Efficient Event-Based Object Detection: A Hybrid Neural Network with Spatial and Temporal Attention

[(1,2)]Soikat Hasan Ahmed*, [(1,2)]Jan Finkbeiner† and [(1,2)]Emre Neftci

[(1)] Forschungszentrum Jülich, [(2)] RWTH Aachen University

Event cameras offer high temporal resolution and dynamic range with minimal motion blur, making them promising for robust object detection. While Spiking Neural Networks (SNNs) on neuromorphic hardware are often considered for energy efficient and low latency event-based data processing, they often fall short of Artificial Neural Networks (ANNs) in accuracy and flexibility.
Here, we introduce Attention-based Hybrid SNN-ANN backbones for event-based object detection to leverage the strengths of both SNN and ANN architectures.
A novel Attention-based SNN-ANN bridge module captures sparse spatial and temporal relations from the SNN layer and converts them into dense feature maps for the ANN part of the backbone. Additionally, we present a variant that integrates DWConvLSTMs to the ANN blocks to capture slower dynamics. This multi-timescale network combines fast SNN processing for short timesteps with long-term dense RNN processing, effectively capturing both fast and slow dynamics.
Experimental results demonstrate that our proposed method surpasses SNN-based approaches by significant margins, with results comparable to existing ANN and RNN-based methods. Unlike ANN-only networks, the hybrid setup allows us to implement the SNN blocks on digital neuromorphic hardware to investigate the feasibility of our approach.
Extensive ablation studies and implementation on neuromorphic hardware confirm the effectiveness of our proposed modules and architectural choices.
Our hybrid SNN-ANN architectures pave the way for ANN-like performance at a drastically reduced parameter, latency, and power budget.

*Conceptual design, algorithm development and experimentation.
†Conceptual design, hardware analysis and deployment.