A robust bio-inspired approach to event-driven object recognition

Antoine Grimaldi[1], Victor Boutin[1], Sio-Hoi Ieng[2], Laurent Perrinet[1] & Ryad Benosman[2]

[1] Institut de Neurosciences de la Timone (UMR 7289); Aix Marseille Univ, CNRS; Marseille, France

[2] Sorbonne Université, INSERM, CNRS, Institut de la Vision, Paris, France

antoine.grimaldi@univ-amu.fr

Abstract

We propose a neuromimetic architecture able to perform pattern recognition. To achieve this, we extended the existing event-based algorithm from [1] which introduced novel spatio-temporal features: time surfaces. Built from asynchronous events acquired by a neuromorphic camera, these time surfaces allow to code the local dynamics of a visual scene and to create an efficient hierarchical event-based pattern recognition architecture. Inspired by biological findings and the efficient coding hypothesis, our main contribution is to integrate homeostatic regulation to the Hebbian learning rule. Indeed, in order to be optimally informative, average neural activity within a layer should be equally balanced across neurons. We used that principle to regularize neurons within the same layer by setting a gain depending on their past activity and such that they emit spikes with balanced firing rates. The efficiency of this technique was first demonstrated through a robust improvement in spatio-temporal patterns which were learned during the training phase. We validated classification performance with the widely used N-MNIST dataset [2] reaching 87% accuracy with homeostasis compared to 70% accuracy without homeostasis. Finally, studying the impact of input jitter on classification highlights resilience of this method. We expect to extend this fully event-driven approach to more naturalistic tasks, notably for ultra-fast object categorization.

Acknowledgments

Antoine Grimaldi and Laurent Perrinet received funding from the European Union ERA-NET CHIST-ERA 2018 research and innovation program under grant agreement No ANR-19-CHR3-0008-03.

Event-Based Signal

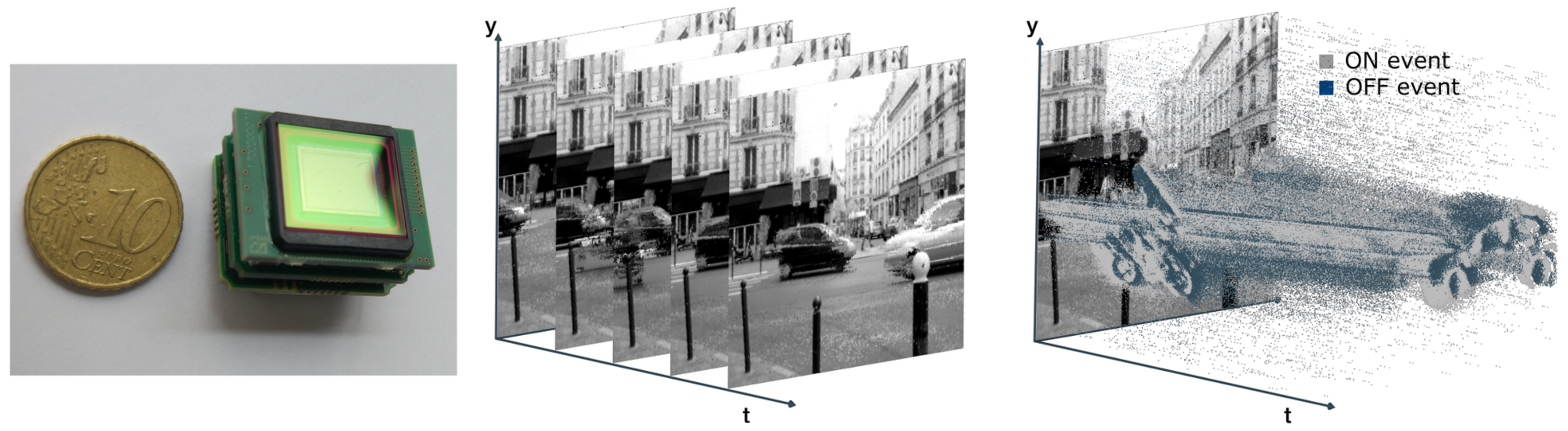

Figure 1: A miniature event-based ATIS sensor (left) which, compared to classical frame-based representations (middle), outputs an event-based representation of the scene (right).

