Efficient Event-Based Object Detection

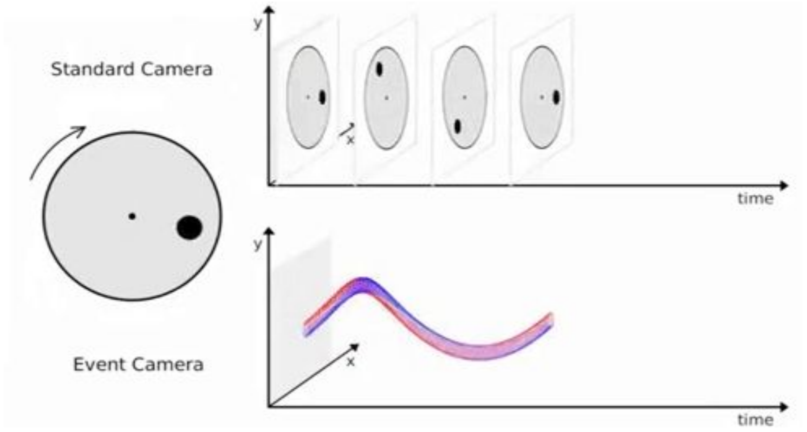
A Hybrid SNN -ANN Neural Network with Spatial and Temporal Attention

