Escuela AI HUB de verano
Sobre IA y Educación

VALENCIA, DEL 1 AL 5 DE JULIO

Studying the effect of synaptic delays in neuromorphic spatio-temporal datasets

Mario Chacón Falcón[1], Alberto Patiño Saucedo[1], Luis Camuñas Mesa[1], Teresa Serrano Gotarredona[1], Bernabé Linares Barranco[1]

[1] Instituto de Microelectrónica de Sevilla, IMSE-CNM, CSIC/Universidad de Sevilla, Seville, Spain

Introduction

Spiking Neural Networks (SNNs) are emergent models that offer a biologically plausible alternative to artificial neural networks (ANNs) by computing in terms of discrete spikes (a.k.a. events) instead of real-valued activations. Some advantages of using SNNs are:

- **Sparsity and efficiency:** Spike encoding make them more energy efficient if deployed in

appropriate hardware.

- **Temporal dinamics**: Spiking neurons have a self-recurrent state, the **membrane**
- **potential**, which evolves upon incoming spikes, allowing them to process and learn from time-series naturally.

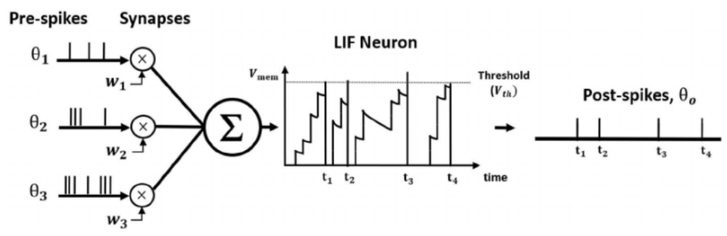
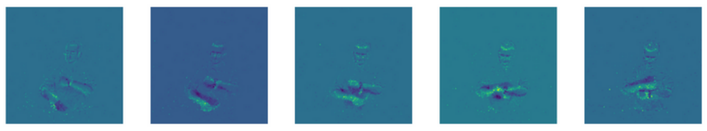


Figure 1: An ilustration of a LIF neuron dinamics [3].

Dataset

Neuromorphic datasets are event-based datasets. In this work, we use the **DVS128Gestures**, recorded with an **event camera**, which detects illumination changes and stream events consisting of a **timestamp, address** and **polarity**. The dataset has recordings of 11 hand gestures from 29 subjects under 3 illumination conditions.

Subfigure (a): Frame representation of a hand clapping instance of the DVSGestures dataset.

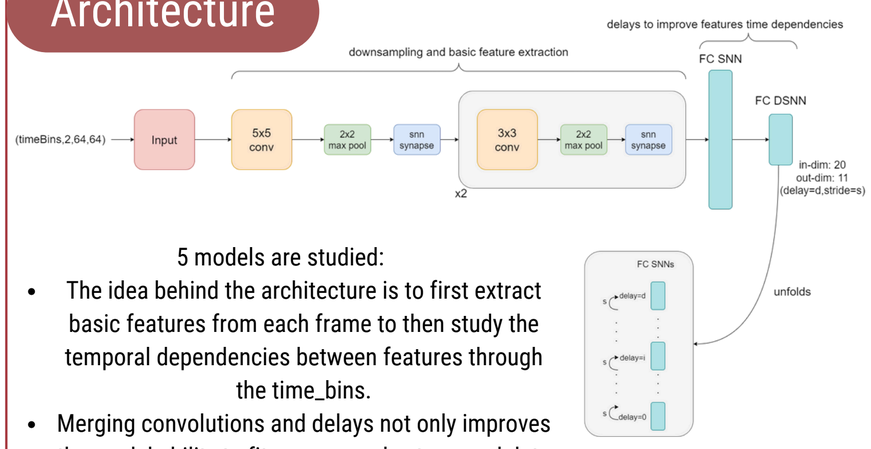


Subfigure (b): Frame representation of an arm roll instance of the DVSGestures dataset.

Subfigure (c): Frame representation of a right hand wave instance of the DVSGestures dataset.