Some advantages of a neuromorphic event-based camera (Dynamic Vision Sensor):

- high temporal resolution
- reduction of redundancy
- energy efficiency
- high dynamic range

Events are recorded asynchronously on the pixel grid when observing a significant change in brightness. An event is defined as $\delta = (a, t)$ where t is the time of the event, $a = (x, y, \mathbf{p})$ where x and y define its position and $\mathbf{p}$ its polarity. Polarities can be ON or OFF to represent the sign of brightness change.

HOTS network from [1]

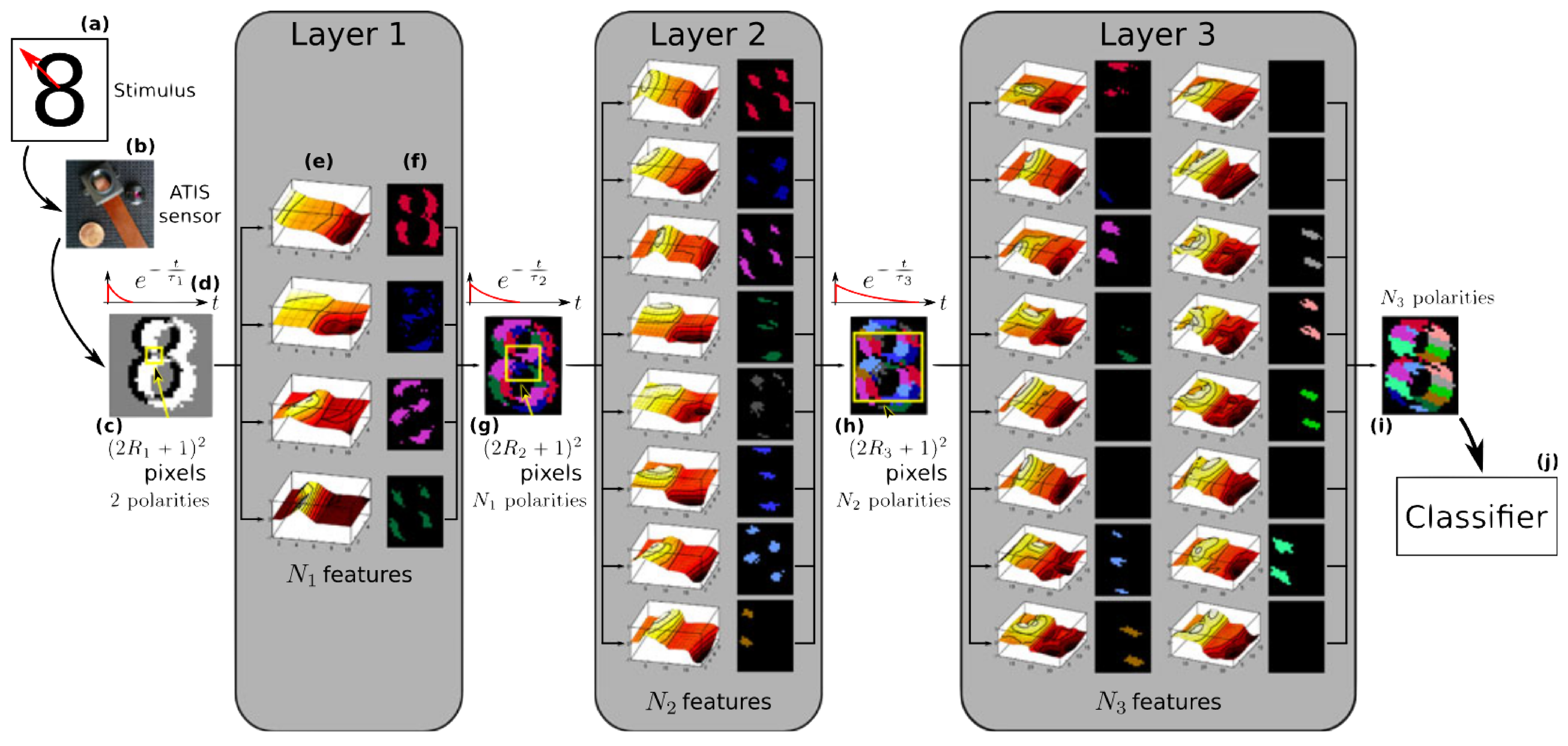

Figure 2: Illustration of the HOTS network extracted from [1]

