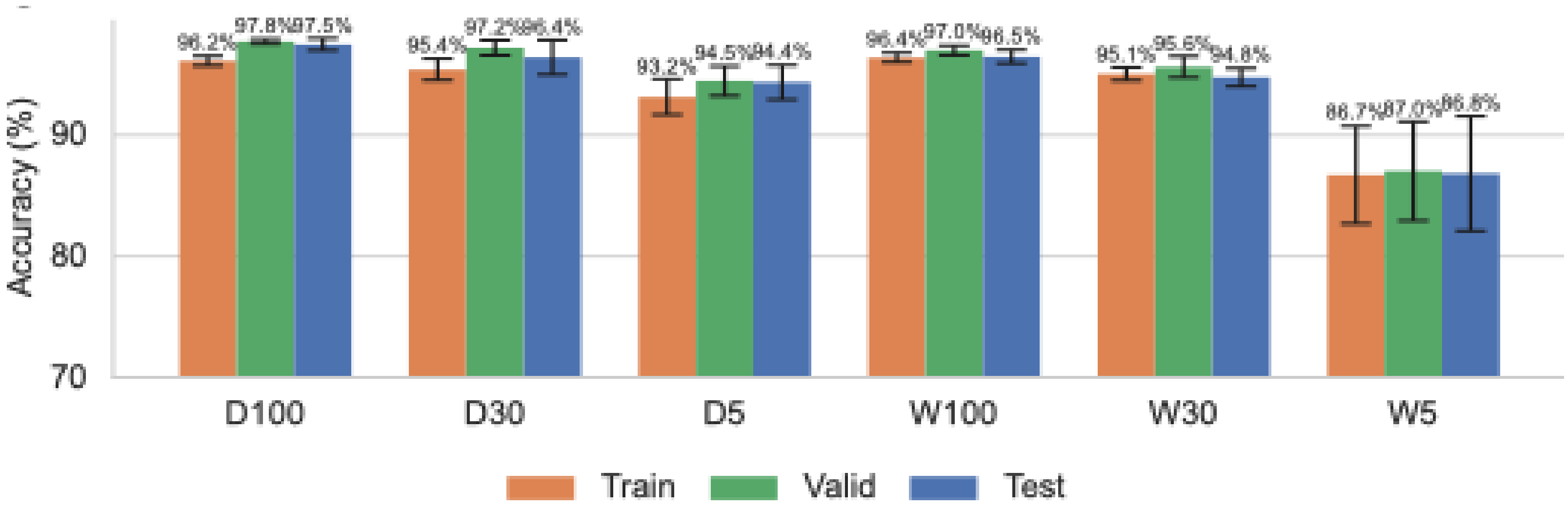
Mészáros, B., Knight, J. C., & Nowotny, T. (2025). Efficient Event-based Delay Learning in Spiking Neural Networks. *arXiv preprint arXiv:2501.07331*.

