From event-based computations to a bio-plausible Spiking Neural Network

antoine.grimaldi@univ-amu.fr

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

[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

Abstract

We propose a neuromimetic online classifier for always-on digit recognition. To achieve this, we extend an existing event-based algorithm [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 create an efficient hierarchical event-based pattern recognition architecture. Its formalism was previously adapted in the computational neuroscience domain by showing it may be implemented using a Spiking Neural Network (SNN) of leaky integrate-and-fire models and Hebbian learning [2]. Here, we add an online classification layer using a multinomial logistic regression which is compatible with a neural implementation [3]. A decision can be taken at any arbitrary time by taking the arg max of the probability values associated to each class. We extend the parallel with computational neuroscience by demonstrating that this classification layer is also equivalent to a layer of spiking neurons with a Hebbian-like learning mechanism. Our method obtains state-of-the-art performances on the N-MNIST dataset [4] and we show that it is robust to both spatial and temporal jitter. As a summary, we were able to develop a neuromimetic SNN model for online digit classification. We aim at pursuing the study of this architecture for natural scenes and hope to offer insights on the efficiency of neural computations, and in particular how mechanisms of decision-making may be formed.

Acknowledgments

AG and LP received funding from the European Union ERA-NET CHIST-ERA 2018 research and innovation program under grant agreement N° ANR-19-CHR3-0008-03 ("APROVIS3D"). SI, RB and LP received funding from the ANR project N° ANR-20-CE23-0021 ("AgileNeuroBot").

Event-Based Signal

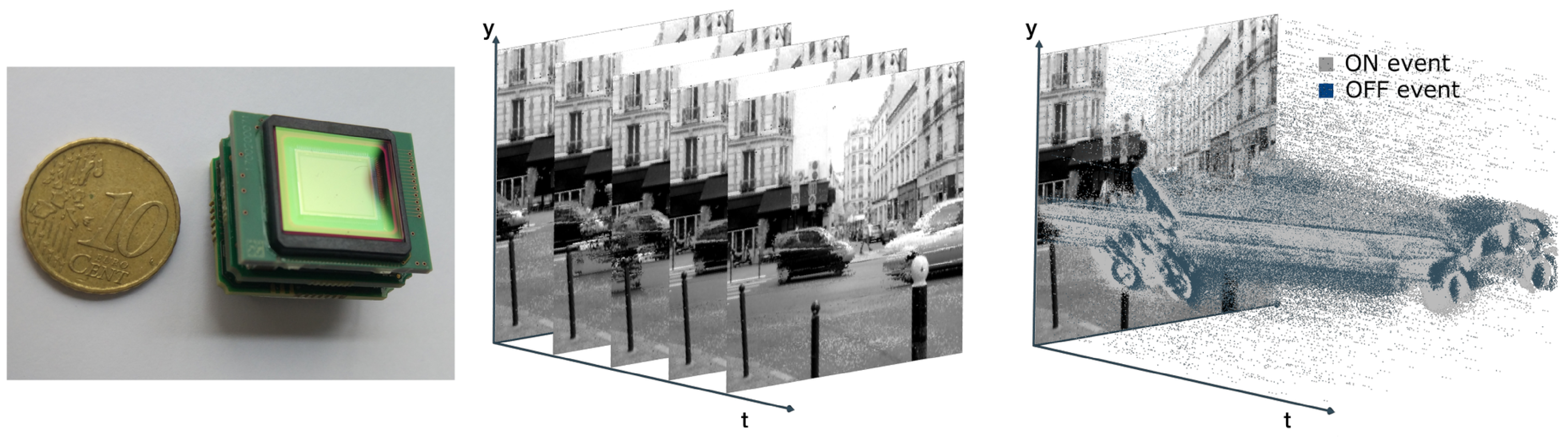

Figure 1: A miniature event-based ATIS sensor (Left) which, compared to a classical frame-based representation (Middle), outputs an event-based representation of the scene (Right). The output of an event-based camera is a discrete stream of events which can be formalized as an ordered set of addresses: $\{a_i\}_{i\in[0,N_{ev})}$ where N_{ev} is the total number of events in the data stream. Each address is typically in the form $a_i = (x_i, y_i, \mathbf{p}_i)$, where (x_i, y_i) defines its position on the pixel grid and $\mathbf{p}_i$ its polarity. ON and OFF polarities correspond respectively to an increase or decrease in luminance at a specific address.