Materials and Methods

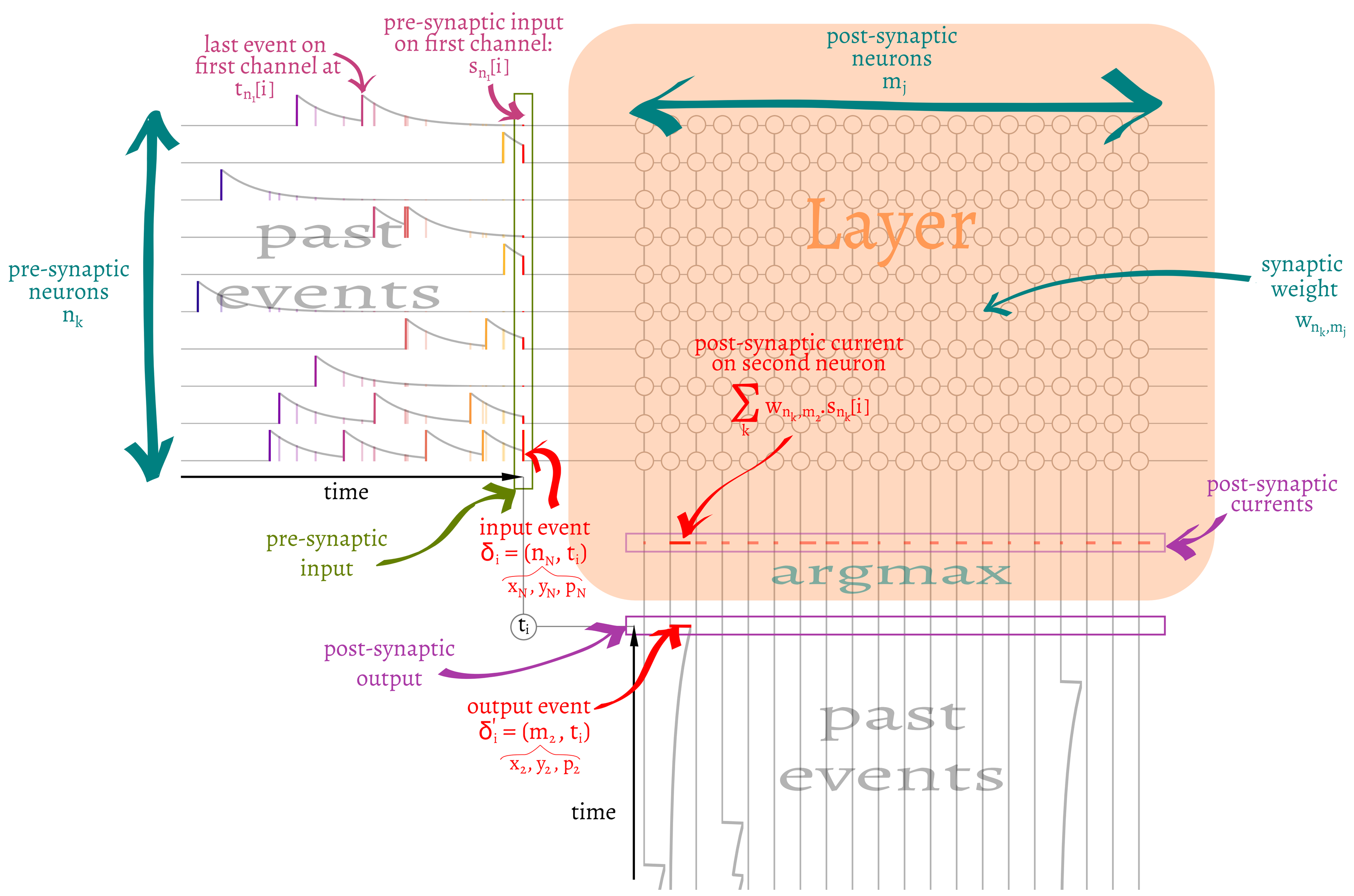

Figure 3: Simplified illustration of neuronal computations within the SNN. On the left, incoming events are illustrated as spikes with colors related to their timing, and overlaid grey curves represent the evolution of pre-synaptic inputs as a function of time, following the differential equation (5). Input of the layer at the time of event δ_i (indicated by a green frame) gathers the contributions of all pre-synaptic neurons n_k. Columns are the post-synaptic neurons m_j within the layer and circles are the synaptic weights w_{n_k,m_j} associated to the different input neurons. These weights adapt through Hebbian learning and result in kernels, also called time surfaces, observed in figure 4. All weighted contributions are summed for each neuron to yield post-synaptic currents (framed in purple). These post-synaptic currents are, for each neuron, a measure of similarity between the synaptic weights and the pre-synaptic input. Output neuron $\mathbf{m}^*$ which gets the highest score ($\arg\max_{m_j}$ nonlinearity) will spike giving rise to an output event. This event gets the same address (x, y) and timing as the input but with a new polarity associated to $\mathbf{m}^*$: $\mathbf{p}*$.

Event-based formalism

For each layer of the network, computations are done discretely on an event-based timestamp. For a given event $\delta_i = (n_i, t_i)$, we update the timing of the last event on each input neuron n_k:

$$t_{n_k}[i] = \max_{j\le i}(t_j | n_j = n_k) \qquad (1)$$

Then, elements of the time surface are computed with the following formula:

$$s_{c_a}[i] = e^{-\frac{t - t_{c_a}[i]}{\tau_{\mathbf{L}}}} \qquad (2)$$

where $\tau_{\mathbf{L}}$ is increased in the next layer such that $\tau_{\mathbf{L+1}} = k_\tau \tau_{\mathbf{L}}$ (usually, $k_\tau = 2$).

For each output neuron m_j of the layer, we compute the scalar product of its weights with the time surface as input:

