Optimising Graph Representation for Hardware Implementation of Graph Convolutional Networks for Event-based Vision

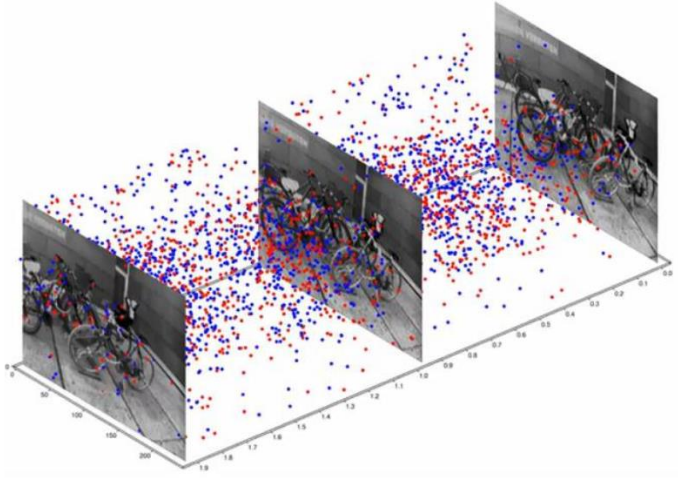
Kamil Jeziorek[1], Piotr Wzorek[1], Krzysztof Błachut[1], Andrea Pinna[2], Tomasz Kryjak[1,2]

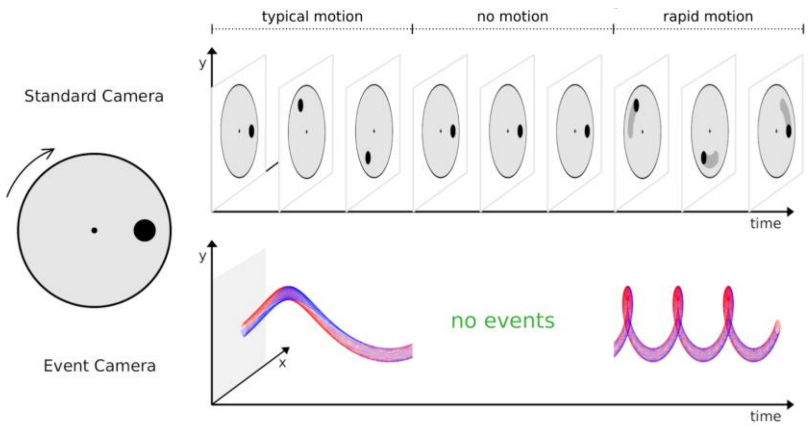
1 Embedded Vision Systems Group,
Department of Automatic Control and Robotics,
AGH University of Krakow, Poland

2 Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

Dynamic Vision Sensor

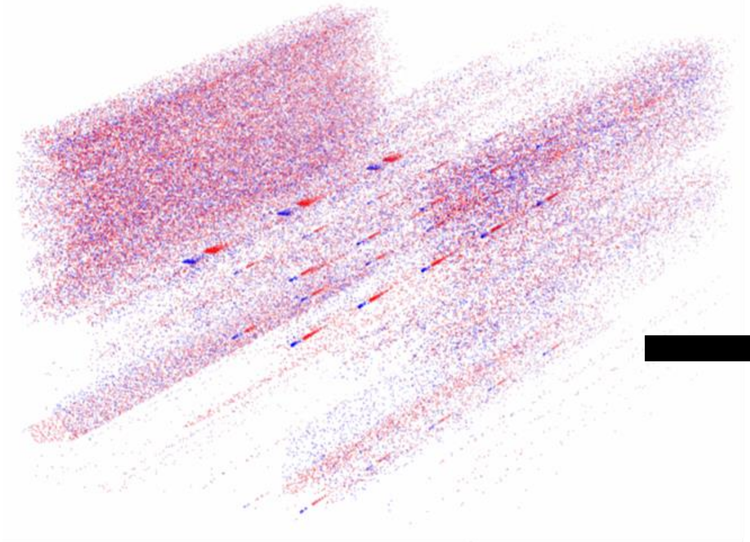
- Low latency (~1 µs vs 16 ms)
- High dynamic range (140 dB vs 60 dB)
- Increased resistance to motion blur
- Low power consumption (1 mW vs 1 W average)