Yin-Yang



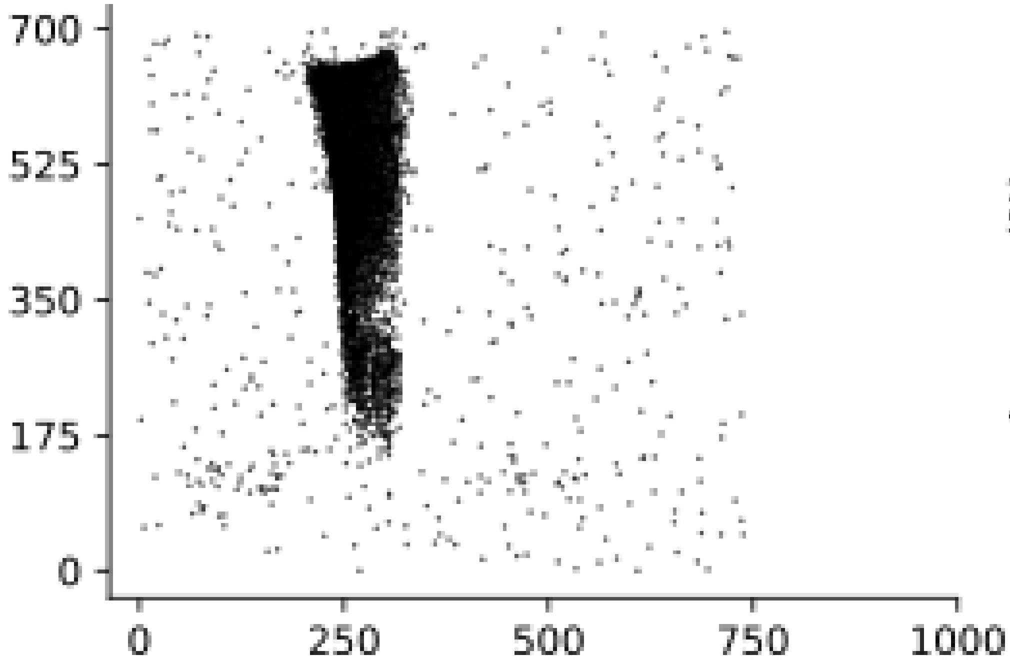


Bias
X
Y
1-X
1-Y

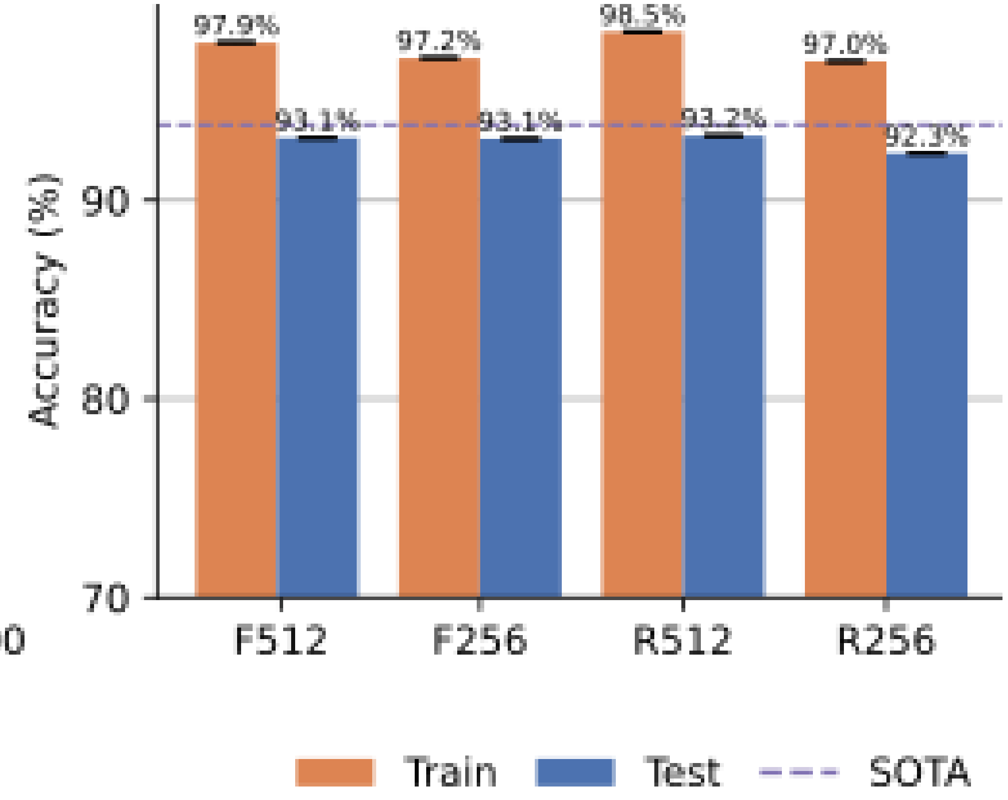


Accuracy (%)
70
80
90
96.2%
97.8%
97.5%
95.4%
97.2%
96.4%
93.2%
94.5%
94.4%
96.4%
97.0%
96.5%
95.1%
95.6%
94.8%
86.7%
87.0%
86.8%
D100
D30
D5
W100
W30
W5
Train
Valid
Test