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

- high temporal resolution (typically µs)
- energy efficiency (a few mW)
- redundancy reduction in time and space
- high dynamic range

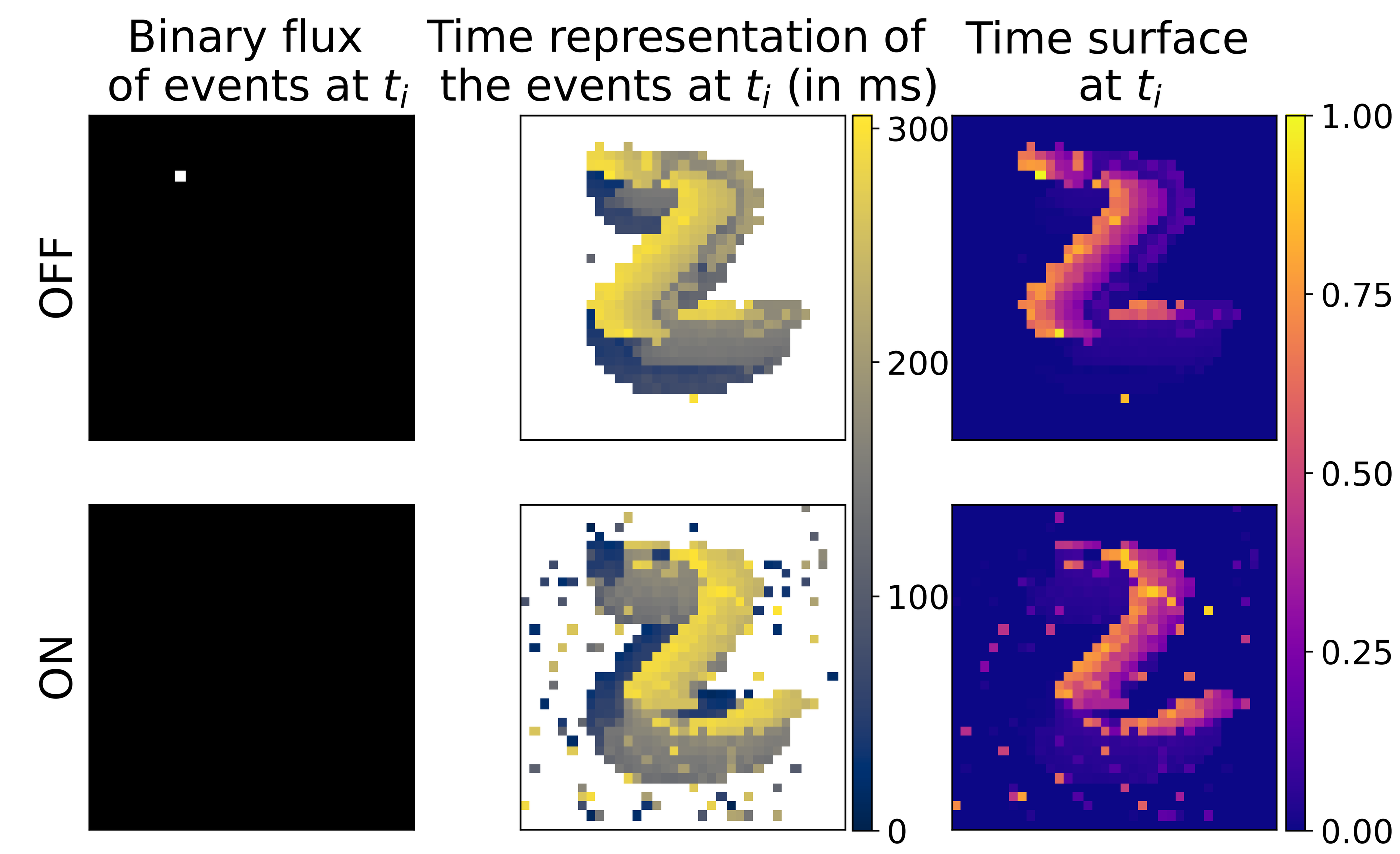

Figure 2: Illustration of the different event-based data types used in the HOTS network at a given event time. The two rows correspond respectively to the OFF and ON polarities of the events as output of the event-based camera. (Left) Screenshot of one single event (in white) at t_i. (Middle) Timings of the latest events recorded, or *time context*, at time t_i, forming the matrix $T(t_i)$ (white represents $-\infty$). (Right) Time surface at t_i as the matrix:

$$\boxed{S(t) = e^{-\frac{t-T(t)}{\tau}}}$$

Core computations of the HOTS network [1]