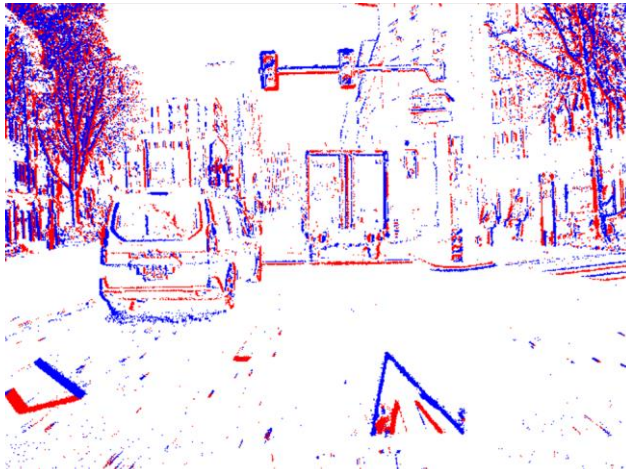


Standard methods of event data processing

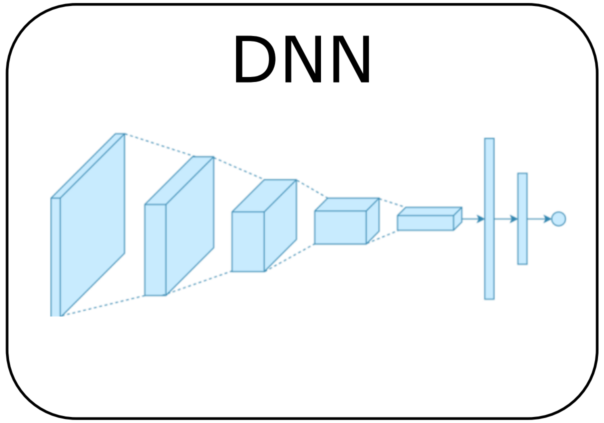
- The events are projected onto a 2D plane to match the CNN models.



