Technische Hochschule Nürnberg

Open Neuromorphic

SPICES Lab

Hybrid Learning for Event-Based Visual Motion Detection and Tracking of Pedestrians

Cristian Axenie
Nuremberg Institute of Technology, Germany

ohm

History, facts, and figures

A renowned patron

Proud of our roots

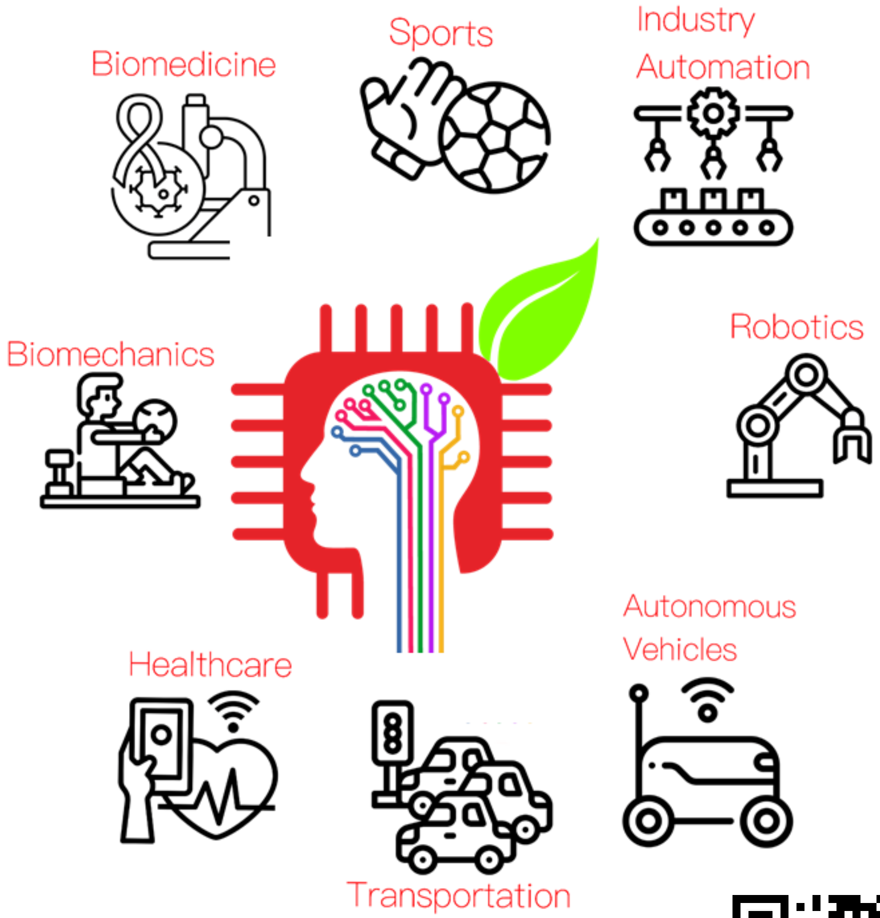
The university's roots (history) can be traced back to the year 1823 and the founding of the Städtische Polytechnikum (Municipal Polytechnic) – the oldest of our forerunners.

A renowned patron

The university is named after the world-renowned physicist, Georg Simon Ohm (biography), who was a physics and mathematics professor in Nuremberg between 1833 and 1849, and also fulfilled the role of rector.

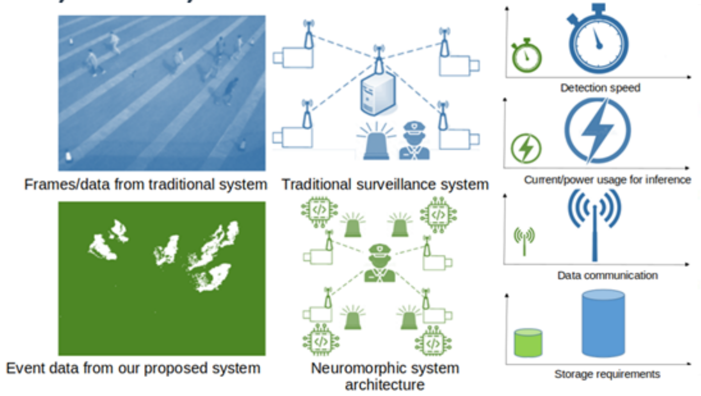
The famous omega

The Greek letter omega in the university's logo is a tribute to Georg Simon Ohm's greatest discovery – his law concerning electrical resistance.

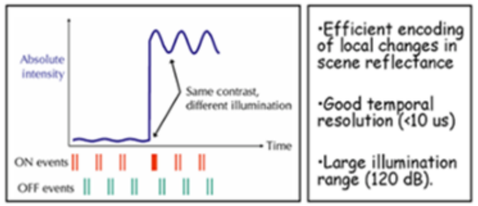


Sensorimotor Processing, Intelligence, and Control in Efficient compute Systems (SPICES) Lab

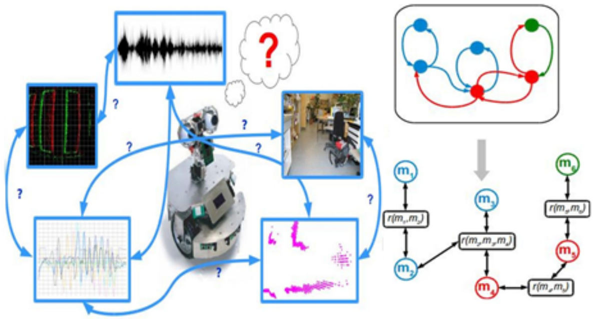
Fast, Efficient, Robust and Green Inference



Spike-based/Event-based data representation

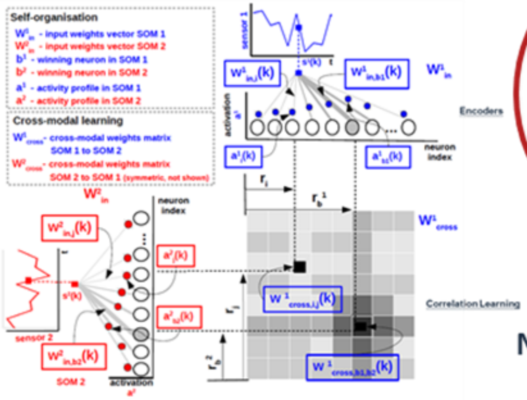


Sensor Fusion



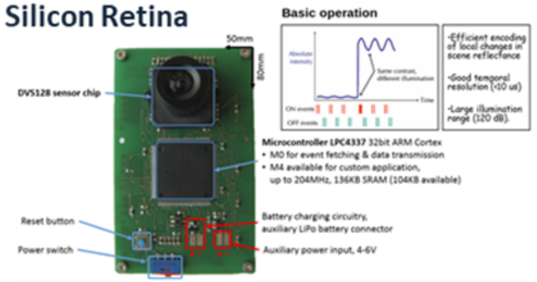
Closed-loop Sensorimotor Control

Learning algorithms



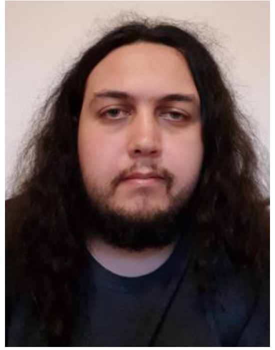
Neuromorphic Chips

Silicon Retina



ohm

Team



Ertan Halilov

Julian Main

David Weiss

Cristian Axenie

SPICES Lab

Outline

- Goal
- Solution overview
- Sensing and algorithmics
- Performance evaluation
- Deployment
- BOM

ohm

Goal

“**Vision Zero**” as a street safety policy that strives for the elimination of traffic fatalities for all transportation modes.

