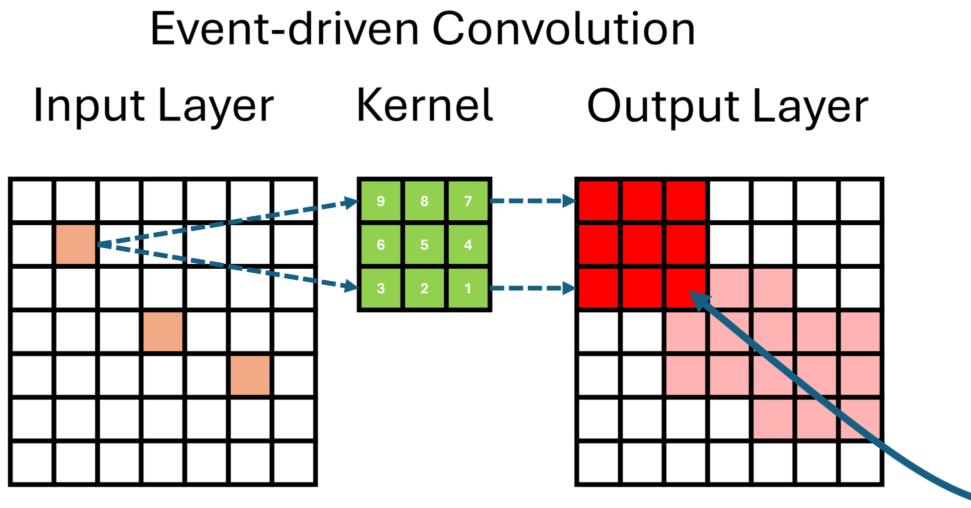
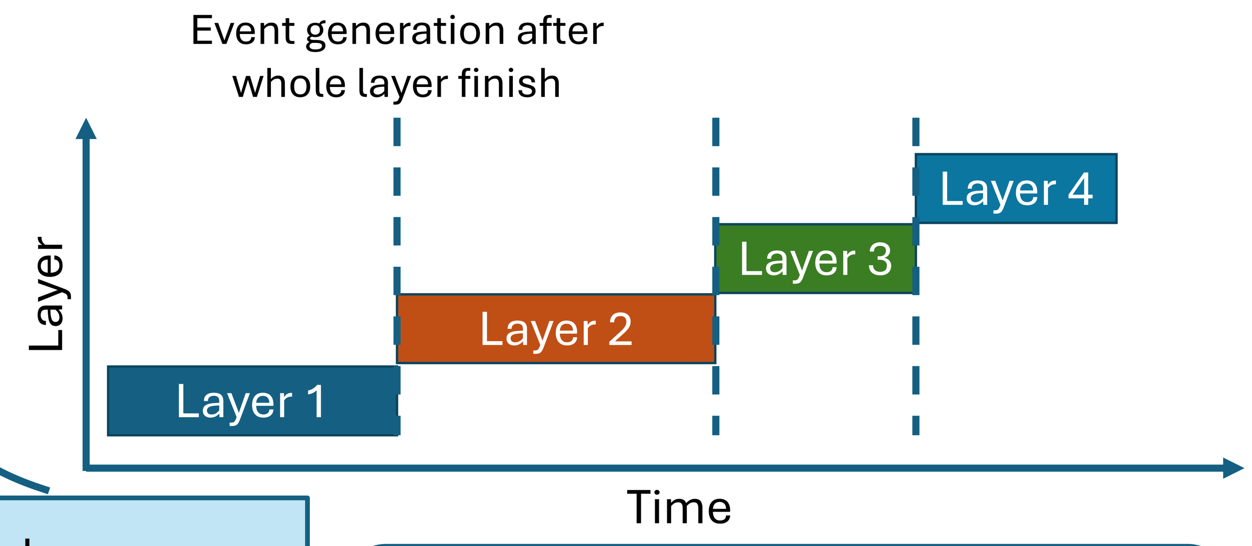
Event-based Optical Flow on Neuromorphic Processor with Efficient Event-driven Depth First Convolution

Yingfu Xu[1], **Guangzhi Tang**[2], Amirreza Yousefzadeh[3], Guido C.H.E. de Croon[4], Manolis Sifalakis[1]

[1] imec the Netherlands, [2] Maastricht University, [3] University of Twente, [4] Delft University of Technology