$$\beta_{m_j} = \frac{\langle W_{m_j}, S[i]\rangle}{\| W_{m_j} \| \| S[i] \|} = \frac{\sum_{n_k} w_{n_k,m_j} \cdot s_{n_k}[i]}{\| W_{m_j} \| \| S[i] \|} \qquad (3)$$

$$\mathbf{m}^* = \arg\max_{m_j}(\beta_{m_j}) \qquad (4)$$

Equation (4) indicates that neuron $\mathbf{m}^*$ will send a spike defining the event's new polarity. The output of layer $\mathbf{L}$ is the event $\delta_i' = (\mathbf{m}^*, t_i)$ where $a_i' = (x_i, y_i, \mathbf{p}^*)$

ODE formalism

For each input neuron n_k the pre-synaptic input can be described on a time continuum by the following differential equation:

$$\tau_{\mathbf{L}} \frac{du_{n_k}(t)}{dt} = -u_{n_k}(t) + (1 - u_{n_k}(t))\delta(t - t_{n_k}(t)) \qquad (5)$$

where

$$t_{n_k}(t) = \max_{t_j \le t}(t_j | n_j = n_k) \qquad (6)$$

It corresponds to a nonlinear Leaky Integrate-and-Fire model where $u_{n_k}(t)$ is the contribution to membrane potential from pre-synaptic neuron n_k. One solution to the equation is

$$u_{n_k}(t) = e^{-\frac{t - t_{n_k}(t)}{\tau_{\mathbf{L}}}} \qquad (7)$$

Synaptic current response on neuron m_j is then

$$I_{m_j}(t) = \frac{\sum_{n_k} w_{n_k,m_j} \cdot u_{n_k}(t)}{\sum_{n_k} w^2_{n_k,m_j} \sum_{n_k} u_{n_k}(t)^2} \qquad (8)$$

- competition between neurons is added by using arg max function
- inhibition of every neuron compensates their activity between every timestamp

Homeostasis

For each layer, a homeostatic gain γ_{m_j} controls the activation of each post-synaptic neuron. We use the simple heuristic derived in [3] to redefine the post-synaptic activation as $\mathbf{m}^* = \arg\max(\gamma_{m_j} \beta_{m_j})$, with $\gamma_{m_j} = e^{\lambda \cdot (p_{\mathbf{P}} \cdot N_L - 1)}$ where λ is a regularization parameter, p_{m_j} is the relative activation frequency of polarity m_j and $N_{\mathbf{L}}$ the total number of features of the layer. This regulation rule allows to train the different features in a balanced fashion and avoid response of some neurons for too specific features.