McKee Road & Jackson Avenue, San Jose, California

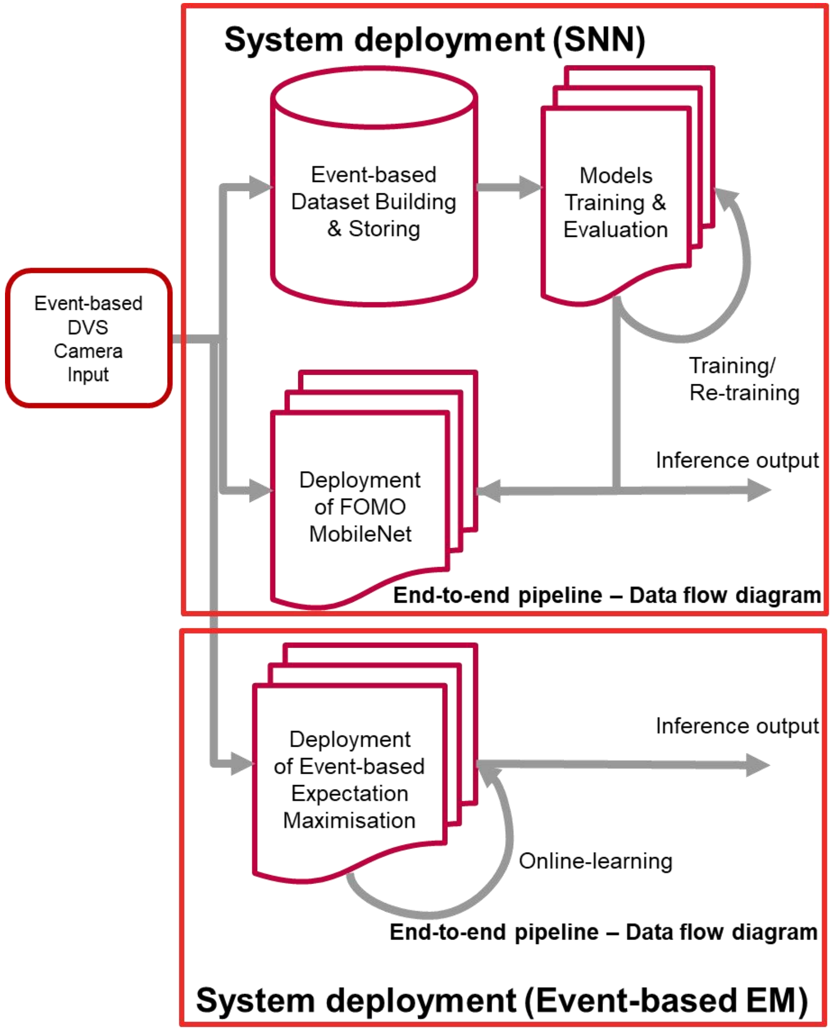
Tully Road & La Ragione Avenue, San Jose, California

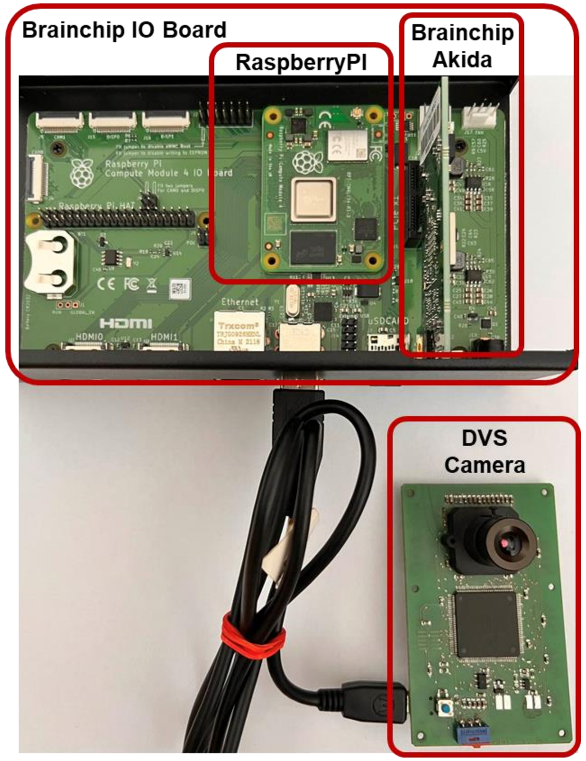
Cost effective and **accurate solutions** are needed to **detect pedestrians** during the **day** and especially at **nighttime** to implement safety measures. **Solutions** need to have a **very good energy footprint**, **robustness**, and a **budget** that allows scaling to city level.

CITY OF SAN JOSE
CAPITAL OF SILICON VALLEY

VZSJ
A S A P !