Problems of **ConvNet** on Neuromorphic Processor



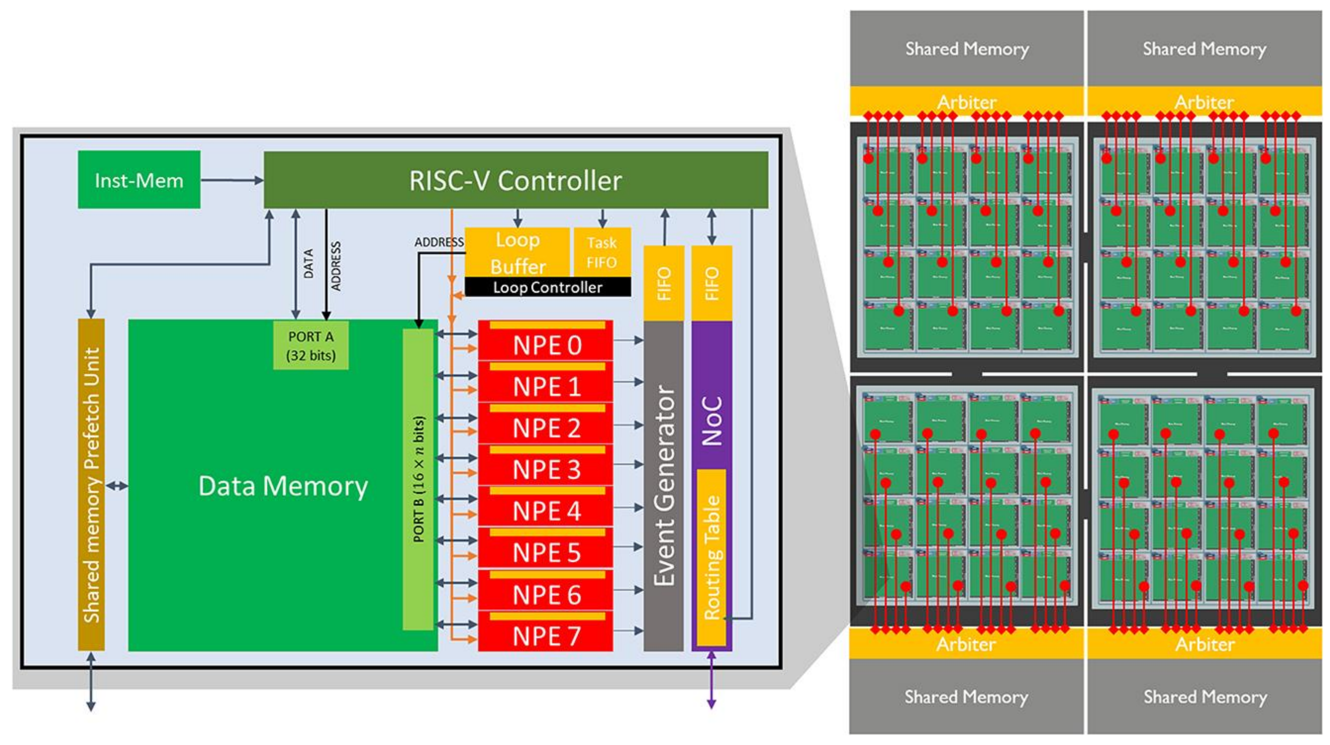


❑ Partial sums
❑ Wait for complete

Problem 1: State Memory
- Store all states on-chip
- Unbearable for high-res

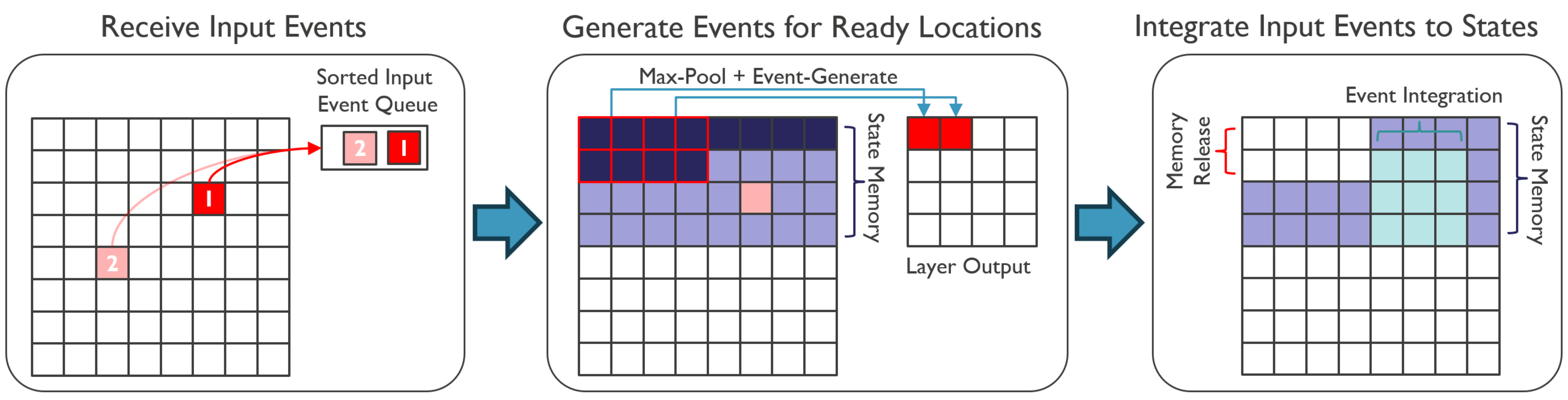
Problem 2: Layer Latency
- Event generate after step sync
- Additional latency per layer

SENECA Neuromorphic Architecture *[1]*



- **Scalable and flexible** architecture design
- Programable **controller** and **hierarchical memory**

