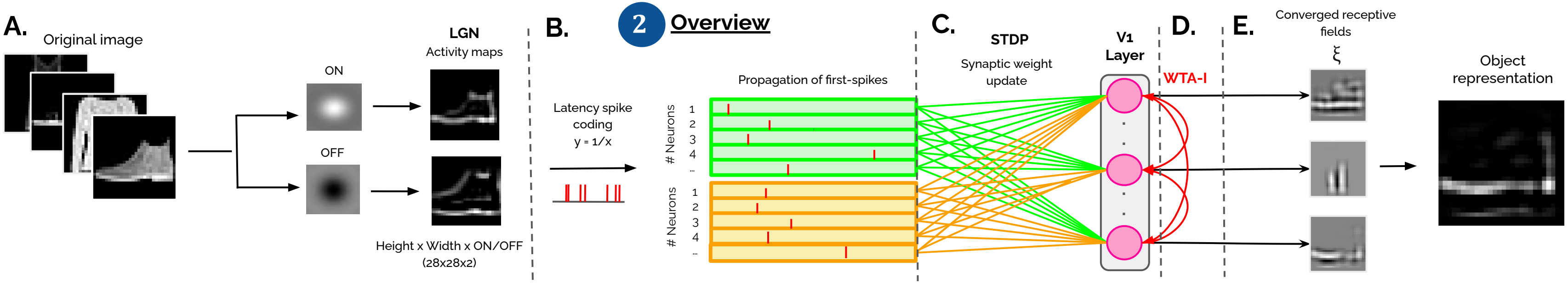
Efficient visual object representation using a biologically plausible spike-latency code and winner-take-all inhibition

Melani Sanchez-Garcia, Tushar Chauhan, Benoit R. Cottereau, Michael Beyeler

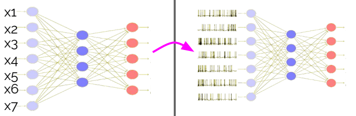
COMPUTER SCIENCE UC SANTA BARBARA · UCSB · BIONIC VISION LAB · CerCo UMR5549 · Université de Toulouse

Contribution: We present a Spiking Neural Network (SNN) model that uses spike-latency encoding and winner-take-all inhibition (WTA-I) to efficiently represent stimuli from the Fashion MNIST dataset.

2 Overview



1 Motivation



Deep Convolutional Neural Networks:
- Large amount of energy
- Computation
- Memory

Spiking Neural Networks:
- Low power consumption
- Fast inference
- Event-driven processing
- Asynchronous operation

Scan the **QR Code** to see our paper!

3 Network Architecture

A. Modeling the retinotopic pathway

- The **LGN layer** consisted of simulated firing-rate neurons with center-surround receptive fields, implemented using a direct application of a 6x6 difference of Gaussian filter on the image.

B. Spike-latency code

- We converted the LGN activity maps to **first spike relative latencies**.
- The LGN spikes contributed to an increase in the membrane potential of V1 neurons, until one of the **V1 membrane potentials** reached threshold.

$$E_n(t) = \begin{cases} \sum_{m \in LGN} w_{mn} \cdot H(t - t_m), & t < \min_t \left\{ t \mid \max_{n \in V1} E_n(t) \geq \theta \right\} \\ 0, & \text{otherwise.} \end{cases}$$

C. Spike-timing dependent plasticity (STDP)

- The weights of plastic synapses connecting LGN and V1 were updated using STDP, as a function of the relative timing of pre- and postsynaptic spikes: Long-term potentiation **(LTP) ($\Delta t > 0$)** and long-term depression **(LTD) ($\Delta t \leq 0$)**.

$$\Delta w = \begin{cases} -\alpha^- \cdot w^{\mu^-} \cdot K(\Delta t, \tau_-), & \Delta t \leq 0 \\ \alpha^+ \cdot (1 - w)^{\mu^+} \cdot K(\Delta t, \tau_+), & \Delta t > 0, \end{cases}$$

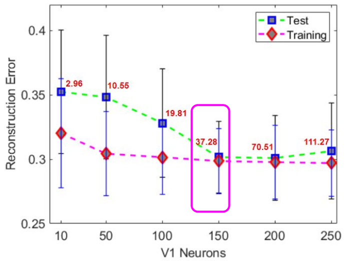
D. Winner-take-all inhibition

- We used a hard **WTA-I** scheme such that, if any V1 neuron fired during a certain iteration, it simultaneously prevented other neurons from firing until the next sample.

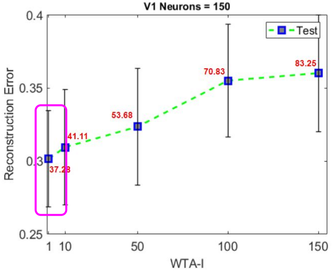
$$y_i = \begin{cases} 1, & \text{if } x_i > x_j \text{ for all } j \neq i \\ 0, & \text{if } x_j > x_i \text{ for some } j \neq i\ . \end{cases}$$

4 Object Representation and Results

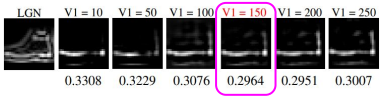
The reconstruction error of the test set went through a minimum (at roughly **150 V1 neurons**).



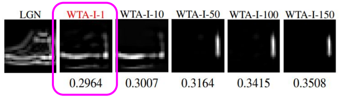
One or more than one neuron was active for each sample.



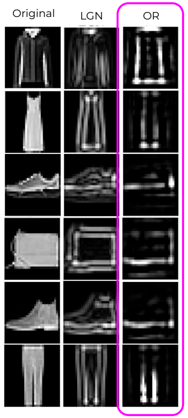
Every spiking neuron became **selective for a particular feature** allowing a rich object representation.



Multiple neurons ended up **learning similar visual features**, which resulted in poor object reconstructions.



E. Stimulus reconstruction



$$OR_k = \sum_{j \in V1} r_{kj} \xi_j ,$$

5 Conclusions

- Efficient object representations can be learned with **STDP** rule .
- Our network is able to represent objects with as little as **150 spiking neurons** and at most **40 spikes**.

6 Related work

[1] M Beyeler, EL Rounds, KD Carlson, N Dutt, and JL Krichmar. Neural correlates of sparse coding and dimensionality reduction. PLoS Computational Biology, 2019.

[2] T Chauhan, T Masquelier, A Montlibert, and BR Cottereau. Emergence of binocular disparity selectivity through Hebbian learning. Journal of Neuroscience, 2018.