Solution overview

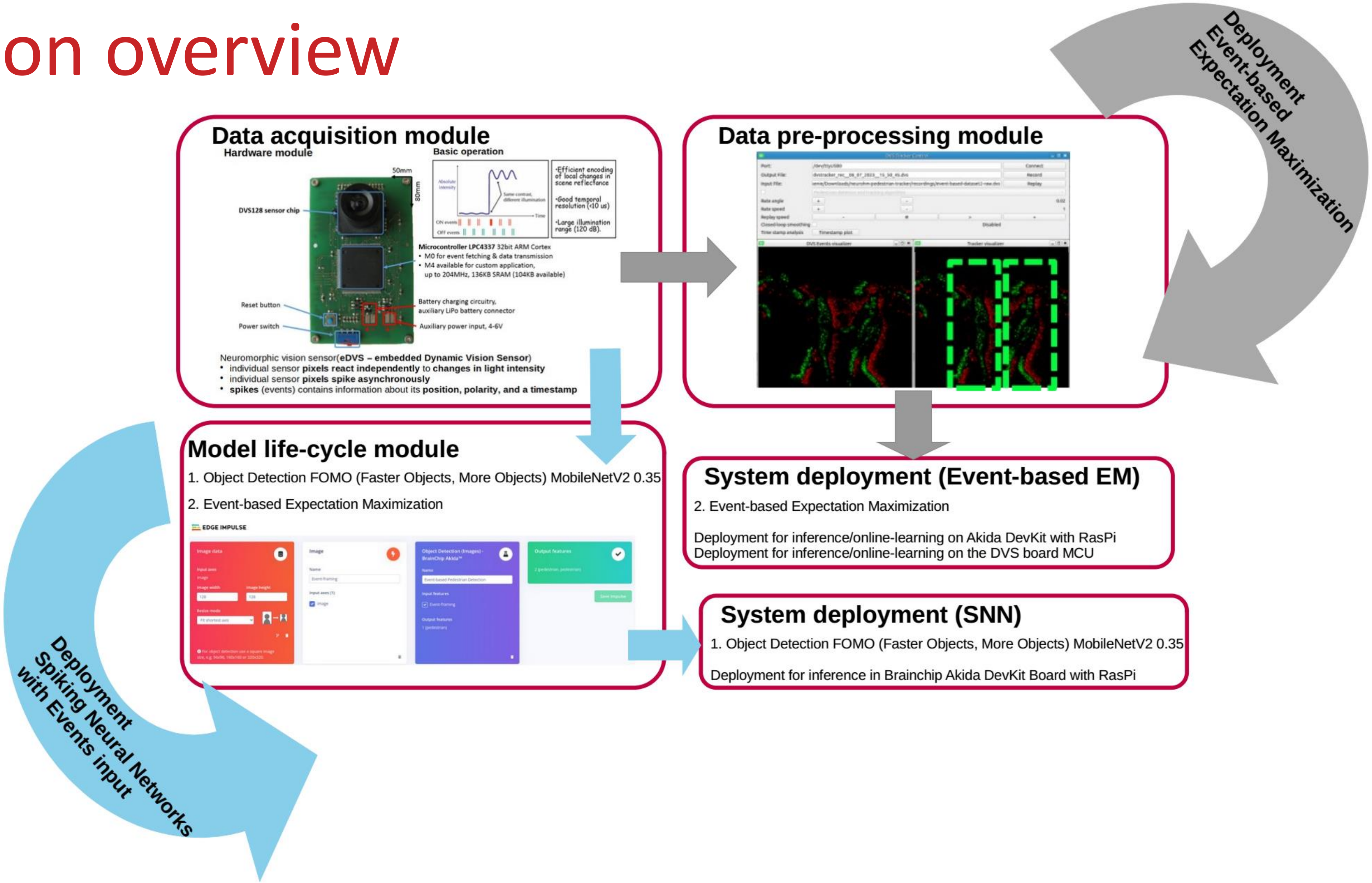




End-to-end pipeline – System hardware

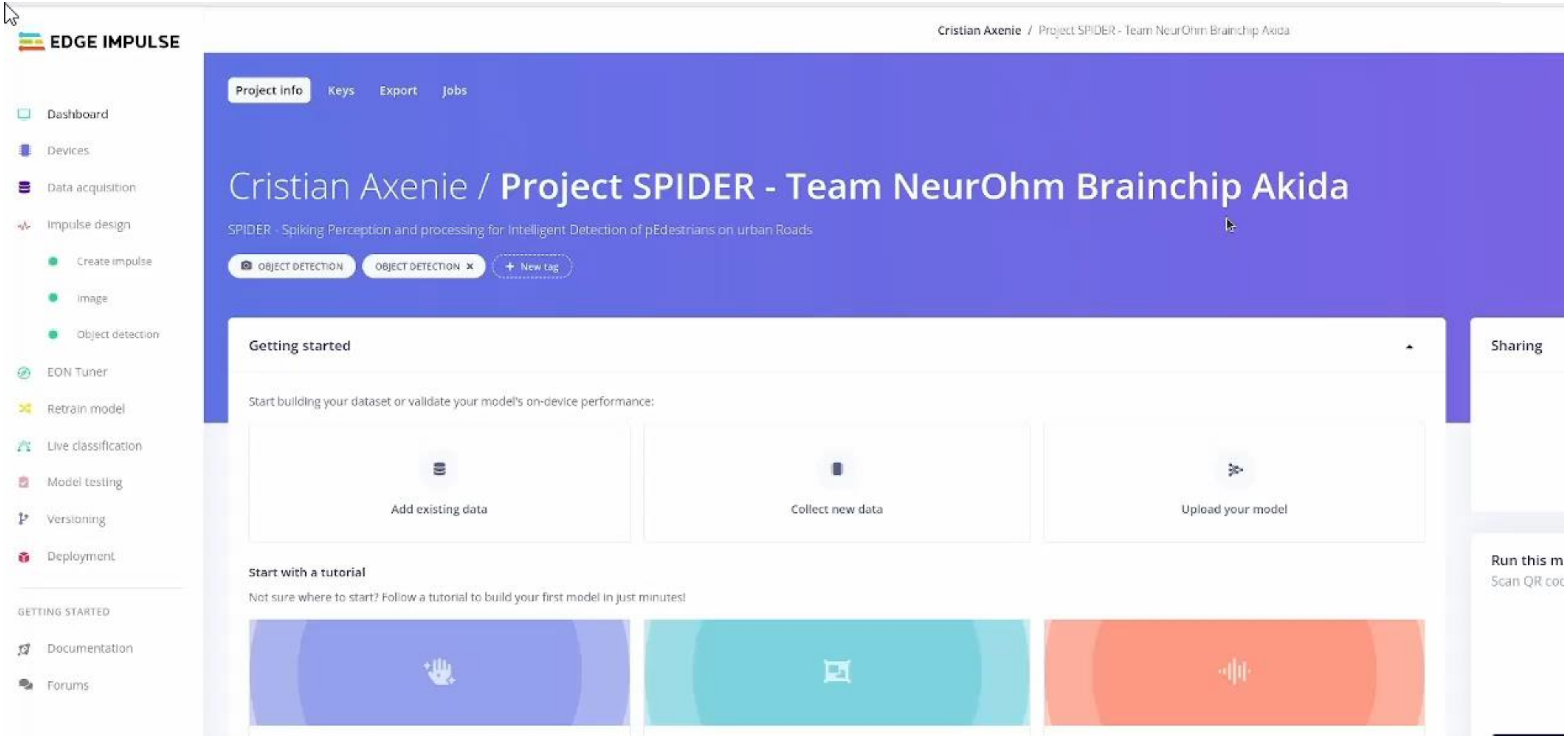
ohm

Solution overview



ohm

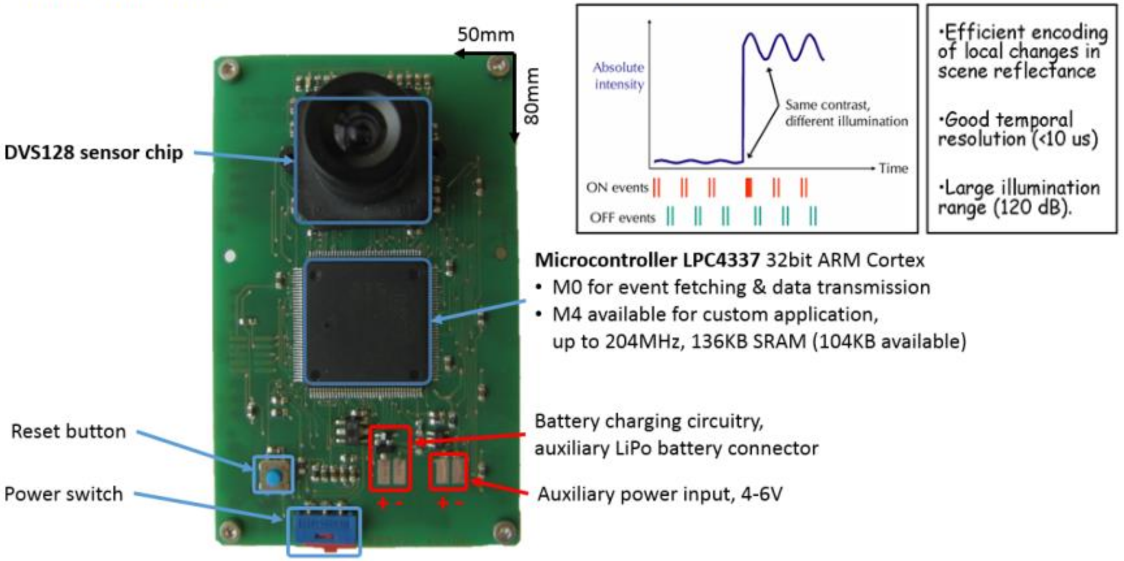
Solution development and life-cycle



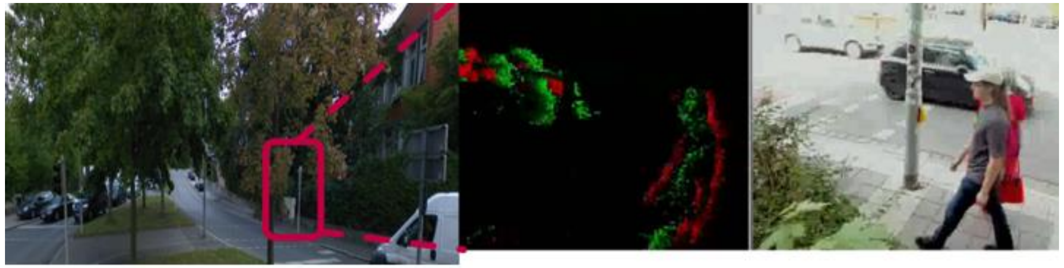
Approach

- Sensing

Hardware module **Basic operation**

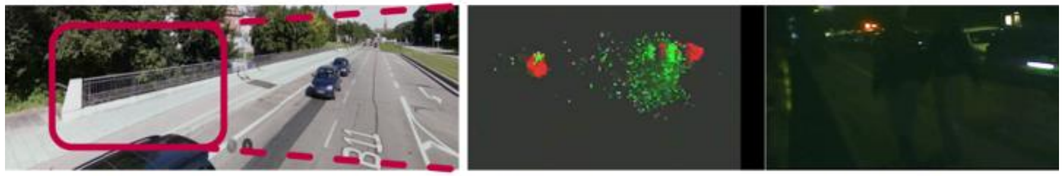


Neuromorphic vision sensor(**eDVS – embedded Dynamic Vision Sensor**)
- individual sensor **pixels react independently** to **changes in light intensity**
- individual sensor **pixels spike asynchronously**
- **spikes** (events) contains information about its **position, polarity, and a timestamp**