Results

Clustering

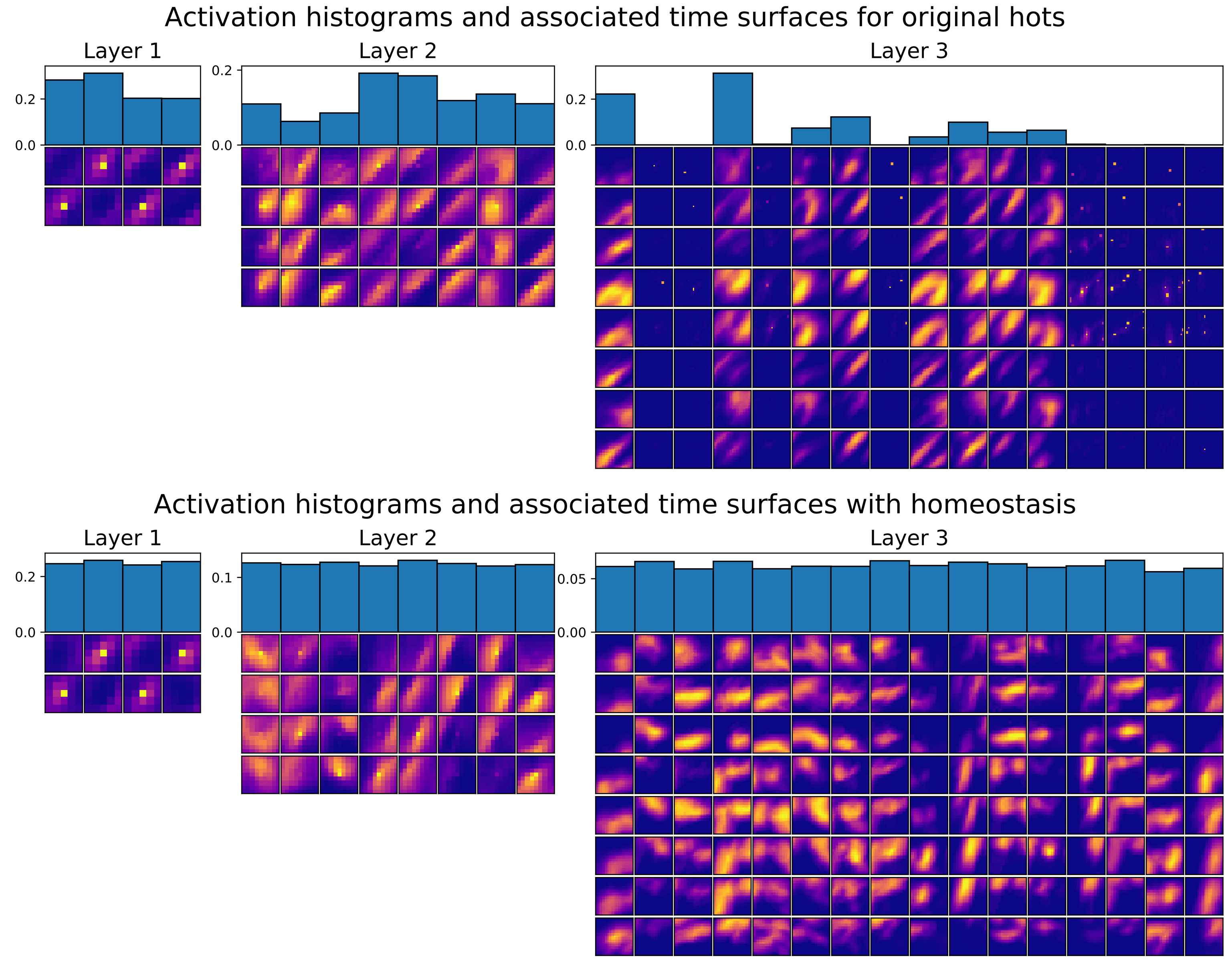

Figure 4: Activation histograms and time surfaces obtained in the self-supervised learning algorithm for the original HOTS network (a) (replicated from [1]) and for the bio-plausible version with homeostasis (b). Associated time surfaces are plotted below histogram bins. The different lines are the different polarities of the features (ON and OFF for the first layer), that is, the output neurons of the previous layer for the next one. Note the unbalanced histograms in the left figure. Some time surfaces are so rarely activated that they remain close to the first inputs they were initialized with. These unevenly matured clusters lead to an inefficient coding within the network.