Spiking Heidelberg Digits

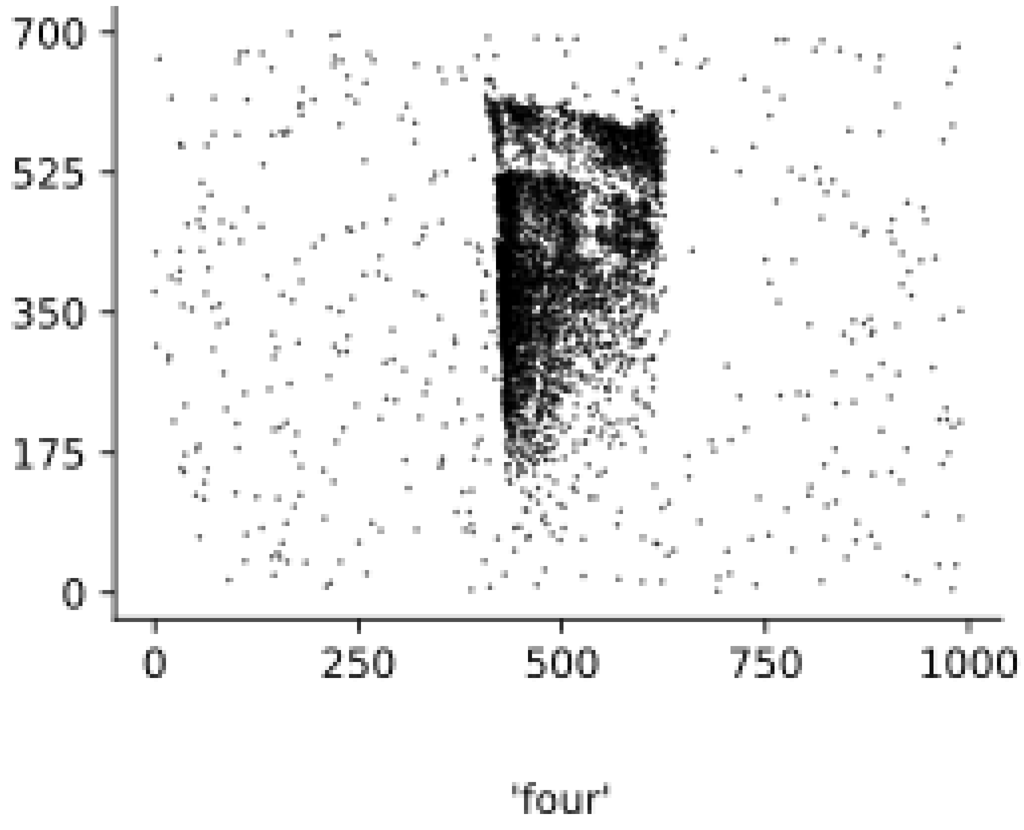


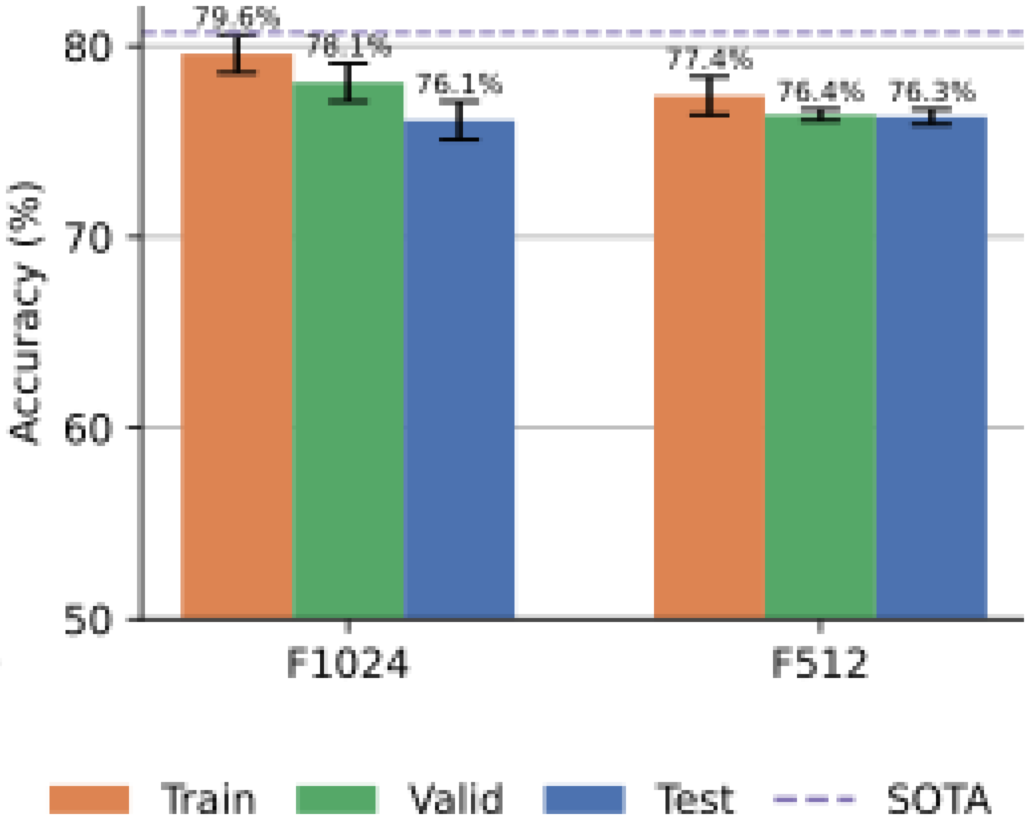
'six'



Maximilian Baronig, Romain Ferrand, Silvester Sabathiel, and Robert Legenstein. "Advancing Spatio- Temporal Processing in Spiking Neural Networks through Adaptation". In: arXiv preprint arXiv:2408.07517 (2024).