Event-driven Depth First Convolution *[2]*



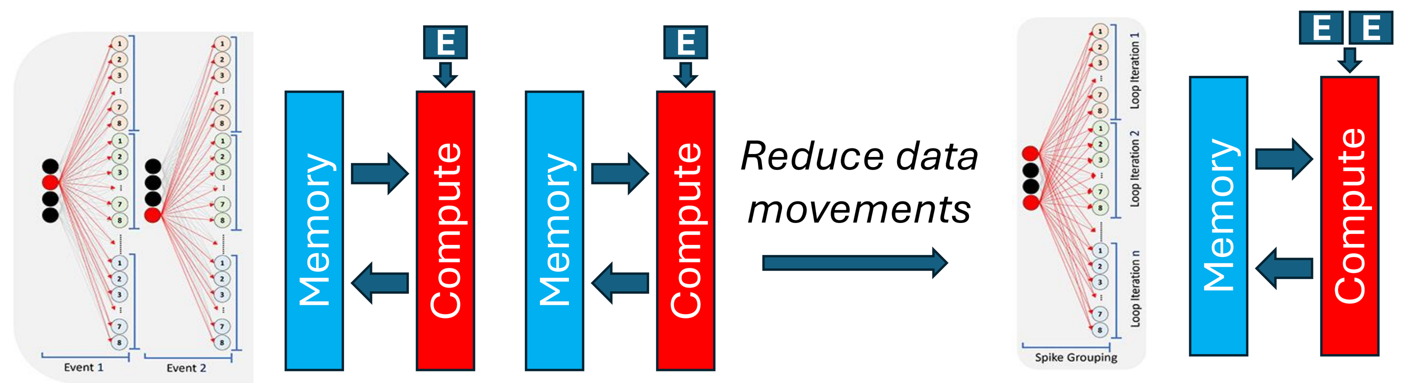
Reduce state memory cost
X*Y*C to **min(X, Y)*C*4**

Reduce layer latency
Full layer to **First event**

Improve vs	Energy	Latency	Area
Loihi	>50x	>3x	>3x
TrueNorth	>9x	>3x	>100x

** 5-layer CNN: (8c3k2p)-(16c3k2p)-(32c3k2p)-128-class*

Spike-Grouping for Reducing Memory Access

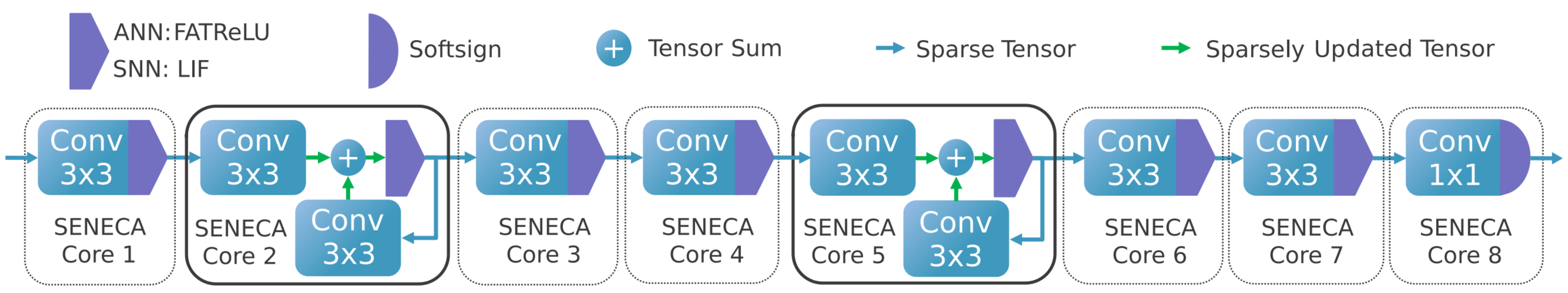


Event-by-event
2 state mem access / event

Spike-grouping
2 state mem access / group

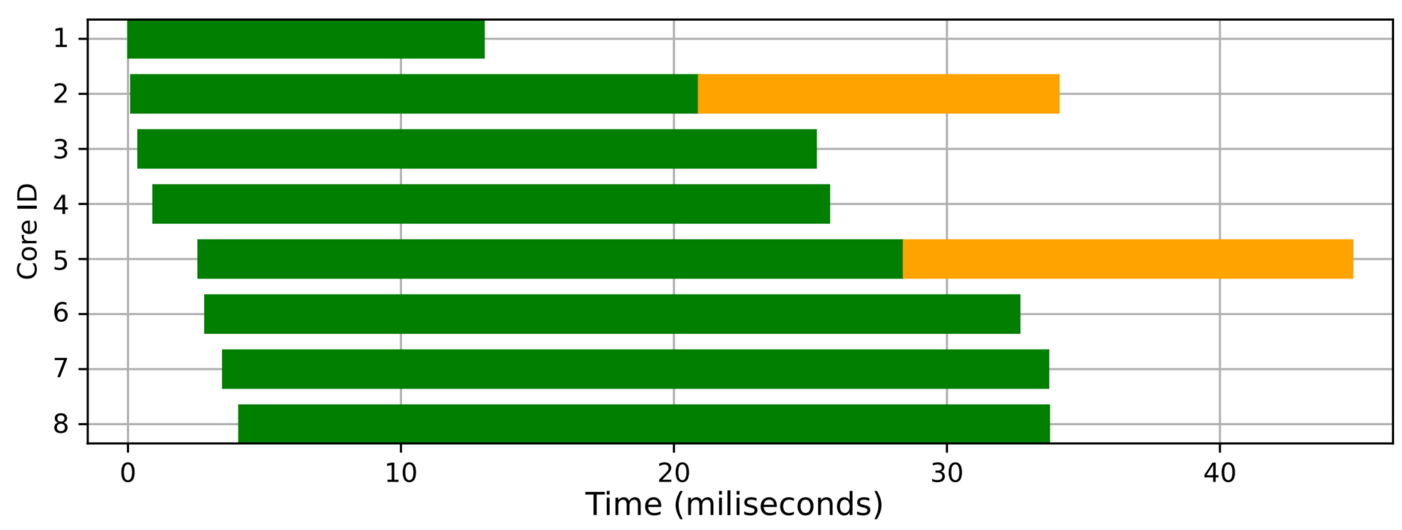
- Reduce **neural state** memory movements
- Combine with Convolution for events in **same pixel**