Event flattening



Event frame

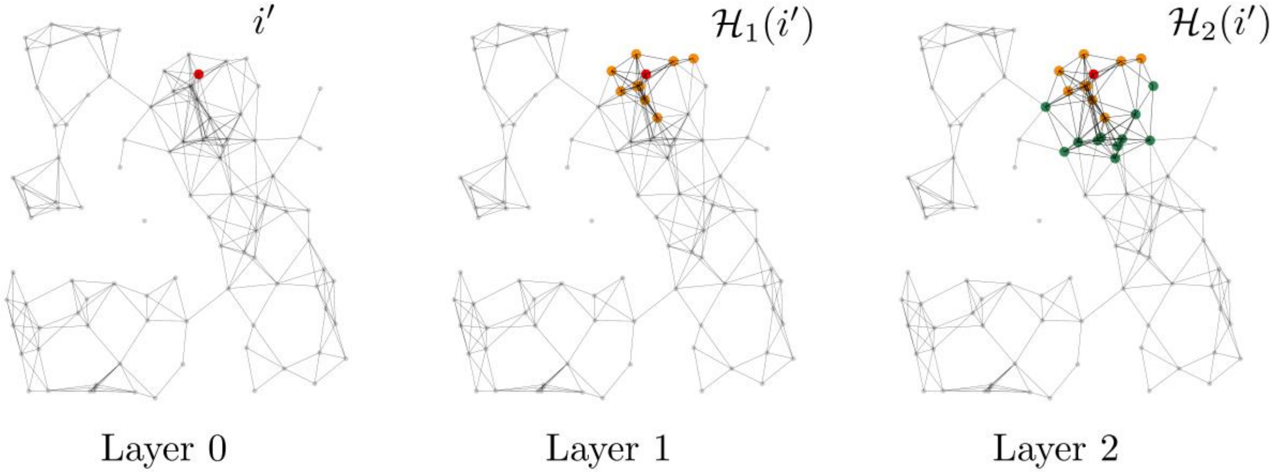


Processing

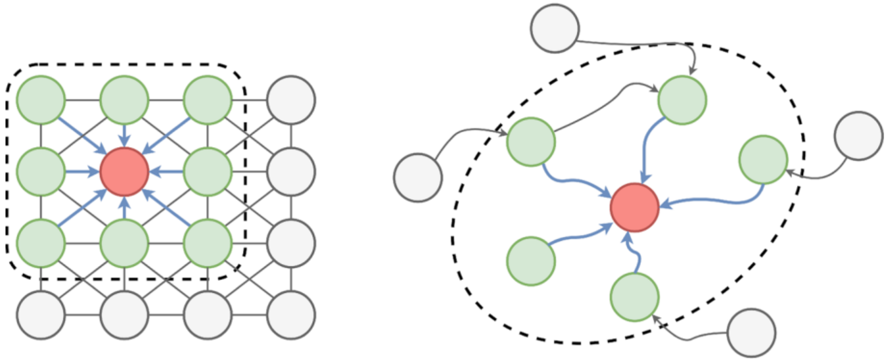
We lose sparsity of events and temporal information!

GCN for event processing

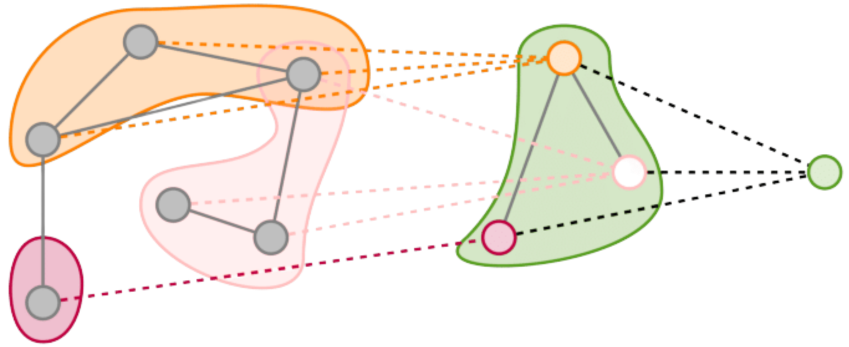
- Temporal information is preserved
- We process sparse data structure
- Recent work presents the possibility of updating graph asynchronously
- Currently only on the GPU!