Classification

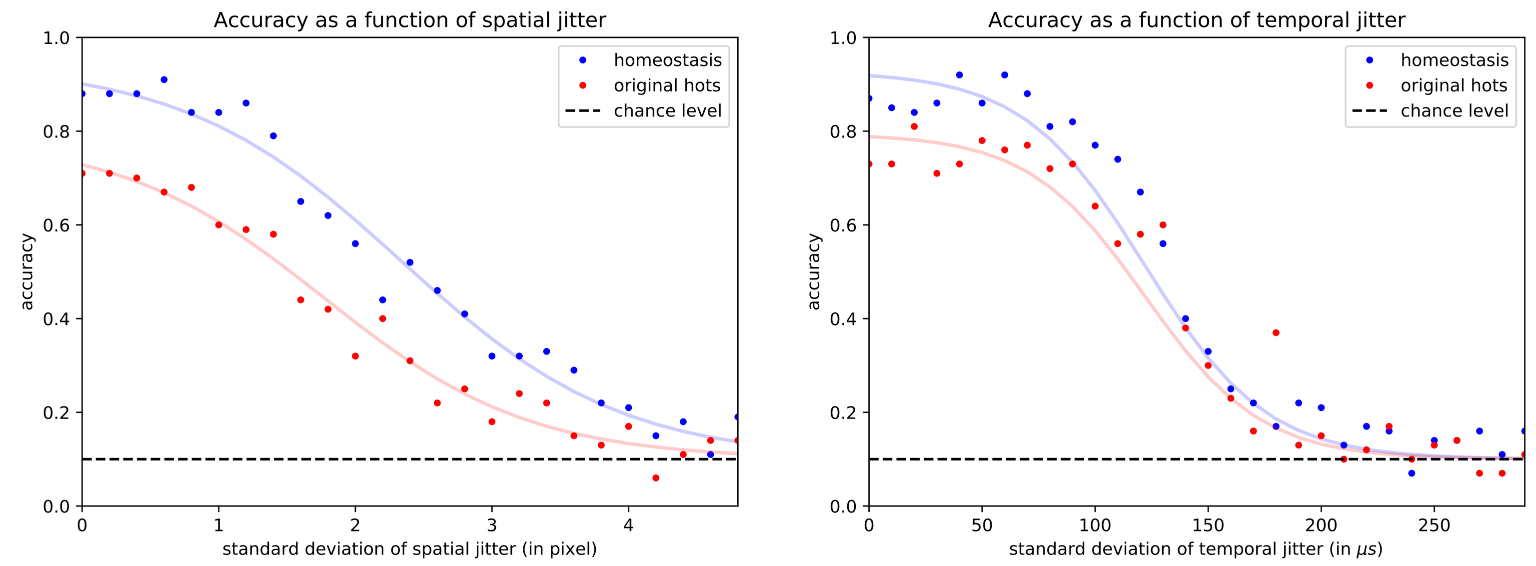

Figure 5: Classification performances on N-MNIST using histogram distances and as a function of (a) spatial or (b) temporal jitter. Blue dots show results of the HOTS method with homeostasis and these are compared to the performance of the original HOTS algorithm (Red dots). Colored lines are the corresponding sigmoid fits of the results.

Conclusion

In this work, we have presented the implementation of a neuromimetic SNN model which is capable to perform online digit classification. Instability of the original method was compensated with a simple homeostatic regulation rule on post-synaptic neurons' activation. Its implementation is available at `https://github.com/SpikeAI/HOTS` and allows to reproduce all results presented here. These results demonstrate the role of competition and cooperation in models of neural computations.

References

[1] X. Lagorce et al. "HOTS: A Hierarchy of Event-Based Time-Surfaces for Pattern Recognition". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.7 (2017), pp. 1346–1359.

[2] G. Orchard et al. "Converting static image datasets to spiking neuromorphic datasets using saccades". In: *Frontiers in neuroscience* 9 (2015), p. 437.

[3] L. U. Perrinet. "An adaptive homeostatic algorithm for the unsupervised learning of visual features". In: *Vision* 3.3 (2019), p. 47.