Urban intersection dataset (Nuremberg, Germany) daytime

Wide urban road dataset (Nuremberg, Germany) daytime



Wide urban road dataset (Munich, Germany) night-time

NCS HTTP://NEUROCOMPUTING.SYSTEMS CST - EECS - KTH STOCKHOLM, SWEDEN

ohm

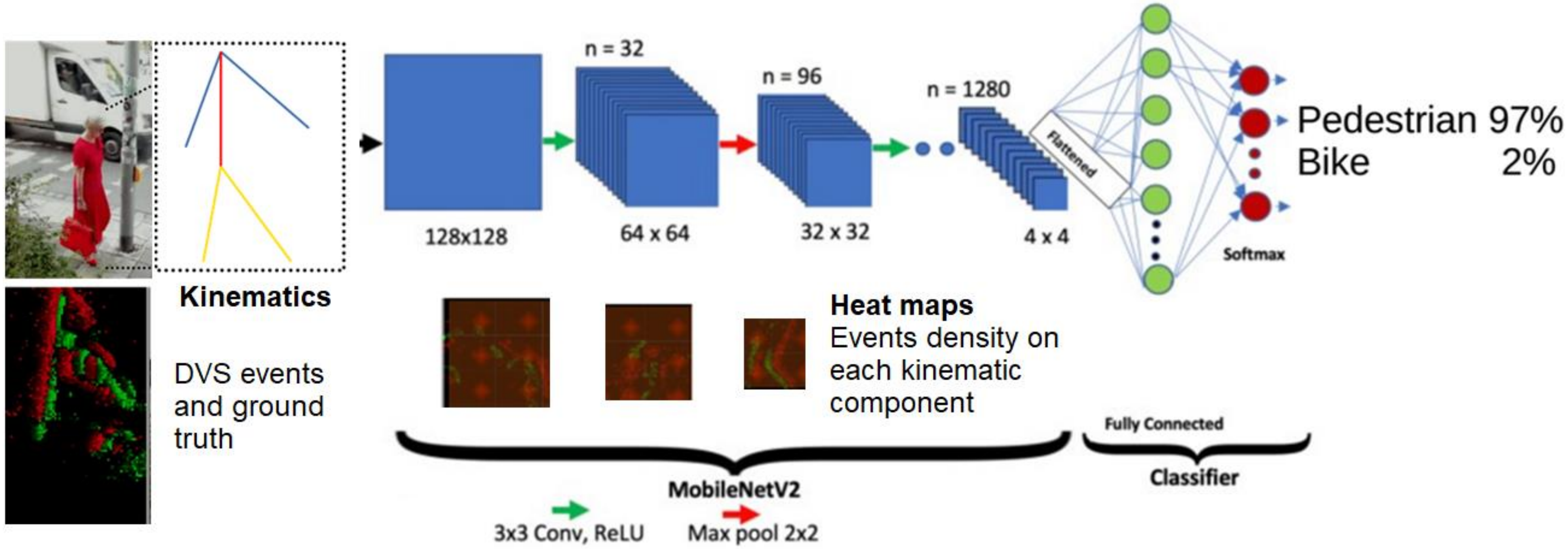
Approach

- Sensing

ohm

Approach

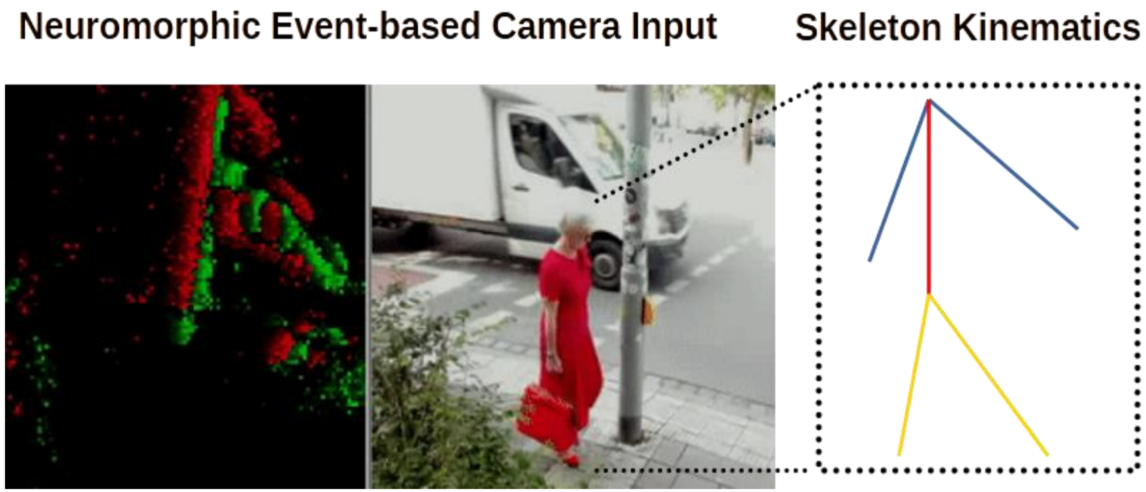
- Algorithmics – Fast Objects More Objects (FOMO) ConvNet, Spiking Neural Network (SNN)

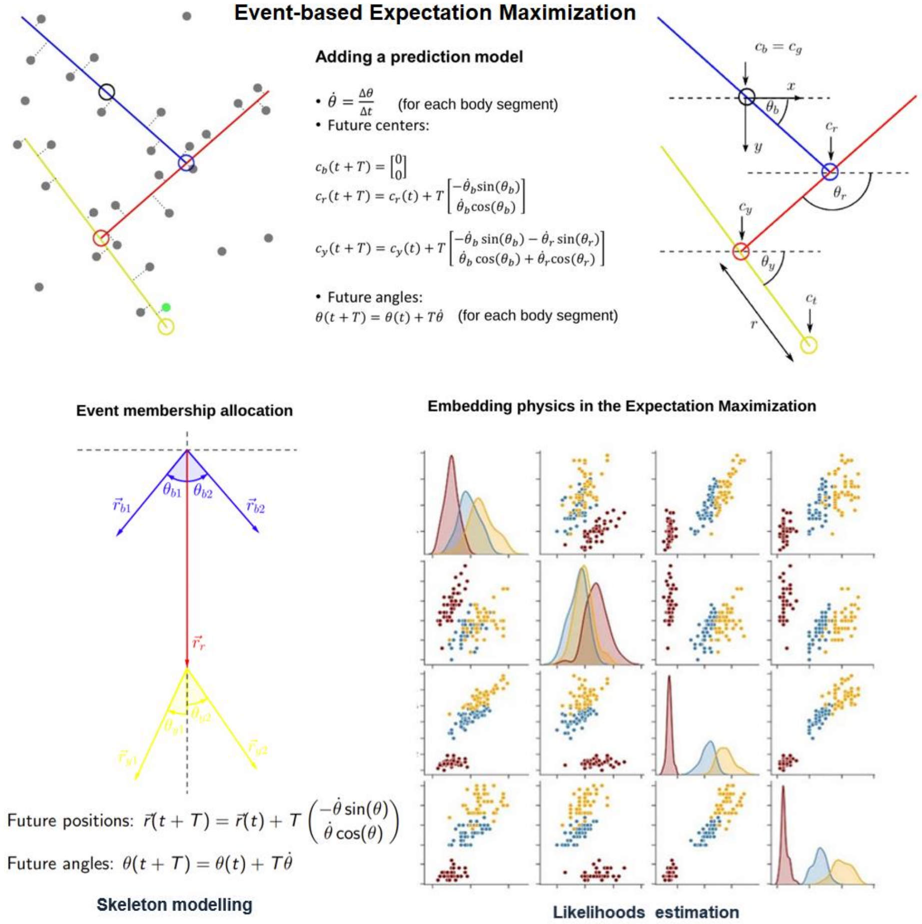


Demo - Spiking Neural Network (SNN)

Approach

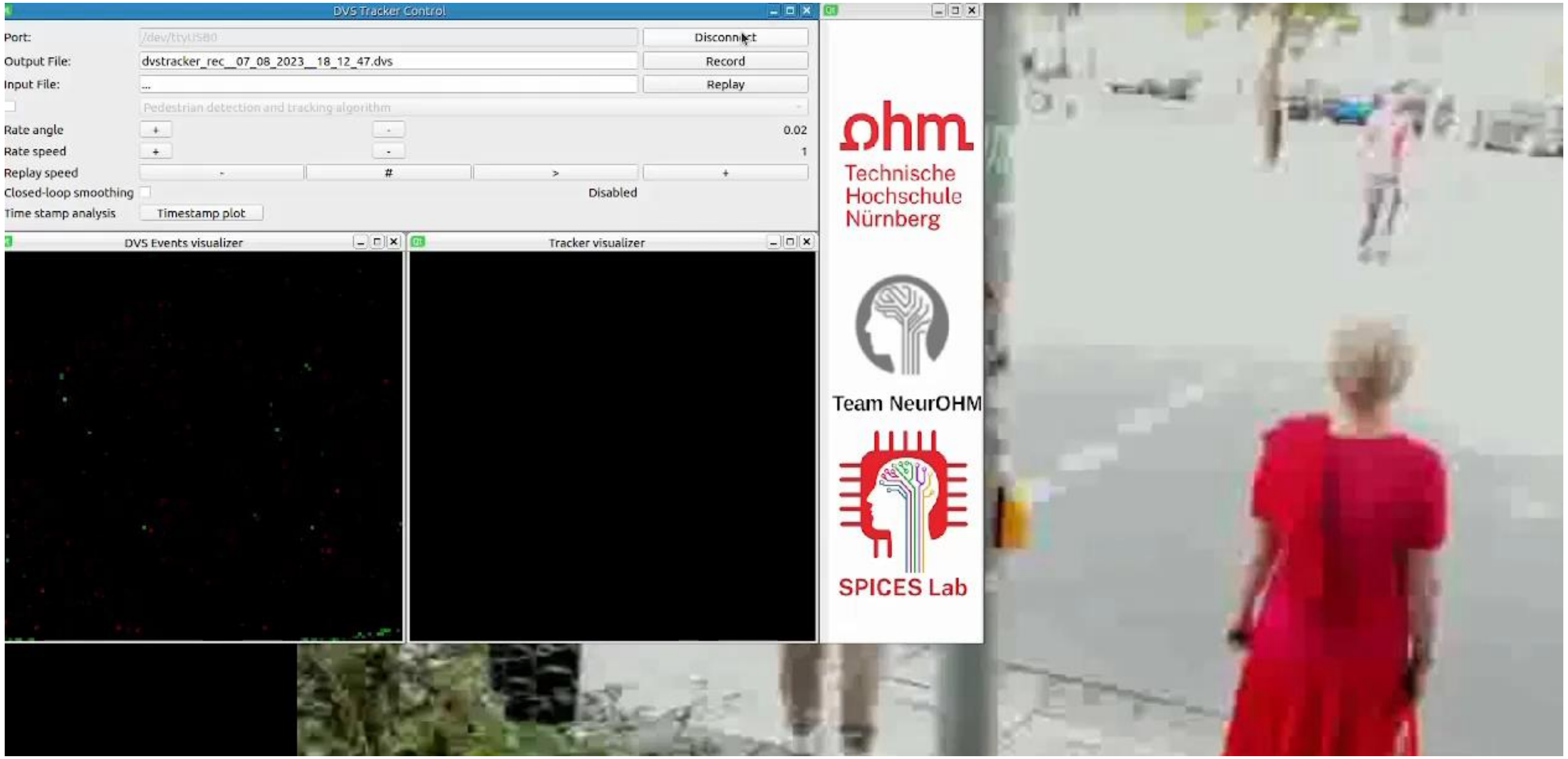
- Algorithmics – Event-based Expectation Maximization (EM)