Soikat Hasan Ahmed*, Jan Finkbeiner†, Emre Neftci

CVPR Nashville JUNE 11-15, 2025

JÜLICH Forschungszentrum Wandel gestalten

Event Camera vs Frame Camera



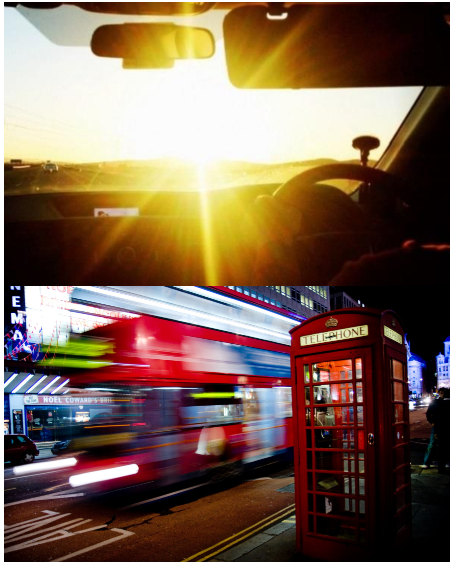
Event camera

Event camera outputs

Standard camera outputs

Reference: https://youtu.be/Qrnpj2FD1e4

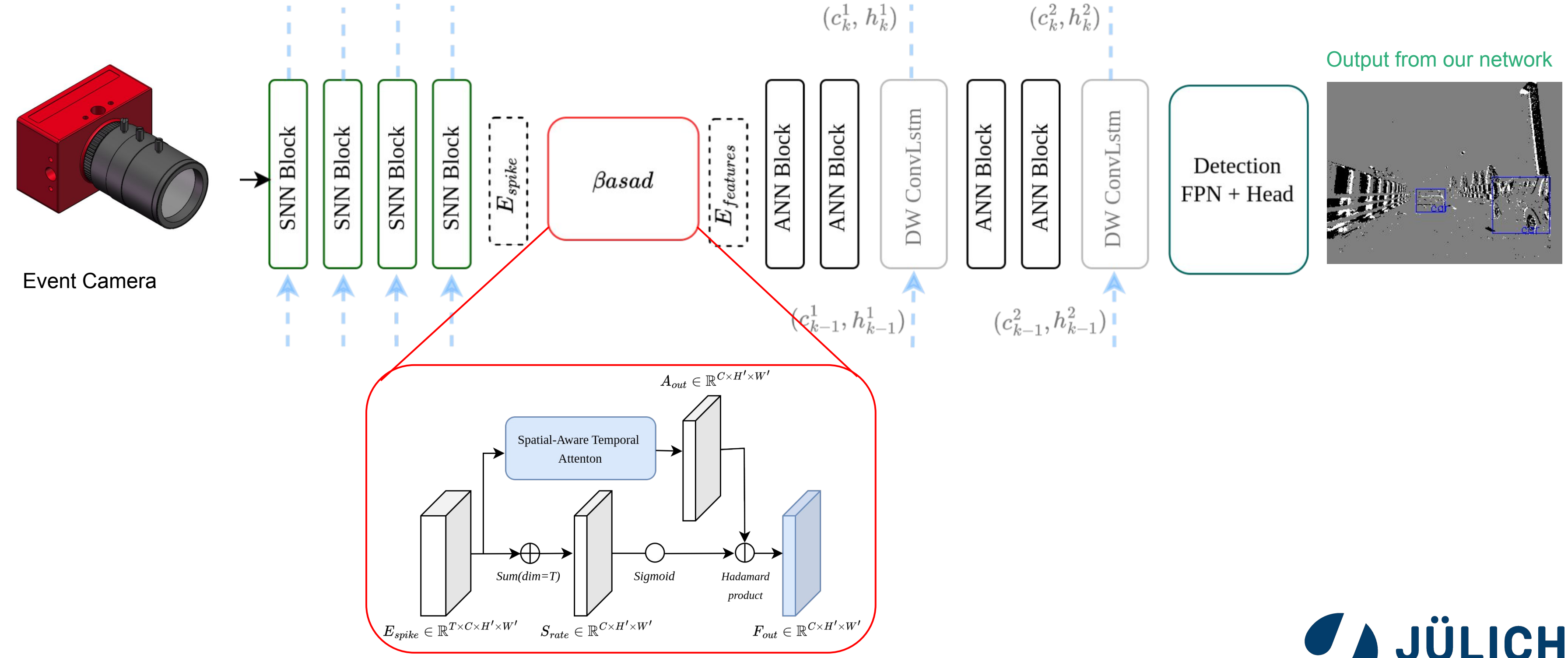
Event Camera vs Frame Camera



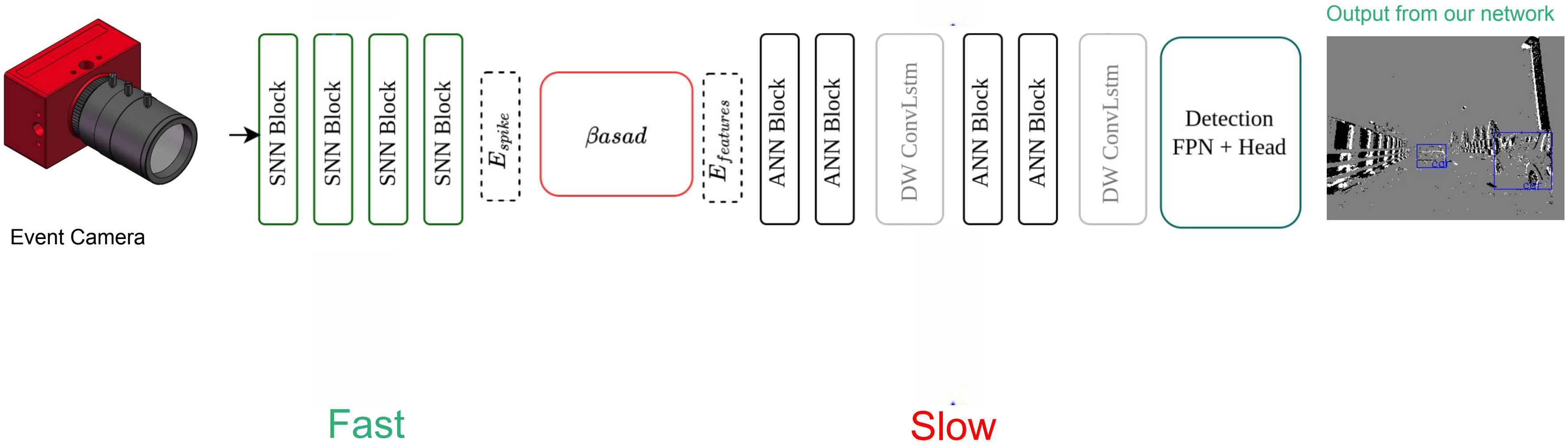
Source: Prophesee

Source: 1, 2

Attention-based Hybrid SNN-ANN Network

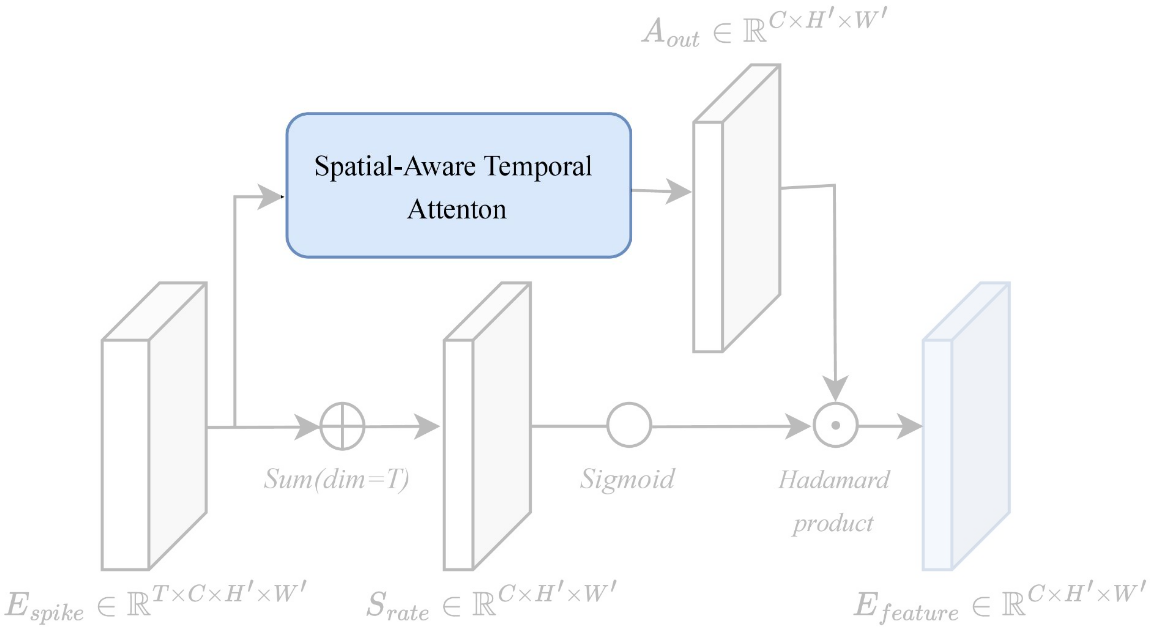


Attention-based Hybrid SNN-ANN Network