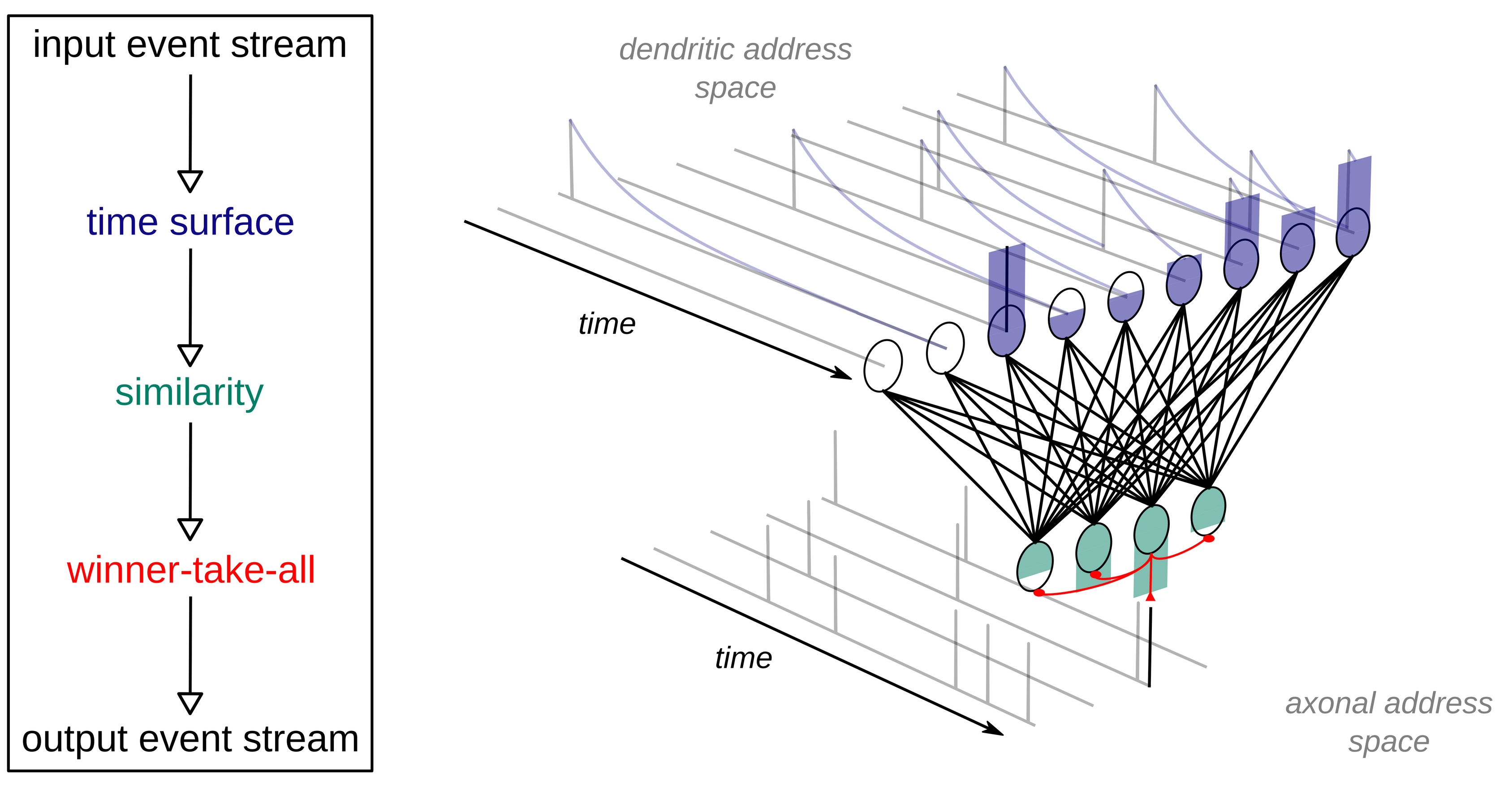

Figure 3: Illustration of the core computation made within one layer of the HOTS algorithm. On the top of the plot, we show the dendritic stream of events convolved by an exponential decay which forms the time surface. Time surfaces are computed at the timestamp of each event/spike. The time surface at present is represented with the colored bar plot on the top. In the vertical slice, computations made within one layer at time t_i are illustrated. The time surface is compared to all the kernels of the layer with the similarity measure resulting in the membrane potential of the postsynaptic neuron represented in green. As an illustration, the layer contains only 4 neurons associated to 4 different kernels and with 10 dendritic inputs. At last, a winner-take-all rule (or arg max non-linearity) will choose at time t_i the most activated neuron. This will emit a spike and prevent the others from being activated through lateral inhibitions (in red). Note that for each event as input of the layer, a new event will be emitted with the same timing as the incoming event.

Multinomial Logistic Regression (MLR) as a classification layer

Similarity mesure:

$$\boxed{\beta_c(t) = \langle W_c, S(t)\rangle = \textstyle\sum_{a\in\mathcal{D}} w_{a,c}\cdot S_a(t)}$$

where S_a is the value of the time surface at address a: $S_a(t) = e^{-\frac{t-T_a(t)}{\tau}}$