Asynchronous update



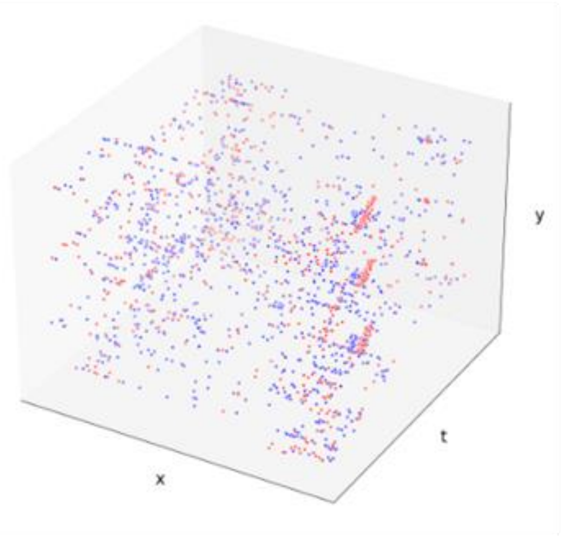
CNN vs GCN



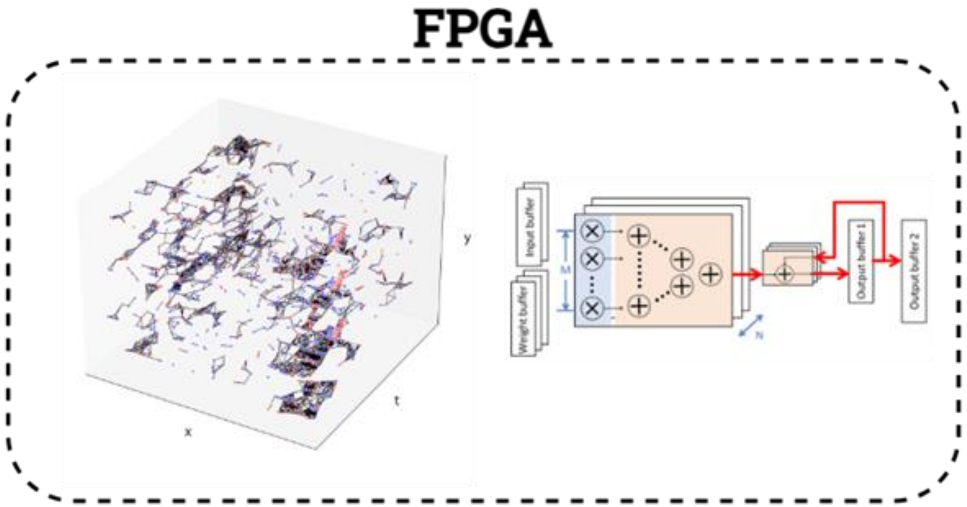
Graph Pooling

Towards Energy-Efficient Vision Systems

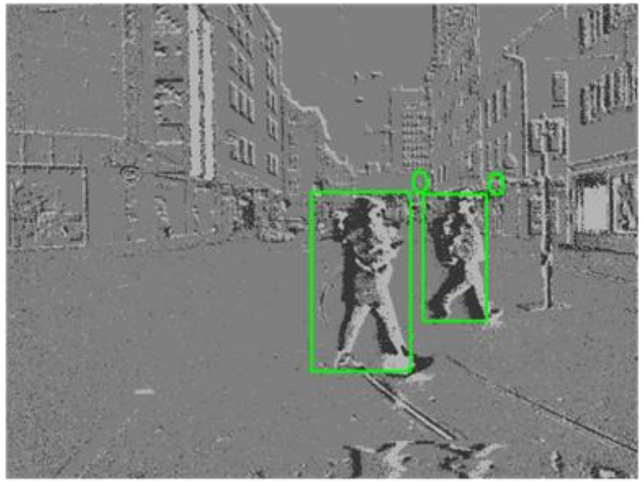
- Combining event cameras and FPGA platforms – energy efficiency
- Combining event cameras and GCN - only important data processing



Raw events



Hardware graph generator and GCN Accelerator



Object detection

Problems with HW graph generation

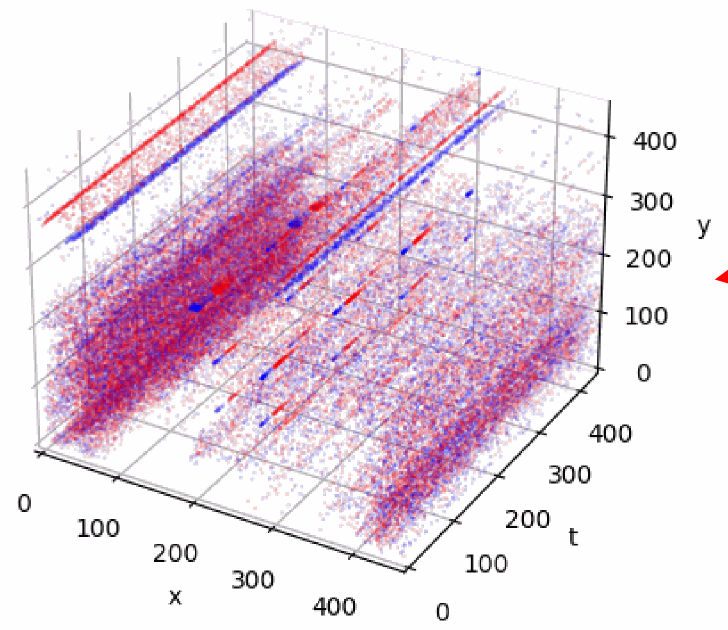
- We do not know where events will be generated
- We do not know how many events will be generated
- We do not know when an event will be generated