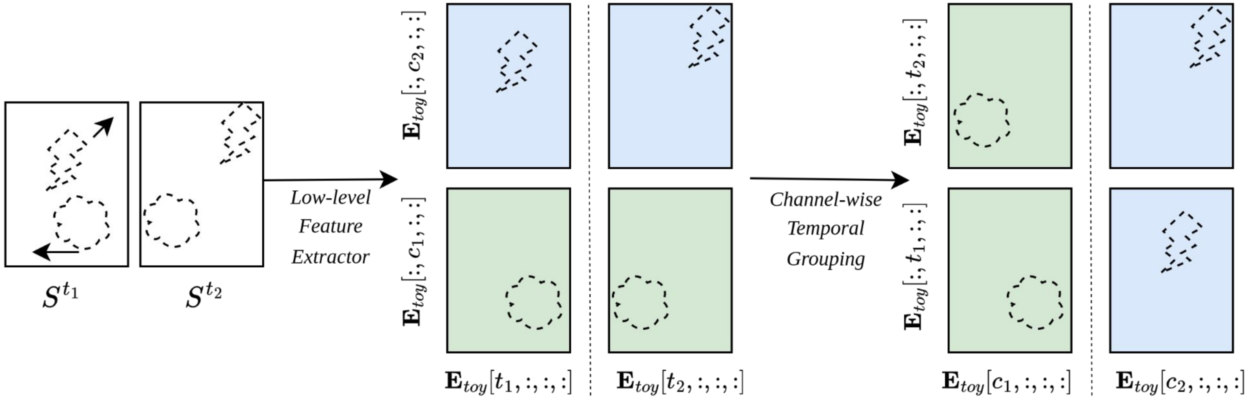
Attention-based SNN-ANN bridge module

Spatial-Aware Temporal Attention and Event-rate Spatial Attention



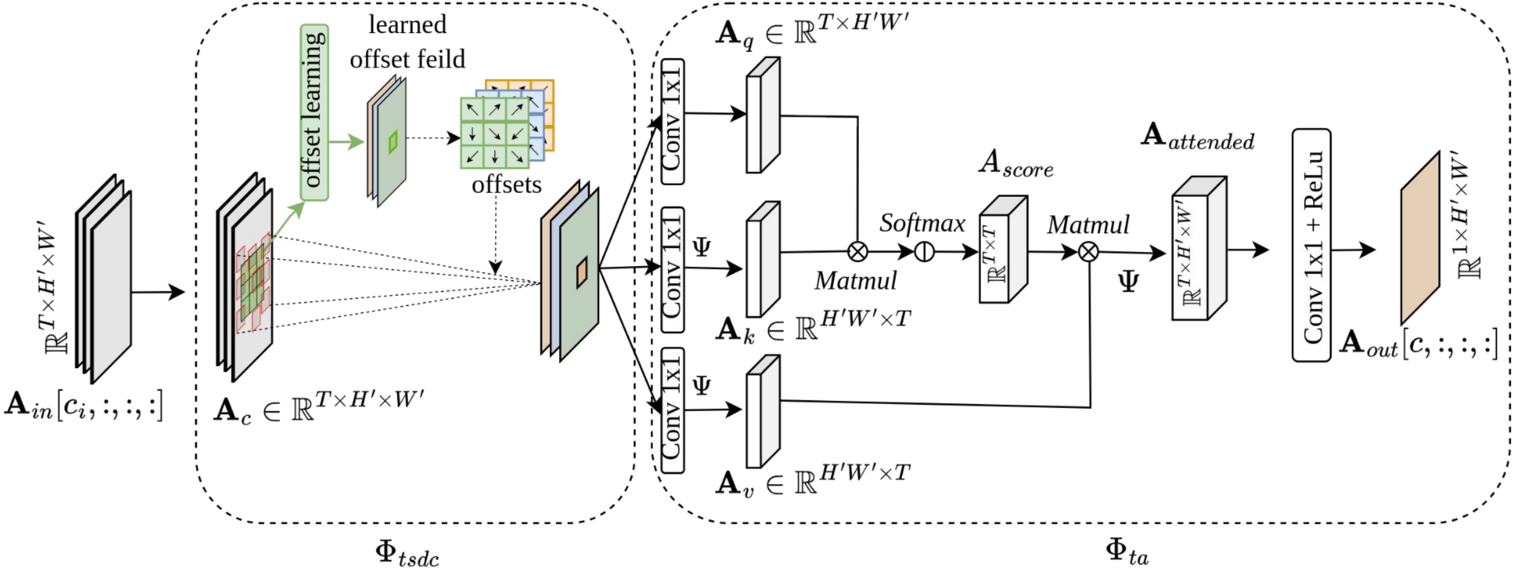
Spatial-Aware Temporal Attention

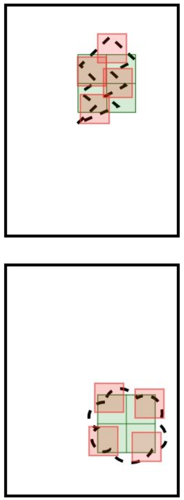
Channel-wise Temporal Grouping (toy example)