and $w_{a,c}$ is the weight associated to class c at address a defined in the dendritic address $\mathcal{D}$.

Probability to predict class c:

$$\forall c \in \{1,\ldots,N_{\text{class}}\},\quad \boxed{Pr(y=c|t;W) = \frac{e^{\beta_c(t)}}{\sum_{j=1}^{N_{\text{class}}} e^{\beta_j(t)}}} = \sigma_c(t)$$

Always-on decision process:

$$\boxed{\hat{y}(t) = \arg\max_c \sigma_c(t) = \arg\max_c \beta_c(t)}$$

Loss function (binary cross-entropy):

$$J(t) = -\textstyle\sum_{c=1}^{N_{\text{class}}} \delta_{\{y(t)=c\}}\cdot \log(\sigma_c(t))$$

where $\delta_{\{y(t)=c\}}$ is the "indicator function" and $y(t_i)$ is the true class.

Update rule during supervised learning:

$$\boxed{\Delta W_c(t) = \begin{cases} \eta\cdot S(t)\cdot(1-\sigma_c(t)), & \text{for } c = y(t) \\ -\eta\cdot S(t)\cdot\sigma_c(t) & \text{for } c \neq y(t) \end{cases}}$$

Analogy with a layer of the HOTS network:

- The decision process of the MLR model is the same as the spiking mechanism used in the HOTS network.

Bio-plausibility:

- The similarity measure corresponds to the integration of the presynaptic spikes with different synaptic weights;
- A stochastic spiking WTA can be built from the softmax function [5];
- Logistic regression is a neurally plausible computation to read out a population code [3];
- The update rule of the weights corresponds to a Hebbian-learning mechanism.

Always-on decision process:

- For each event as input of the network, a decision for digit recognition is taken.