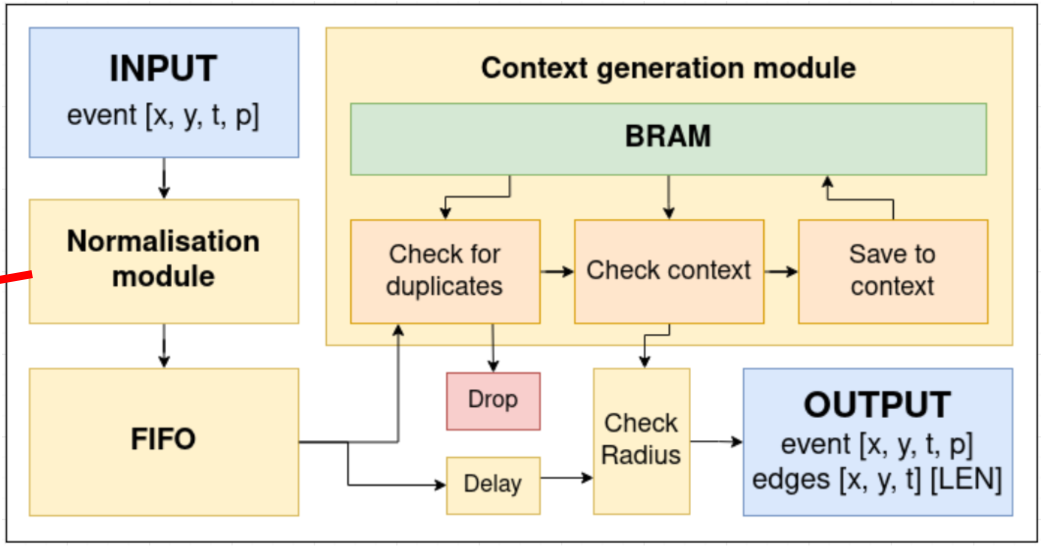
In previous works, graphs are generated by searching for neighbours among all existing vertices, which is inefficient in terms of read/write operations.

To solve this problem, we present hardware module for event-graph generation.

Event Normalisation

- Spatio-temporal event normalisation (e.g. to 256)
- Quantisation to a fixed resolution

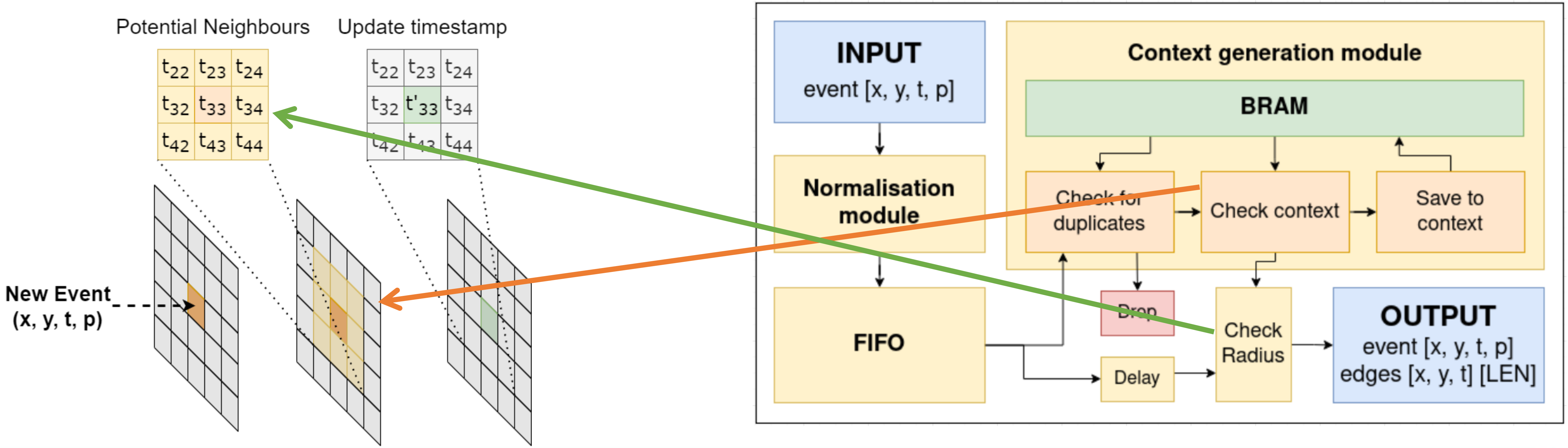





```
register.t <= ((timestamp % TIME_WINDOW)*GRAPH_SIZE/TIME_WINDOW);
register.x <= (x_coord*GRAPH_SIZE/MAX_X_COORD);
register.y <= (y_coord*GRAPH_SIZE/MAX_Y_COORD);
register.p <= polarity;
```