FireNet with **Sparse ANN or SNN** for Event-based Optical Flow Prediction *[3]*



- **Event-based Optical Flow Prediction**: Estimation optical flow using event camera
- **Fair Comparison of ANN and SNN**: Similar architecture, sparsity, deploy hardware
- **Hardware-aware Training**: Novel activation sparsity finetuning for ANN and SNN

SNN FireNet Deploy on SENECA

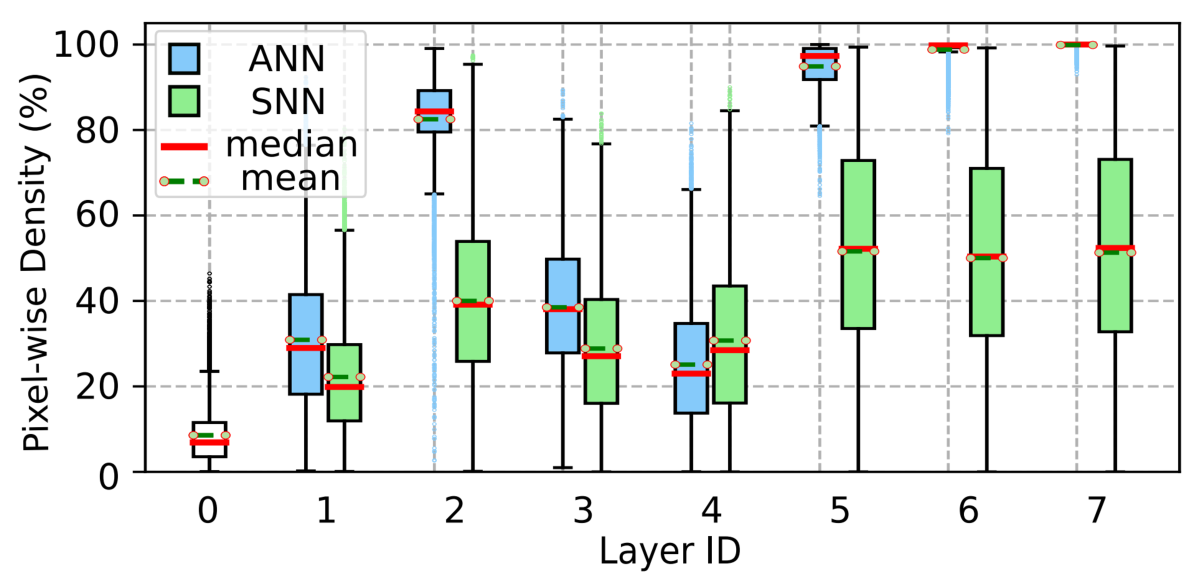


- Deploy SNN on 8 interconnected SENECA cores
- Layers operate in pipeline fashion without waiting

Fair Comparisons of Sparse ANN and SNN on Neuromorphic Optical Flow

Network	outdoor_day1		indoor_flying1		indoor_flying2		indoor_flying3		Average		
	AEE	$\%_{Out.}$	AEE	$\%_{Out.}$	AEE	$\%_{Out.}$	AEE	$\%_{Out.}$	AEE	$\%_{Out.}$	Dens.(%)
EV-FlowNet(GRU)Hagenaars et al. (2021)	**1.69**	12.50	2.16	21.51	3.90	40.72	3.00	29.60	2.94	29.35	-
RNN-EV-FlowNet	**1.69**	12.96	**2.02**	**18.74**	3.84	38.17	2.97	**27.91**	2.88	**27.32**	16.90
RNN-EV-FlowNet-S (smaller λ_s)	1.92	17.34	2.06	18.83	**3.56**	**37.02**	**2.88**	28.94	**2.79**	27.76	5.35
RNN-EV-FlowNet-S (bigger λ_s)	1.73	**12.20**	2.03	19.03	3.83	39.70	3.02	30.58	2.90	28.71	4.78
LIF-EV-FlowNet	1.99	15.99	2.47	26.79	4.94	50.51	3.91	39.59	3.68	37.47	9.46
LIF-EV-FlowNet-S (smaller λ_s)	2.01	16.88	2.69	32.00	4.77	51.29	3.84	41.85	3.66	39.90	5.81
LIF-EV-FlowNet-S (bigger λ_s)	1.88	16.36	2.76	33.63	4.96	52.65	4.06	44.76	3.80	41.68	**3.93**
FireNet(GRU)Hagenaars et al. (2021)	2.04	20.93	3.35	42.50	5.71	61.03	4.68	53.42	4.41	49.92	-
RNN-FireNet	1.94	17.80	3.11	38.79	5.45	57.31	4.47	49.59	4.19	46.22	34.03
RNN-FireNet-S* (no threshold regularizer)	**1.67**	**12.88**	**2.79**	**32.70**	**5.02**	**51.99**	**4.05**	**43.69**	**3.80**	**40.54**	16.57
RNN-FireNet-S	2.16	22.04	3.16	40.09	5.14	55.96	4.24	48.76	4.05	46.25	5.92
RNN-FireNet-S-FT	1.97	18.31	3.24	39.23	5.48	57.00	4.45	49.02	4.22	46.09	**4.52**
LIF-FireNet	1.96	15.82	3.32	41.37	5.99	62.24	4.98	54.63	4.58	49.94	19.54
LIF-FireNet-S	2.15	21.06	3.14	39.17	5.59	57.94	4.63	51.19	4.31	47.36	4.53