Figure 4: Frame representations examples of three DVS128Gestures instances.

Architecture

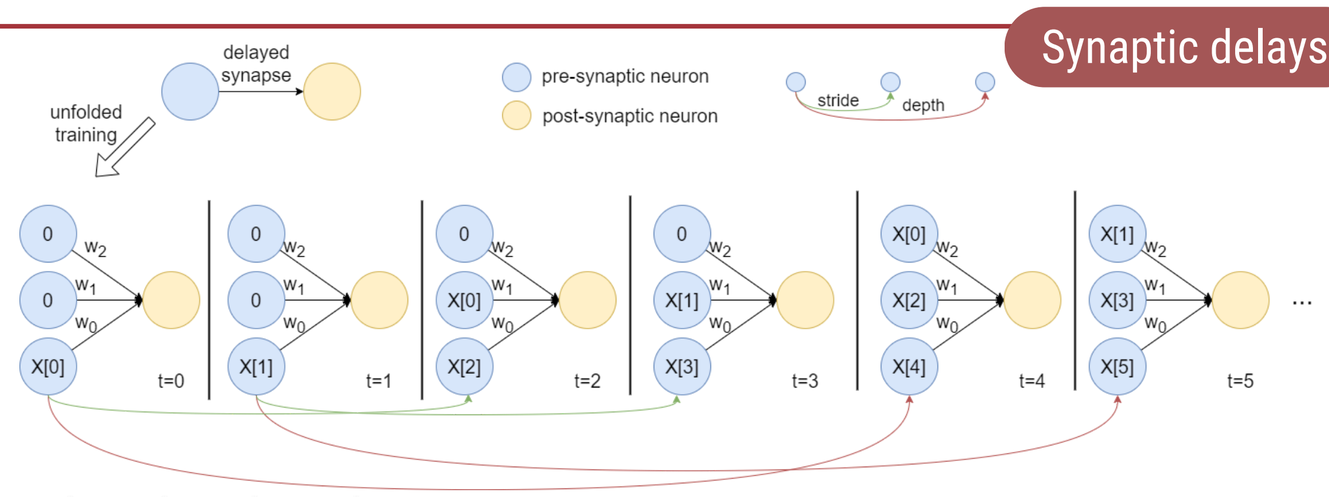


5 models are studied:

- The idea behind the architecture is to first extract basic features from each frame to then study the temporal dependencies between features through the time_bins.
- Merging convolutions and delays not only improves the model ability to fit more complex temporal data but also reduces the number of parameters needed when delays are introduced.
- Each model has the same convolutional structure and only **varies on the number of delays** (10 , 20, 30, 40, 50).
- Every model is trained and **evaluated on 4 different frame representations** of DVS128Gestures dataset (1, 30, 50, 100 time bins).

Figure 4: Overview of the proposed architecture to study delays effect.

Synaptic delays



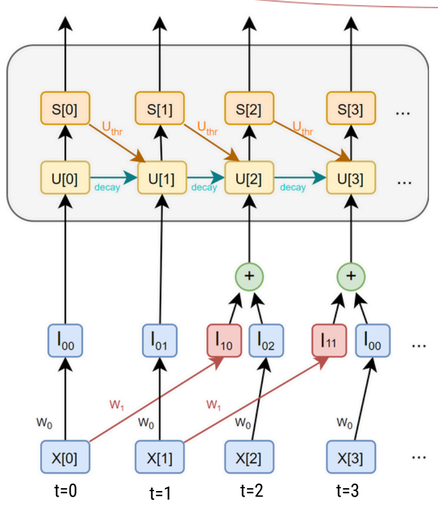


Figure 2: It presents a detailed example of the use of a delayed synapse. In this case, a depth of 4 and stride of 2 delay is shown. The **depth parameter enables the model to learn how long are previus inputs going to be retained** (up to depth). Higher depth will lead to a higher temporal capacity, but also to a higher number of learnable parameters.