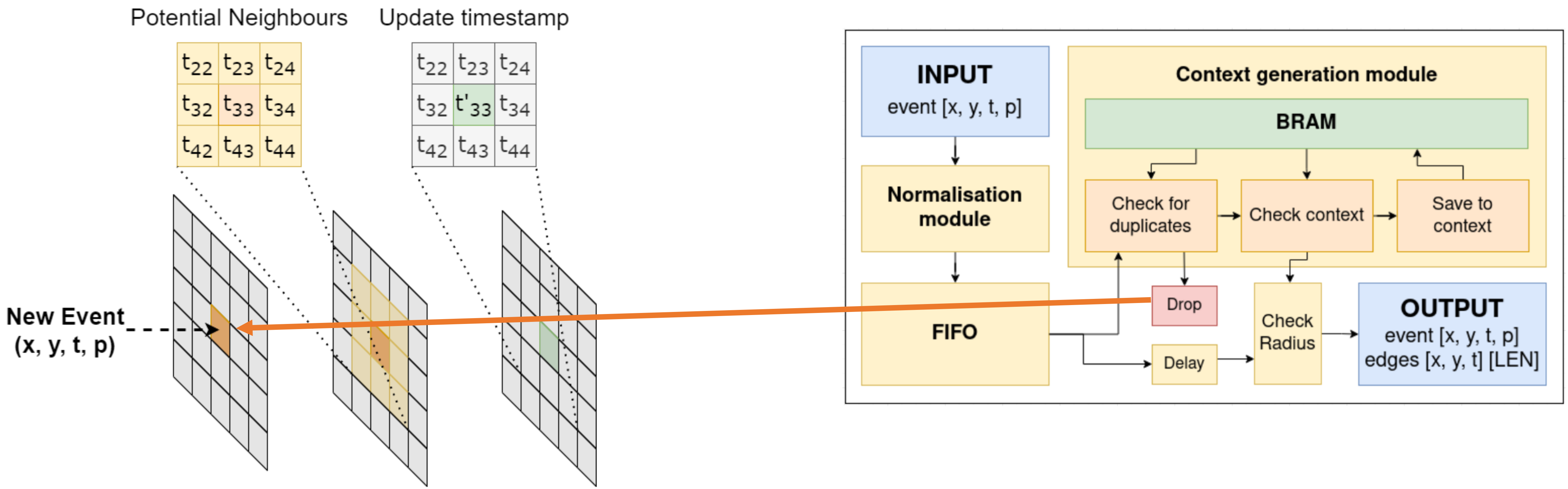
Neighbour matrix

- Neighbours within a specified radius R are searched within a fixed distance.
- New event records its spatial position and time of occurrence in the corresponding cell of the matrix.



Unique events

- Removal of events with the same normalised spatio-temporal position.



if $(x_{33}, y_{33}, t_{33}) == (x'_{33}, y'_{33}, t'_{33})$ *-> drop event*

Results

- **Throughput**: 9.6 events per microsecond, up to 1 million events per second (Mev/s) at a 250 MHz clock rate.
- **Memory Initialisation**: Memory requirements are easy to determine based on normalisation parameters.