Where does SNN work better than ANN?



- SNN having **higher Pixel-wise sparsity** than ANN
- More events in pixels increase **data reuse** chances
- Result in **lower energy and latency** on hardware

[1] Tang, Guangzhi, et al. "SENECA: building a fully digital neuromorphic processor, design trade-offs and challenges." Frontiers in Neuroscience, 2023.
[2] Xu, Yingfu, et al. "Optimizing event-based neural networks on digital neuromorphic architecture: a comprehensive design space exploration." Frontiers in Neuroscience, 2024.
[3] Xu, Yingfu, et al. "Event-based Optical Flow on Neuromorphic Processor: ANN vs. SNN Comparison based on Activation Sparsification." arXiv preprint, 2024.

imec embracing a better life
UNIVERSITY OF TWENTE.
Maastricht University
TU Delft
Chips JU

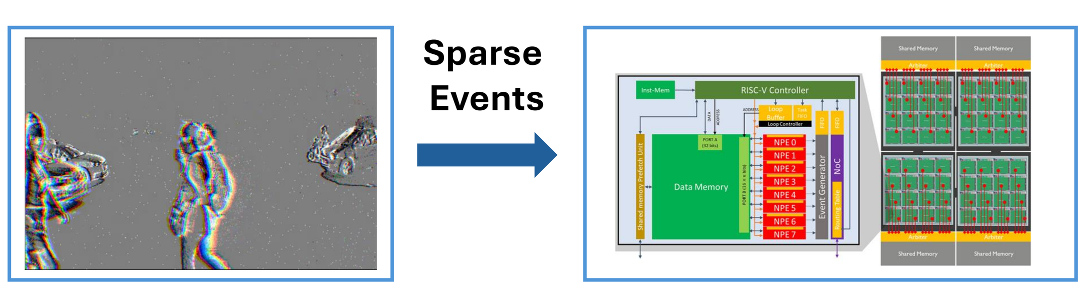
*This work was partially funded by research and innovation projects **REBECCA** (KDT JU under grant agreement No. 101097224), **NeuroKIT2E** (KDT JU under grant agreement No. 101112268)*

TRIP: Trainable Region-of-Interest Prediction for Hardware-Efficient Neuromorphic Processing on Event-based Vision

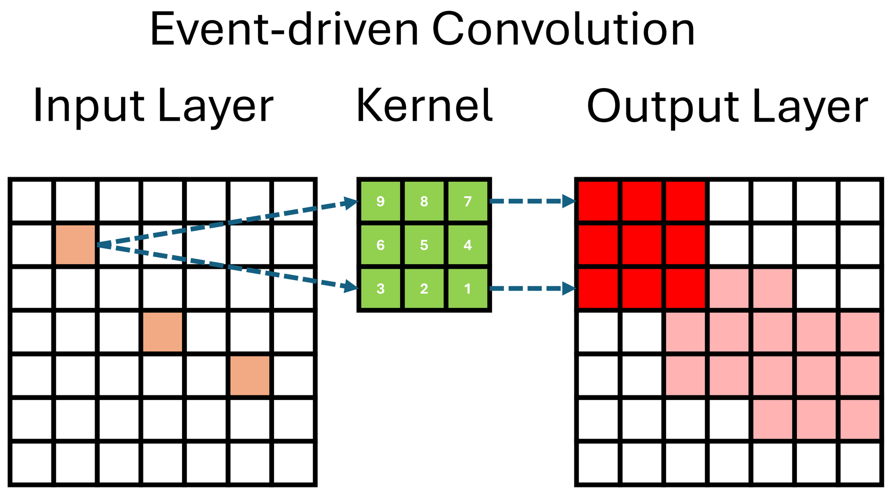
Cina Arjmand[1], Yingfu Xu[1], Kevin Shidqi[1], Alexandra F. Dobrita[1], Kanishkan Vadivel[1], Paul Detterer[1], Manolis Sifalakis[1], Amirreza Yousefzadeh[2], **Guangzhi Tang[3]**