UserFunction.cpp

Demo - Event-based Expectation Maximization (EM)

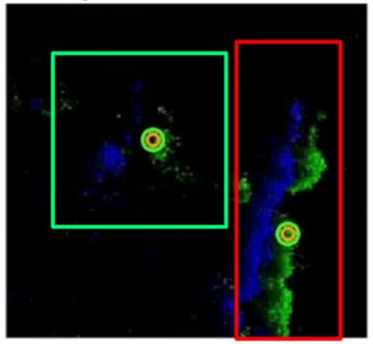


Performance

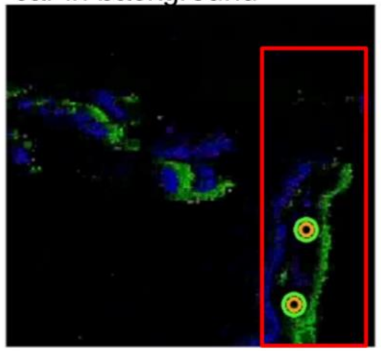
- Spiking Neural Network

Qualitative evaluation

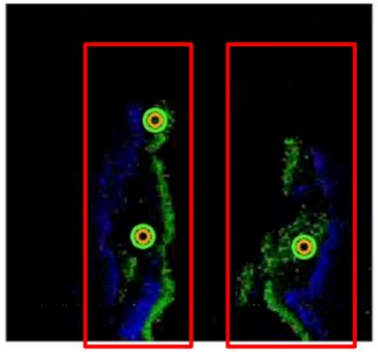
Detection: 1 pedestrian and 1 bicyclist



Detection: 1 pedestrian and car in background



Detection: 2 pedestrians



* For the accuracy evaluation of both the detection and tracking components of the system, we performed multiple experiments in order to get a statistically valid analysis. For the quantitative evaluation, we executed the following protocol: 1) Read relevant detection and tracking data from each of the experiments (i.e. N = 30 experiments); 2) Perform statistical tests (i.e., a combination of omnibus ANOVA and posthoc pairwise T-test with a significance p = 0.05) and adjust the ranking of experiments depending on significance; 3) rank subsets of relevant metrics (i.e., the metrics with a significant difference, e.g., the F1 score for detection and four others tracking specific metrics).

Quantitative evaluation

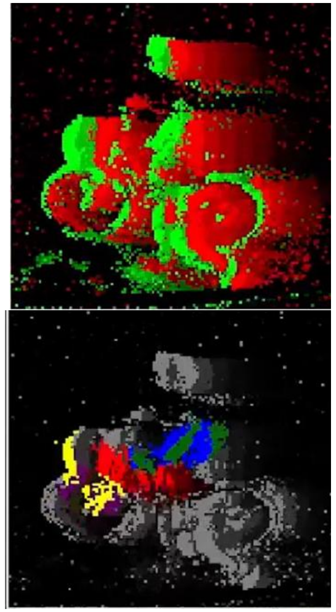
Dataset	Background %	Bicyclist %	Pedestrian %
Dataset 1 (daytime)			
Background	99.70	0.26	0.04
Bicyclist	12.10	87.90	0.00
Pedestrian	6.20	0.00	93.80
F1-Score	1.00	0.62	0.77
Dataset 2 (daytime)			
Background	97.50	2.20	0.80
Bicyclist	10.10	89.90	0.00
Pedestrian	3.20	0.00	96.80
F1-Score	1.00	0.70	0.87
Dataset 3 (night)			
Background	99.07	0.30	0.00
Bicyclist	21.10	78.00	0.90
Pedestrian	10.20	13.00	76.80
F1-Score	0.90	0.60	0.70

Performance

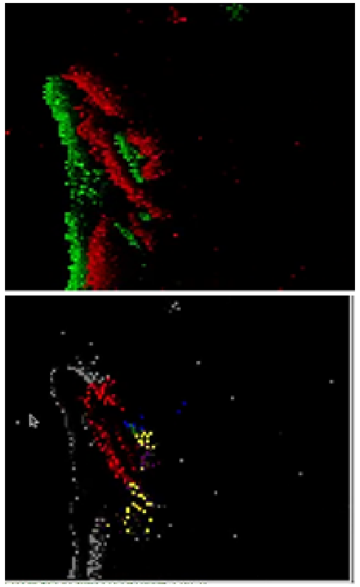
- Event-based Expectation Maximization

Qualitative evaluation

Bicyclist tracking



Pedestrian tracking



* For the accuracy evaluation of both the detection and tracking components of the system, we performed multiple experiments in order to get a statistically valid analysis. For the quantitative evaluation, we executed the following protocol: 1) Read relevant detection and tracking data from each of the experiments (i.e. N = 30 experiments); 2) Perform statistical tests (i.e., a combination of omnibus ANOVA and posthoc pairwise T-test with a significance p = 0.05) and adjust the ranking of experiments depending on significance; 3) rank subsets of relevant metrics (i.e., the metrics with a significant difference, e.g., the F1 score for detection and four others tracking specific metrics).