Results

Online classification performance

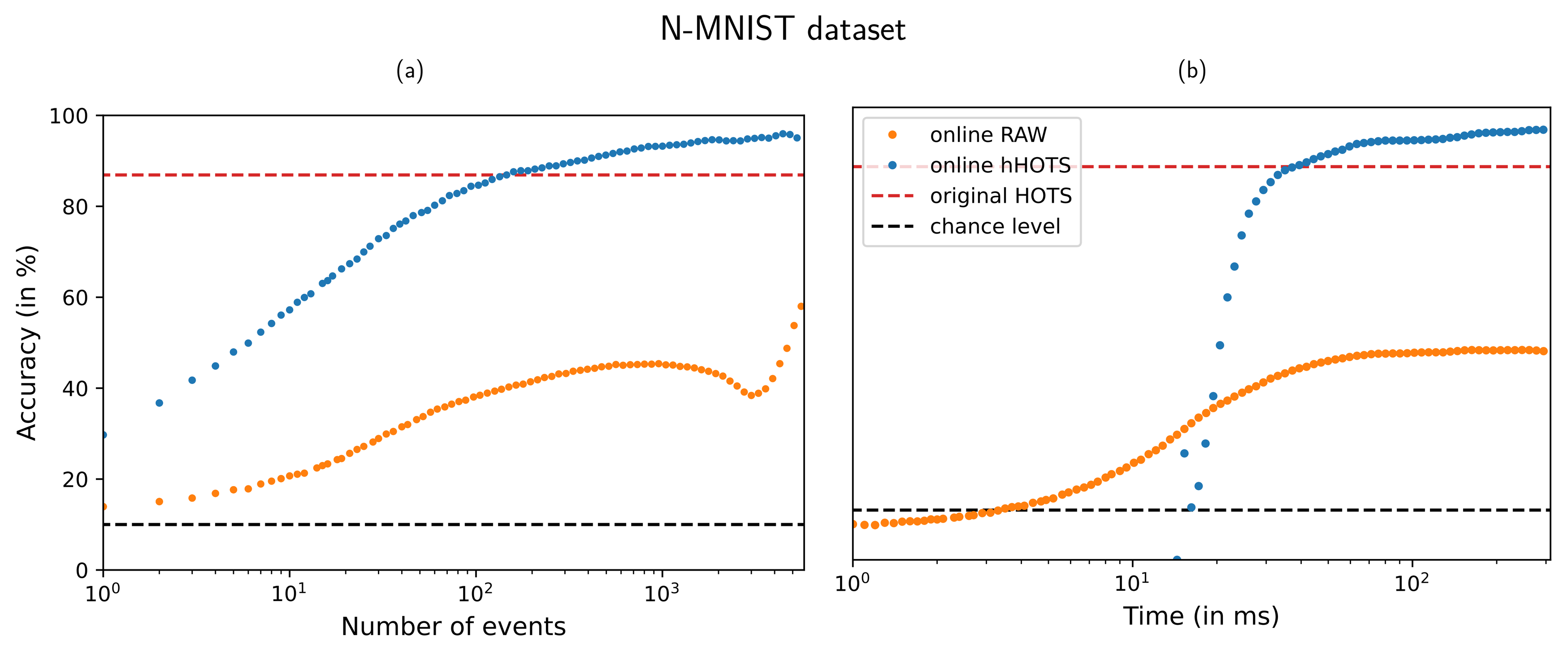

Figure 4: Online classification performance on the N-MNIST dataset. We show the average accuracy computed with respect to the number of events since the beginning of the stream (a) and also as a function of the corresponding time (b). The label *online RAW* corresponds to the MLR classifier fed by the raw stream of events as output of the event-based camera. *online hHOTS* is the method proposed in this work and *original HOTS* is the method described in [1]. In this last method, prediction is performed offline with a k-nearest neighbors classifier on the histograms of activation of the output neurons.