[1] imec the Netherlands, [2] University of Twente, [3] Maastricht University

Neuromorphic Processing on Event-based Vision



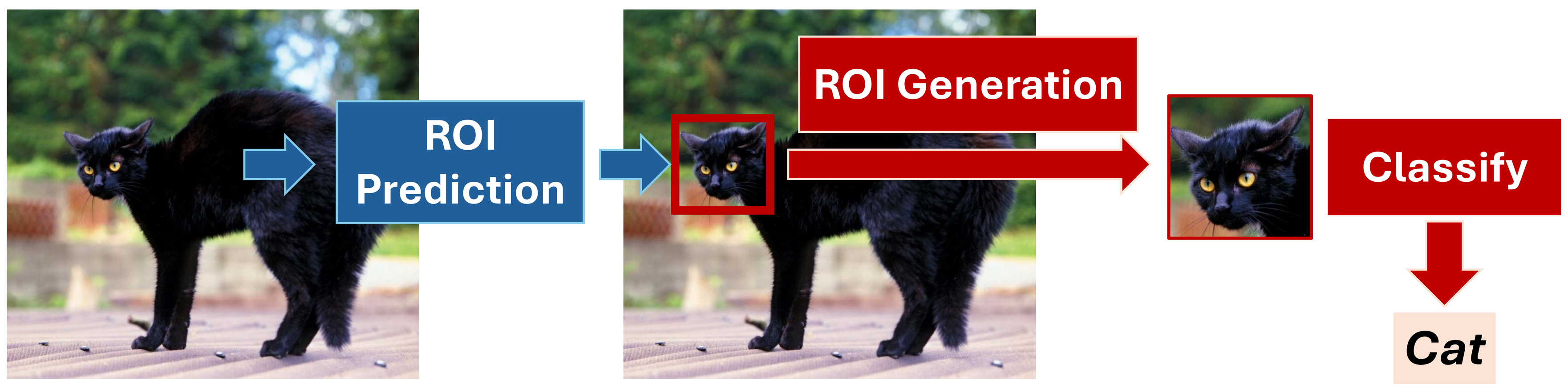
Advantages
- Exploit input sparsity
- Low latency and energy



Challenge 1: Memory Cost
- Large neural state memory
- High area cost

Challenge 2: Compute Cost
- Large number of events
- High computation cost

Hard Attention for Efficient Image Classification

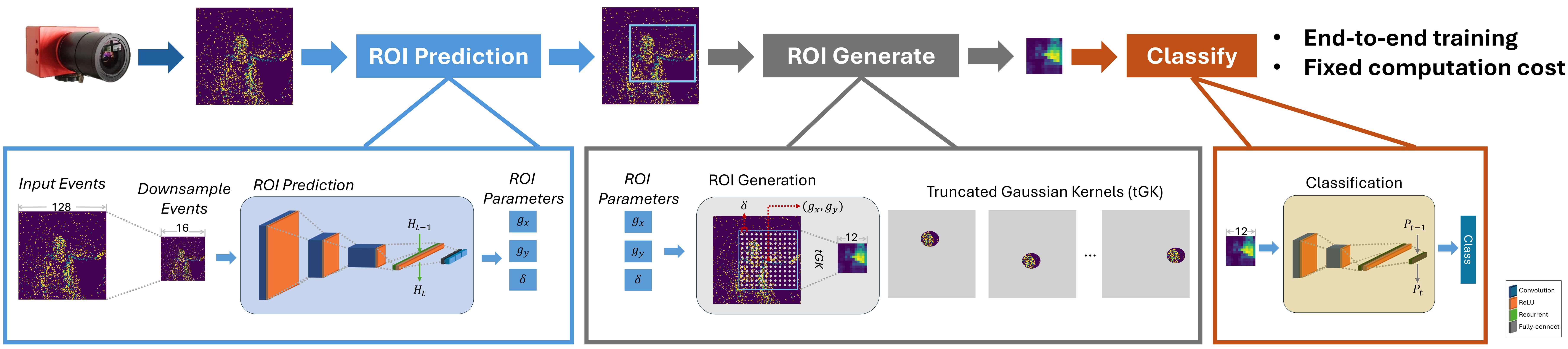


Advantage: Process high resolution image without quadratic complexity to input scale

Challenge 1: Overhead for ROI
Costly ROI prediction and generation for complex scenes