Quantitative evaluation

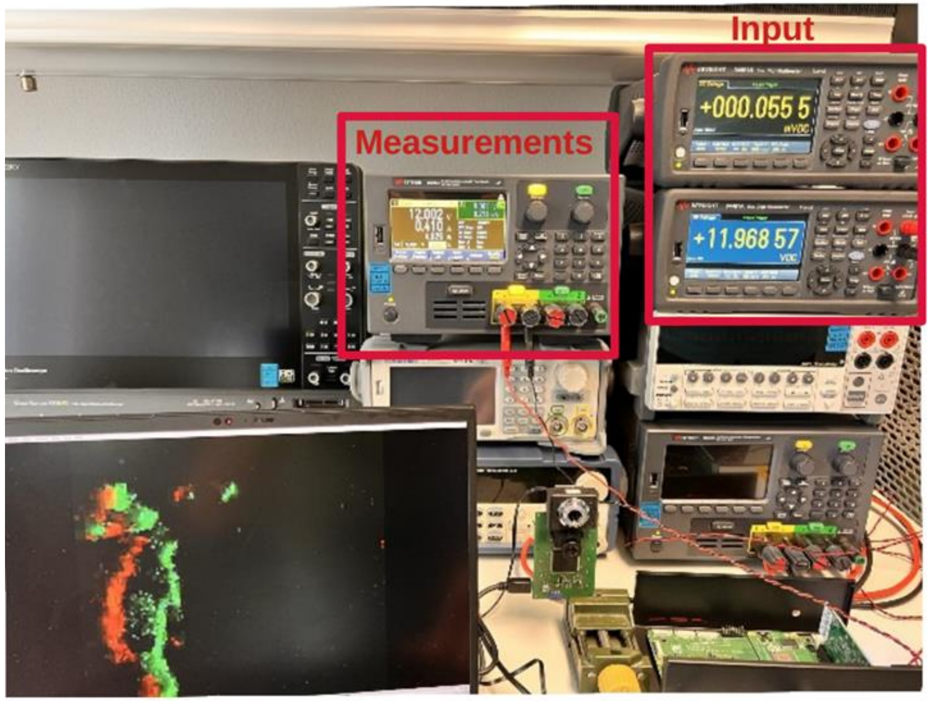
Dataset	Bicyclist	Pedestrian
Dataset 1 (daytime)		
Track Matching Error(%)	13.70	10.10
Tracking Time Delay(s)	0.08	0.03
Tracking Detection Rate(%)	95.00	98.00
Tracking Completeness(s)	0.38	0.25
Dataset 2 (daytime)		
Track Matching Error(%)	11.20	8.21
Tracking Time Delay(s)	0.07	0.02
Tracking Detection Rate(%)	97.00	99.00
Tracking Completeness(s)	0.38	0.25
Dataset 3 (night)		
Track Matching Error(%)	23.30	20.10
Tracking Time Delay(s)	0.09	0.08
Tracking Detection Rate(%)	76.00	79.00
Tracking Completeness(s)	0.84	0.76

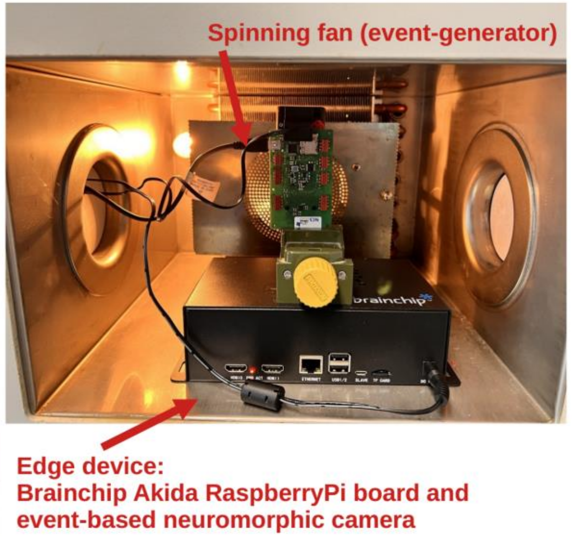
ohm

Deployment evaluation

- Power consumption & weatherization analysis

Measurement setup
Camera pointing to a screen with recorded traffic data





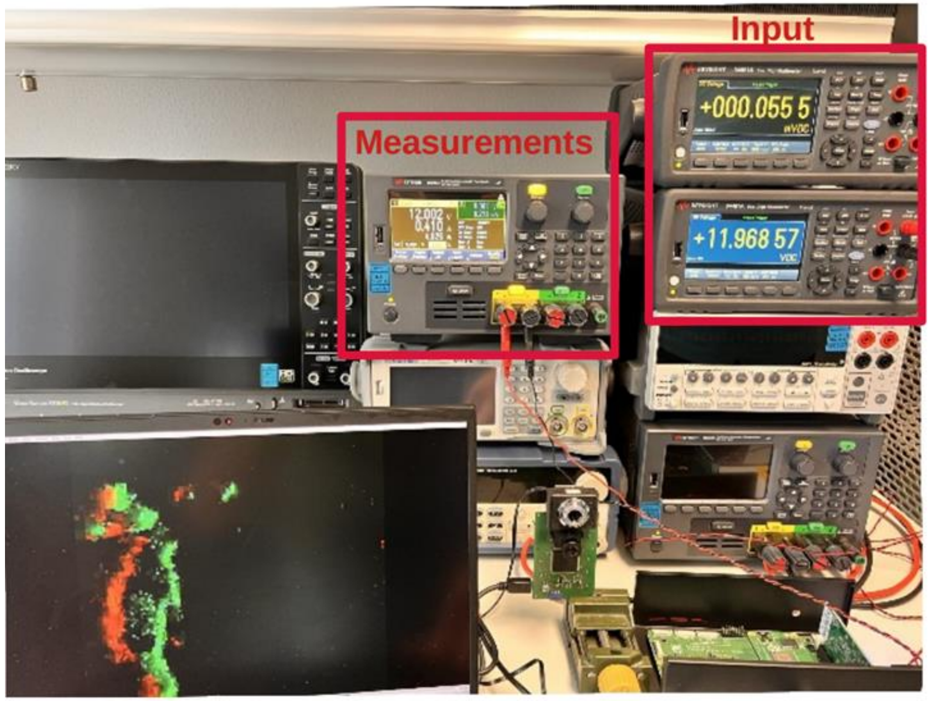
Dataset	Power(W)	Latency(ms)
Dataset 1 (daytime)	7.58	14.32
Dataset 2 (daytime)	4.92	8.21
Dataset 3 (night)	5.65	24.62

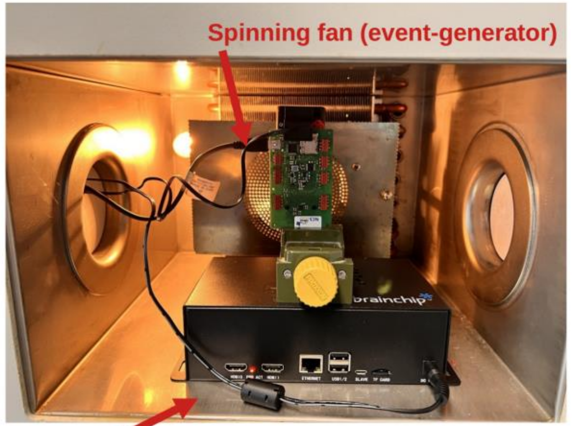
* For the accuracy evaluation of both the detection and tracking components of the system, we performed multiple experiments in order to get a statistically valid analysis. For the quantitative evaluation, we executed the following protocol: 1) Read relevant detection and tracking data from each of the experiments (i.e. N = 30 experiments); 2) Perform statistical tests (i.e., a combination of omnibus ANOVA and posthoc pairwise T-test with a significance p = 0.05) and adjust the ranking of experiments depending on significance; 3) rank subsets of relevant metrics (i.e., the metrics with a significant difference, e.g., the F1 score for detection and four others tracking specific metrics).