Spatial-Aware Temporal Attention

Time-wise Separable Deformable Convolution with Temporal Attention



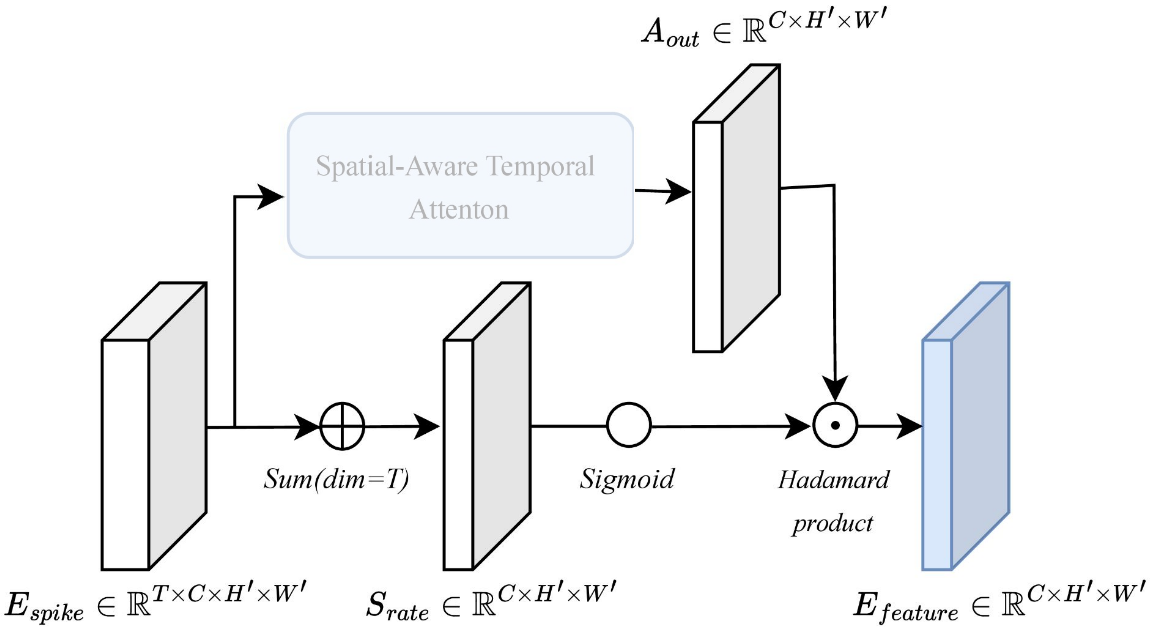


Deformable kernel example

JÜLICH Forschungszentrum

Attention-based SNN-ANN bridge module

Spatial-Aware Temporal Attention and Event-rate Spatial Attention



Performance Comparison

CVPR Nashville JUNE 11-15, 2025

On Prophesee benchmark datasets

Models	Type	Params	Gen 1 mAP	Gen 4 mAP
AEGNN [35]	GNN	20M	0.16	-
SparseConv [30]	ANN	$133M$	0.15	-
Inception + SSD [18]	ANN	$> 60M^*$	0.3	0.34
RRC-Events [5]	ANN	$> 100M^*$	0.31	0.34
Events-RetinaNet [33]	ANN	33M	0.34	0.18
E2Vid-RetinaNet [33]	ANN	44M	0.27	.25
RVT-B W/O LSTM [14]	Transformer	$16.2M^*$	0.32	-
Proposed	Hybrid	6.6M	0.35	.27

Comparison with ANN-based approaches

Models	Type	Params	mAP
VGG-11+SDD [6]	SNN	13M	0.17
MobileNet-64+SSD [6]	SNN	24M	0.15
DenseNet121-24+SSD [6]	SNN	8M	0.19
FP-DAGNet[45]	SNN	22M	0.22
EMS-RES10 [39]	SNN	6.20M	0.27
EMS-RES18 [39]	SNN	9.34M	0.29
EMS-RES34 [39]	SNN	14.4M	0.31
SpikeFPN [46]	SNN	22M	0.22
Proposed	Hybrid	6.6M	0.35

Comparison with SNN-based approaches

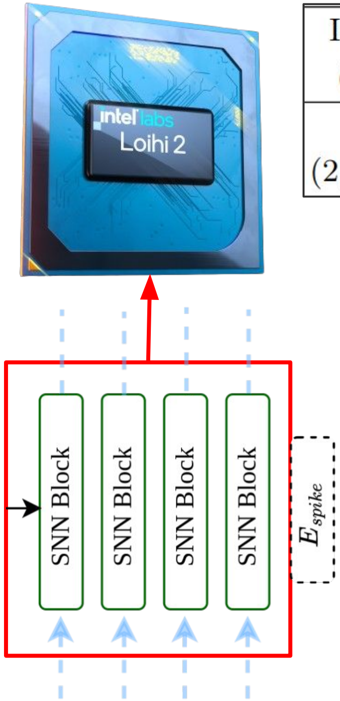
Models	Type	Params	mAP
S4D-ViT-B [48]	TF + SSM	16.5M	0.46
S5-ViT-B [48]	TF + SSM	18.2M	0.48
S5-ViT-S [48]	TF + SSM	9.7M	0.47
RVT-B [14]	TF + RNN	19M	0.47
RVT-S [14]	TF + RNN	10M	0.46
RVT-T [14]	TF + RNN	4M	0.44
ASTMNet [25]	(T)CNN + RNN	100M	0.48
RED [33]	CNN + RNN	24M	0.40
Proposed+RNN	Hybrid + RNN	7.7M	0.43

Comparison with RNN-based approaches

Hardware Implementation

Implementation of the SNN blocks in Intel Loihi 2 - implemented by Jan Finkbeiner

Image by Intel.