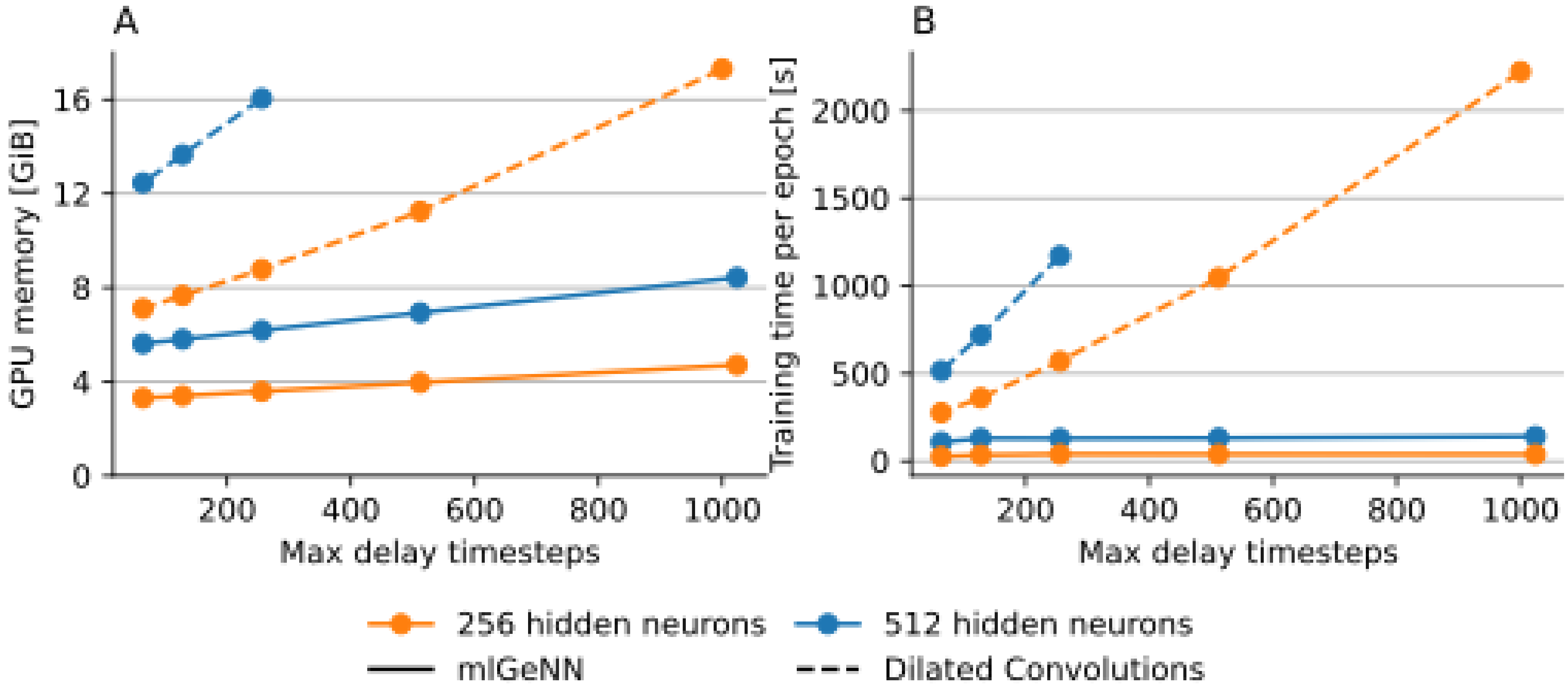
Spiking Speech Commands





Hammouamri, I., Khalfaoui-Hassani, I., & Masquelier, T. Learning Delays in Spiking Neural Networks using Dilated Convolutions with Learnable Spacings. In *The Twelfth International Conference on Learning Representations*.

Benchmarking



Hammouamri, I., Khalfaoui-Hassani, I., & Masquelier, T. Learning Delays in Spiking Neural Networks using Dilated Convolutions with Learnable Spacings. In *The Twelfth International Conference on Learning Representations.*