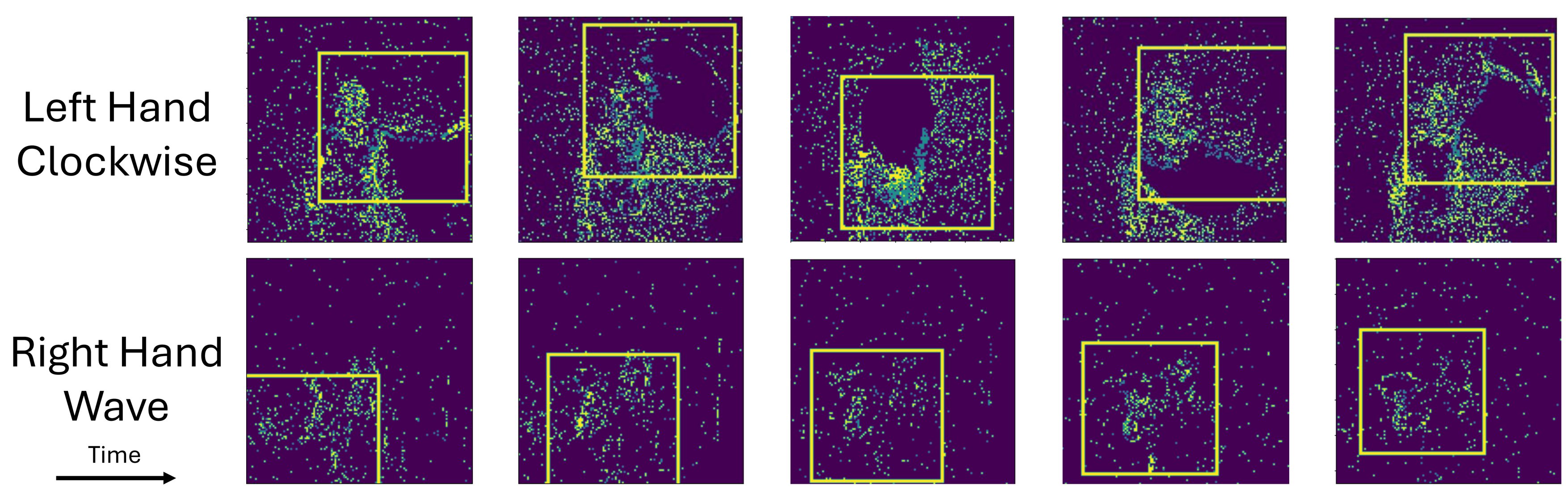
Challenge 2: Training Complexity
Hard to perform end-to-end training with simple architecture

TRIP – Hard Attention Framework for Event-based Vision on Event-driven Neuromorphic Processor *[1]*



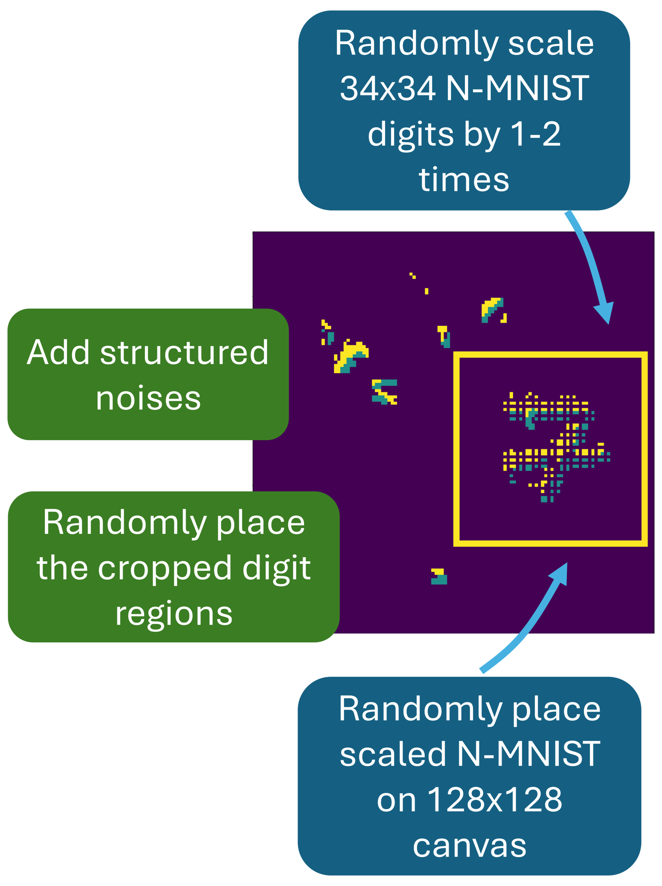
- **End-to-end training**
- **Fixed computation cost**

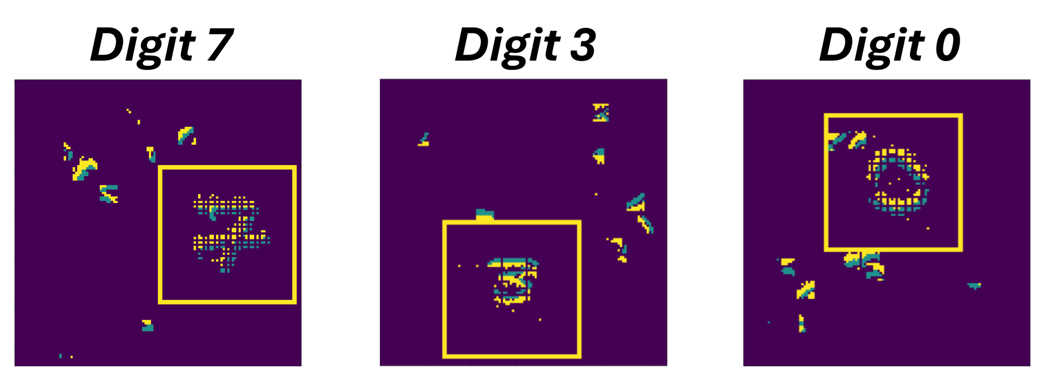
Experiments on the **DVSGesture** Dataset



Architecture	Input Resolution	Param	Effective MACs (Single Timebin)	Accuracy [%] (mean ± std)	Accuracy [%] (Maximum)
LSTM [20]	32×32	7.4M	3.9M	–	86.8
AlexNet+LSTM [21]	128×128	8.3M	601.3M	–	97.7
CNN+EGRU [13]	128×128	4.8M	80.6M	97.3± 0.4	97.8
ConvLIAF [22]	32×32	**0.22M**	113.3M	–	97.6
TRIP (Ours)	16×16+12×12	0.46M	**1.75M**	**97.6 ± 0.5**	**98.6**