Robustness to jitter

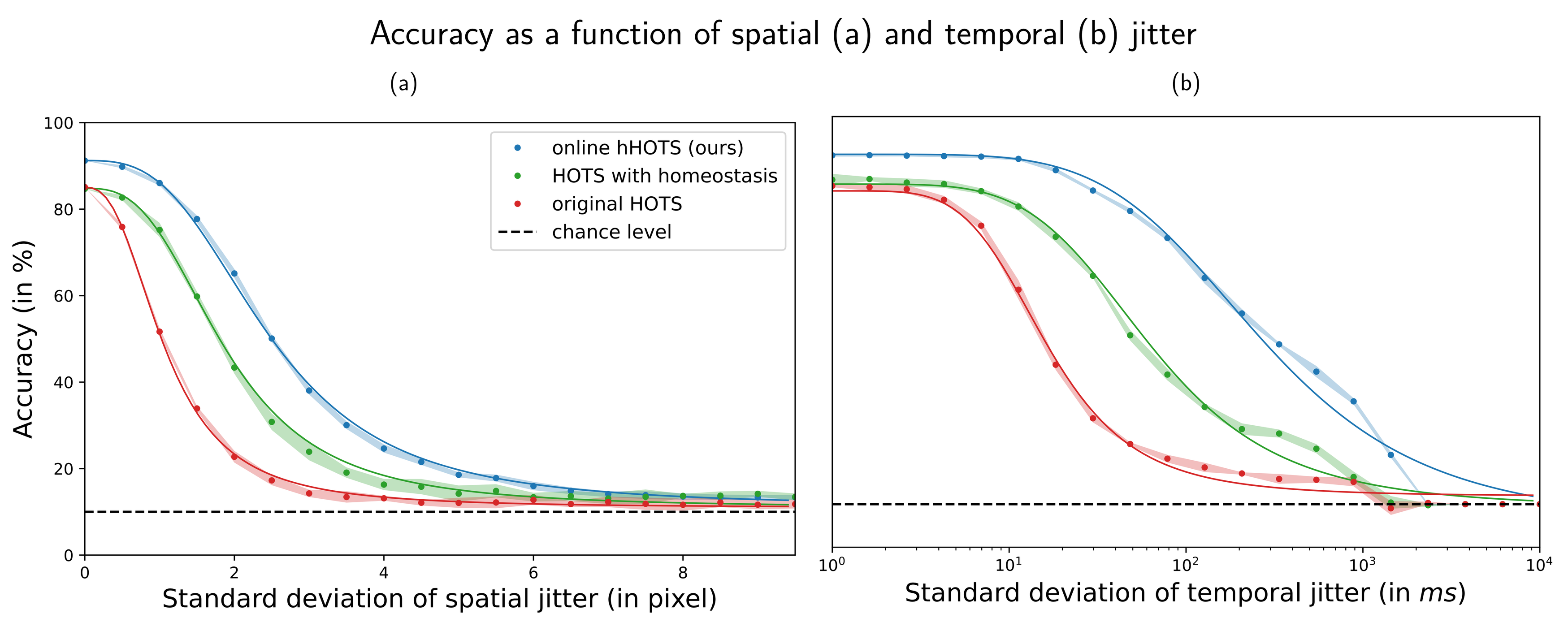

Figure 5: Robustness of classification performances on N-MNIST. The input stream was perturbed by changing the position or timing of the events and we report classification results as a function of (a) standard deviation of spatial jitter or (b) standard deviation of temporal jitter (in logscale). Blue curves show fits for the results of the method presented in this study and these are compared to the performance of the HOTS algorithm with (in green) or without (in red) an homeostatic regulation rule for the activation of the different neurons. With homeostatis, neurons within the same layer are equally activated during the clustering phase (see [6]). Dots are the mean values of accuracy over 10 trials for discrete values of jitter. Transparent outlines represent the 5% and 95% quantiles.

Conclusion

By adding a bio-plausible classifier to an existing event-based algorithm for object recognition, we propose a SNN for online classification. We boost the performances of the original method and allow for robust ultra-fast categorization with this event-driven decision process.

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] A. Grimaldi et al. "A robust bio-inspired approach to event-driven object recognition". In: *Computational and Systems Neuroscience (Cosyne) 2021*. Feb. 26, 2021.
[3] P. Berens et al. "A fast and simple population code for orientation in primate V1". In: *J Neur* 32.31 (2012).
[4] G. Orchard et al. "Converting static image datasets to spiking neuromorphic datasets using saccades". In: *Frontiers in neuroscience* 9 (2015), p. 437.
[5] B. Nessler et al. "Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity". In: *PLoS Comput Biol* 9.4 (2013), e1003037.
[6] A. Grimaldi et al. "A homeostatic gain control mechanism to improve event-driven object recognition". In: *Content-Based Multimedia Indexing (CBMI) 2021*. Apr. 16, 2021.