Figure 3: Schemagic diagram of how a spiking neuron's membrane potential evolves through time, according to the inputs received every timestep (red arrows represent delayed inputs).

Objective & Methodology

Adding sinaptic delays to SNNs improves their ability to learn temporal features, reducing the number of layers and parameters needed [1].

Objective:

Study the **performance of SNNs using sinaptic delays (D-SNNs)** when used to classify a neuromorphic **spatio-temporal** dataset.

Methodology:

- Adapt sinaptic delays framework to be used with spiking CNNs and build a CNN + Delayed Fully Conected (D-FC) architecture.
- Optimize the spiking CNN architecture for a fixed number of delays.
- Fix the CNN architecture and study the effect of various numbers of delays on a wide range of spatio-temporal frame representation of the selected dataset.

Results & discussion

Results:

- As temporal dependencies of the dataset increases, delays effects become more notable.
- Excesive amount of delays (more delays than time bins) leads to poorer performance.
- When the dataset has no temporal dependencies (1 time bin), the number of delays has no efect in performance.
- A simple architecture with 50k parameters is able to reach a 91% accuracy rate.

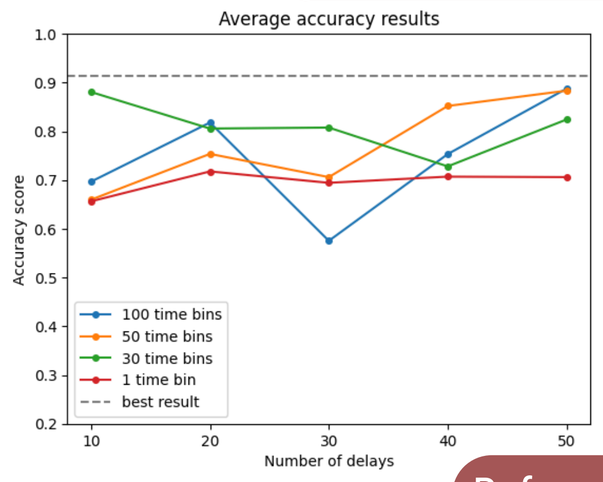


Figure 5: Plot representing the average accuracy results for each model. Each model was trained three times on every frame representation.

Conclusions

Conclusions:

- Study the effect of delays on other spatio-temporal datasets.
- Develop a delayed convolutional layer aiming to improve the memory consumption.
- Perform data fusion on visual and auditive data to create a new dataset and apply convolutions and delays to classify it.

Future work :

- Results show that merging convolutions and delays improves the model performance when the dataset has strong temporal dependencies and reduces the parameters needed when only using delays.
- The number of delays introduced is a hyperparameter to be tuned, since more delays is not necessarily better (it depends on how long temporal dependencies are).

References

[1] A. Patiño-Saucedo, A. Yousefzadeh, G. Tang, F. Corradi, B. Linares-Barranco and M. Sifalakis, "Empirical study on the efficiency of Spiking Neural Networks with axonal delays, and algorithm-hardware benchmarking," 2023 IEEE International Symposium on Circuits and Systems (ISCAS), Monterey, CA, USA, 2023, pp. 1-5, doi: 10.1109/ISCAS46773.2023.10181778.

[2] J. K. Eshraghian et al., "Training Spiking Neural Networks Using Lessons From Deep Learning," in Proceedings of the IEEE, vol. 111, no. 9, pp. 1016-1054, Sept. 2023, doi: 10.1109/JPROC.2023.3308088.

[3] Lee, Chankyu & Sarwar, Syed & Panda, Priyadarshini & Srinivasan, Gopalakrishnan & Roy, Kaushik. (2020). Enabling Spike-Based Backpropagation for Training Deep Neural Network Architectures. Frontiers in Neuroscience. 14. 119. 10.3389/fnins.2020.00119.

CSIC Consejo Superior de Investigaciones Científicas | CSIC HUB Inteligencia Artificial | Fundación "la Caixa" | IFIC Institut de Física Corpuscular | Universitat de València