Experiments on the **Synthetic NMNIST** Dataset



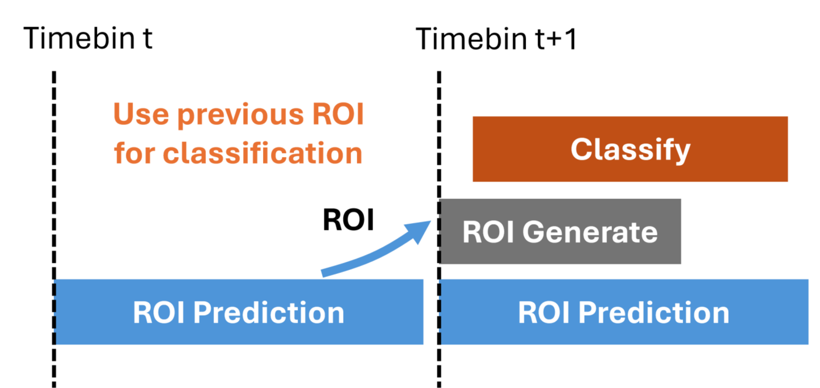


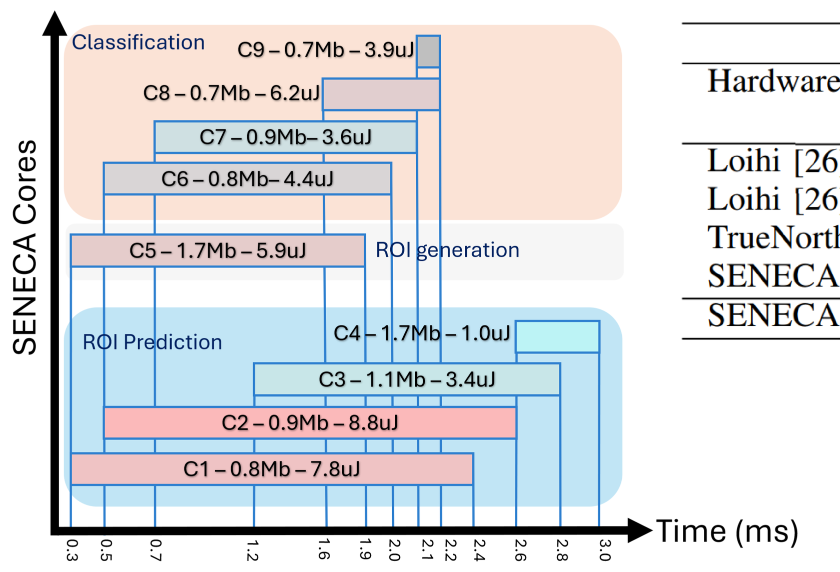
Architecture	Param	FLOPs	Accuracy [%] (mean ± std)
Baseline (16x16)	0.31M	6.0M	71.8± 2.3
Baseline (32x32)	0.67M	24.4M	93.0± 0.6
Baseline (64x64)	0.67M	57.4M	96.2± 0.9
TRIP (16x16)	0.30M	16.0M	95.4 ± 0.4
TRIP (32x32)	0.65M	28.0M	96.1 ± 0.3

TRIP achieves **higher accuracy** compared to CNN with same number of layers and parameters

Deploy TRIP on the **SENECA Neuromorphic Processor**

Remove causal processing in TRIP





					Single Timebin			Multiple Timebins		
Hardware	Solutions	Technology	Core [#]	Area [mm^2]	Latency [ms]	E_{inf} [uJ]	Accuracy [%]	Latency [ms]	E_{inf} [uJ]	Accuracy [%]
Loihi [26]	Spiking CNN [3]	Intel 14 nm	>20	>8.20	11	–	89.6	–	–	–
Loihi [26]	Spiking CNN [14]	Intel 14 nm	59	24.19	–	–	–	**22.0**	2731	96.2
TrueNorth [27]	Spiking CNN [11]	Samsung 28 nm	3838	383.8	–	–	**91.8**	104.6	18702	94.6
SENECA [10]	Event-based CNN	GF FDX 22 nm	7	**3.29**	–	–	–	78.9	1069.2	97.3
SENECA [10]	TRIP	GF FDX 22 nm	9	4.23	**2.7**	**35.86**	91.1	25.8	**430.32**	**98.3**

- **End-to-end deployment on the SENECA neuromorphic processor**
- **Improve latency and reduce energy cost using less hardware area**

[1] Arjmand, Cina, et al. "TRIP: Trainable Region-of-Interest Prediction for Hardware-Efficient Neuromorphic Processing on Event-based Vision." International Conference on Neuromorphic Systems, 2024.

imec embracing a better life
UNIVERSITY OF TWENTE.
Maastricht University
Funded by the European Union
Chips JU

*This work was partially funded by research and innovation projects **REBECCA** (KDT JU under grant agreement No. 101097224), **NeuroKIT2E** (KDT JU under grant agreement No. 101112268), and **NimbleAI** (Horizon EU under grant agreement 101070679)*