Utilisation for ZCU 104 platform with data normalisation to 256

Resource type	Available	Used
LUT	230400	5612 (2.4)
Flip-Flop	460800	950 (0.2)
Block RAM	312	17 (5.5)
DSP	1728	189 (10.9)

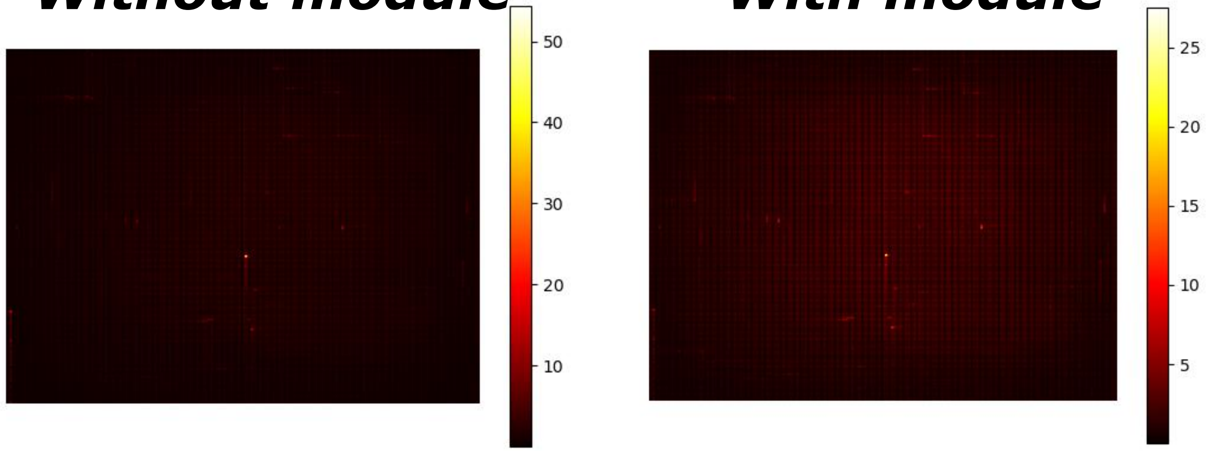
Ablation studies - nodes and edges

- In previous works, the average number of edges was 381,563, achieved by removing over half of the events.
- Using our method, the number of edges is reduced to 65,214, which is 5.85 times smaller, while utilising almost all events (98% of events).

Graph size	Event Preprocessing			Edge Generation		Average Nodes	Average Edges
	Without	Random	Unique	Radius	NM		
128	x			x		56,940	836,304
	x				x		309,711
		x		x		25,000	173,868
		x			x		70,896
			x	x		**50,498**	623,078
			x		x		**262,935**
256	x			x		56,940	138,757
	x				x		67,043
		x		x		25,000	28,892
		x			x		14,051
			x	x		**55,925**	131,648
			x		x		**65,214**

Hot-pixel influence

Without module ***With module***



Ablation studies - accuracy

- The module shows minimal impact on accuracy during training, with only 0.08% mean Average Precision (mAP) drop on the N-Caltech dataset.

Method	Normalisation	Pre-processing	Edge Generation	mAP
Base model[6]	Only time = 100	Random	Radius	53.47
Normalisation value	64	Random	Radius	43.73
	128	Random	Radius	51.99
	256	Random	Radius	54.77
Direction of edges	128	Random	NM - undirected	52.20
	128	Random	NM - time directed	51.40
	128	Random	NM - op. directed	50.23
Unique values	128	Unique	NM - time directed	**51.28**
	256	Unique	NM - time directed	**53.39** ←

Summary

- **Hardware Module**: Presented a hardware module for event data graph generation, optimised for Graph Convolutional Networks (GCNs).
- **Performance**: Achieved a throughput of up to 1 million events per second, maintaining high accuracy.

Future work

- **GCN Accelerator**: Planning to develop a hardware accelerator for the GCN model.
- **Vision System**: Aiming for integration with the generator module to create a comprehensive end-to-end vision system for object detection.