Input size (C,W,H)	Weight qunatization	Number of chips	Total Power [W]	Execution Time Per Step [ms]
(2, 256, 160)	int8	6	1.73 ± 0.10	2.06 ± 0.74
	int6	6	1.71 ± 0.11	2.06 ± 0.74

Models	mAP(.5)	mAP(.5:.05:.95)
Variant 1 (float16)	0.613	0.348
Variant 2 (int8)	0.612	0.349
Variant 3 (int6)	0.612	0.348
Variant 4 (int4)	0.610	0.347
Variant 5 (int2)	0.432	0.224

Model accuracy in different quantization level

Efficiency

- Power consumption was analyzed following [2], demonstrating that the proposed method achieves a better balance between accuracy and energy efficiency.
- **MACs** and **ACs** were compared across different network variants

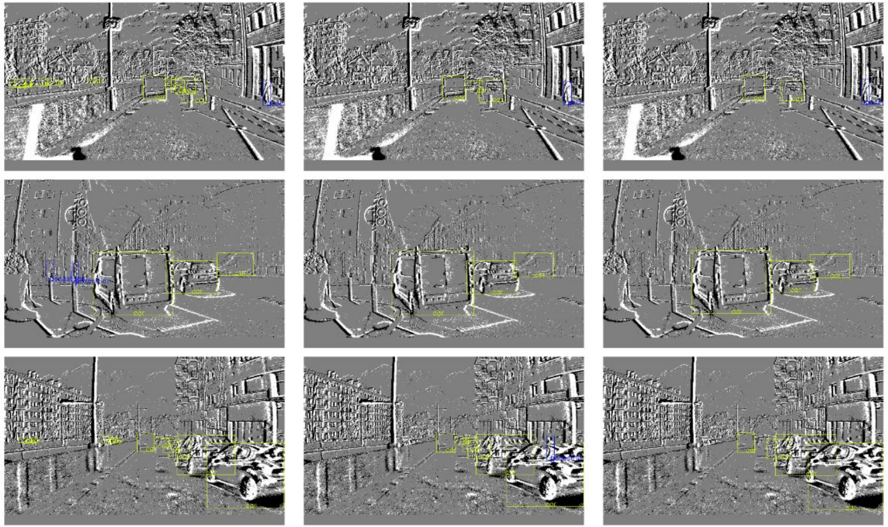
Models	mAP(.5)	MACs	ACs
$Baseline_{ann}$	0.61	$15.34e9$	0.0
$Baseline_{w/o\,\beta_{asab}}$	0.53	$1.18e9$	$0.97e9$
$\mathbf{Proposed_{w/\beta_{asab}}}$	0.61	$\mathbf{1.63e9}$	$\mathbf{0.97e9}$
$Proposed_{snn+}$	0.58	$0.87e9$	$1.59e9$

Models	Gen 1/Gen 4 mAP	MACs / ACs	Energy [mJ]
VGG-11+SDD	0.17 / -	0.0 / $11.1e9$	4.2
MobileNet-64+SSD	0.15 / -	0.0 / $4.3e9$	1.6
DenseNet121+SSD	0.19 / -	0.0 / $2.3e9$	0.9
Inception + SSD	0.3 / 0.34	$11.4e9^*$ / 0.0	19.3
Events-RetinaNet	0.34 / 0.18	$3.2e9^*$ / 0.0	5.4
E2Vid-RetinaNet	0.27 / .25	$> 3.2e9^*$ / 0.0	> 5.4
RVT-B W/O LSTM	0.32 / -	$2.3e9$ / 0.0	3.9
Proposed	0.35 / .27	1.6e9 / $1.0e9$	3.1

Ablation study

- An in-depth ablation study was conducted for each component of the proposed **ASAB** module, along with various configurations of the hybrid architecture.

Models	mAP(.5)	mAP
Variant 1(w/o - Φta)	0.57	0.33
Variant 2 (w/o deform)	0.59	0.34
Variant 3 (w/o - ESA)	0.59	0.34
Variant 4 (w/o - ASAB)	0.53	0.30
Variant 5 (Proposed)	0.61	**0.35**



Without ASAB module   With ASAB module   Ground Truth

SCAN FOR MORE DETAILS

THANK YOU FOR YOUR ATTENTION

JÜLICH
Forschungszentrum
Wandel gestalten