Deployment evaluation

- Power consumption & weatherization analysis

Measurement setup
Camera pointing to a screen with recorded traffic data





Edge device:
Brainchip Akida RaspberryPi board and event-based neuromorphic camera

Dataset	Power(W)	Latency(ms)
Dataset 1 (daytime)	7.58	14.32
Dataset 2 (daytime)	4.92	8.21
Dataset 3 (night)	5.65	24.62

* For the accuracy evaluation of both the detection and tracking components of the system, we performed multiple experiments in order to get a statistically valid analysis. For the quantitative evaluation, we executed the following protocol: 1) Read relevant detection and tracking data from each of the experiments (i.e. N = 30 experiments); 2) Perform statistical tests (i.e., a combination of omnibus ANOVA and posthoc pairwise T-test with a significance p = 0.05) and adjust the ranking of experiments depending on significance; 3) rank subsets of relevant metrics (i.e., the metrics with a significant difference, e.g., the F1 score for detection and four others tracking specific metrics).

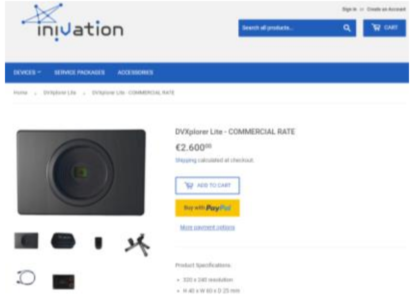
ohm

Complete BOM and costs

Bill of Materials (BOM) Single unit & suggested price for large quantities

Component	Price	Notes
Raspberry PI Compute Module 4 IO Board with RPI CM4Lite	50 $	Price pro unit sold independently from the Brainschip Akida PCIexpress board (see below). https://www.reichelt.de/de/de/raspberry-pi-compute-modul-4-io-board-rpi-cm4-io-board-p290556.html
Brainchip Akida AKD1000 PCIexpress Board	499 $	Price pro unit sold independently from the IO board/carried board. https://shop.brainchipinc.com/products/akida%E2%84%A2-development-kit-pcie-board
IniVation Dynamic Vision Sensor	2500 $	Price per unit, with up to 50 $ if large quantities purchased. https://shop.inivation.com/collections/dvxplorer-lite-1/products/dvxplorer-lite-commercial-rate
USB to miniUSB cable	1 $ (1m long USB cable) – 26 $ (5m long USB cable with signal amplifier)	Length of the cable depends on the gantry layout, we have tried with 1m long USB cable and also with signal amplification 5m long USB cable. https://www.conrad.de/de/p/delock-usb-kabel-usb-3-2-gen1-usb-3-0-usb-3-1-gen1-usb-a-stecker-usb-a-buchse-5-00-m-rot-schwarz-vergoldete-steckkontakte-ul-zertifiziert-82755-649883.html
Total	**3076 $**	Price per unit. When more units are bought a total price of approx. **226 $** for a price per unit **100 $ for Akida Chip**, **50 $ for RaspberryPI boards**, **50 $ for DVS camera**, and long **USB cable 26 $**.

Source:
https://www.proshop.nl/Smart-Home/Samsung-SmartThings-Vision/2786571



Source:
https://shop.inivation.com/collections/dvxplorer-lite-1

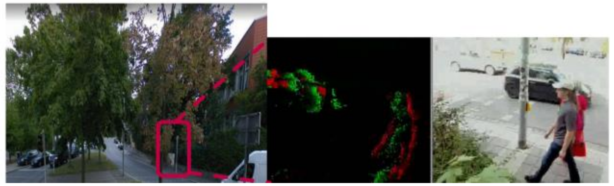
Best price ~226$

Worst price ~3000$

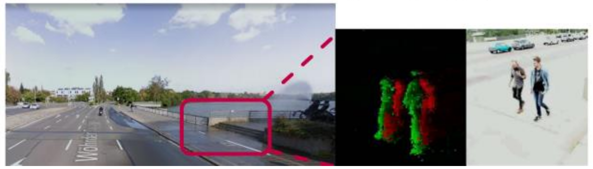
Deliverables

- **Datasets** release

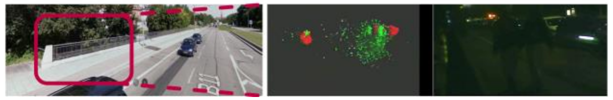
Dataset location 1
- 4 lanes (4 per direction) wide street
- Location: https://goo.gl/maps/JaYGwaTaBHj5H6SL9
- 50 kmh (urban) speed limit
- Near university campus with Pedestrians (people walking), Bicyclists (people biking, scooting, rolling, etc.)
- Ideal Operating Environment



Dataset location 2
- 8 lanes (4 per direction) wide street
- Location: https://goo.gl/maps/jar6AjysZiM2LP5S7
- 50 kmh (urban) speed limit
- Near main train stations of the city and a location with Pedestrians (people walking/running/jogging), Bicyclists (people biking, scooting, rolling, etc.)

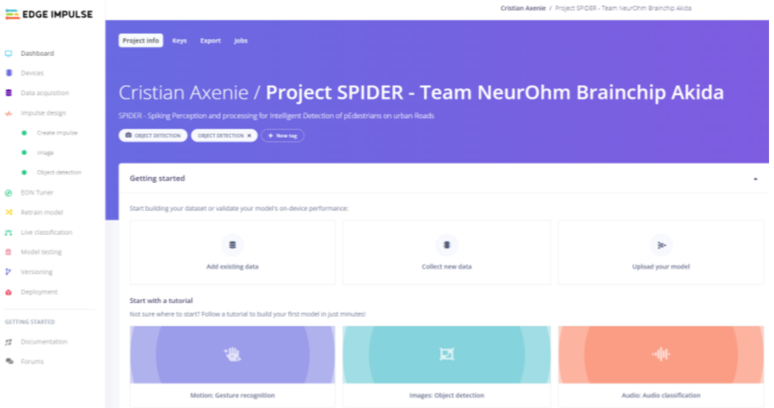


Dataset location 3
- 6 lanes (3 per direction) wide street on bridge
- Location: https://goo.gl/maps/SEEsmpgmLPcD8fG7A
- 50 kmh (urban) speed limit
- Near ring street of Munich and a location with Pedestrians (people walking/running/jogging), Bicyclists (people biking, scooting, rolling, etc.)
- Night time data acquisition

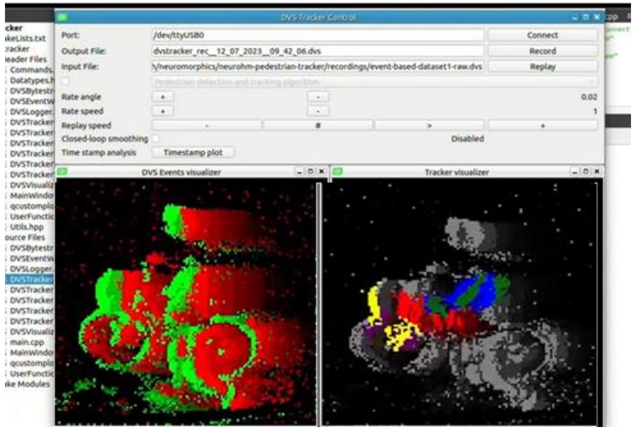


- **Code** release

Model life-cycle and data analysis



Accurate TinyML algorithms



- **Solution** release

Minimal energy footprint

Deployment
Akida Spiking Neural Networks in Event data

Deployment
RaspberryPi Event-based Expectation Maximization

Feasible deployment

ohm

Wrap-up

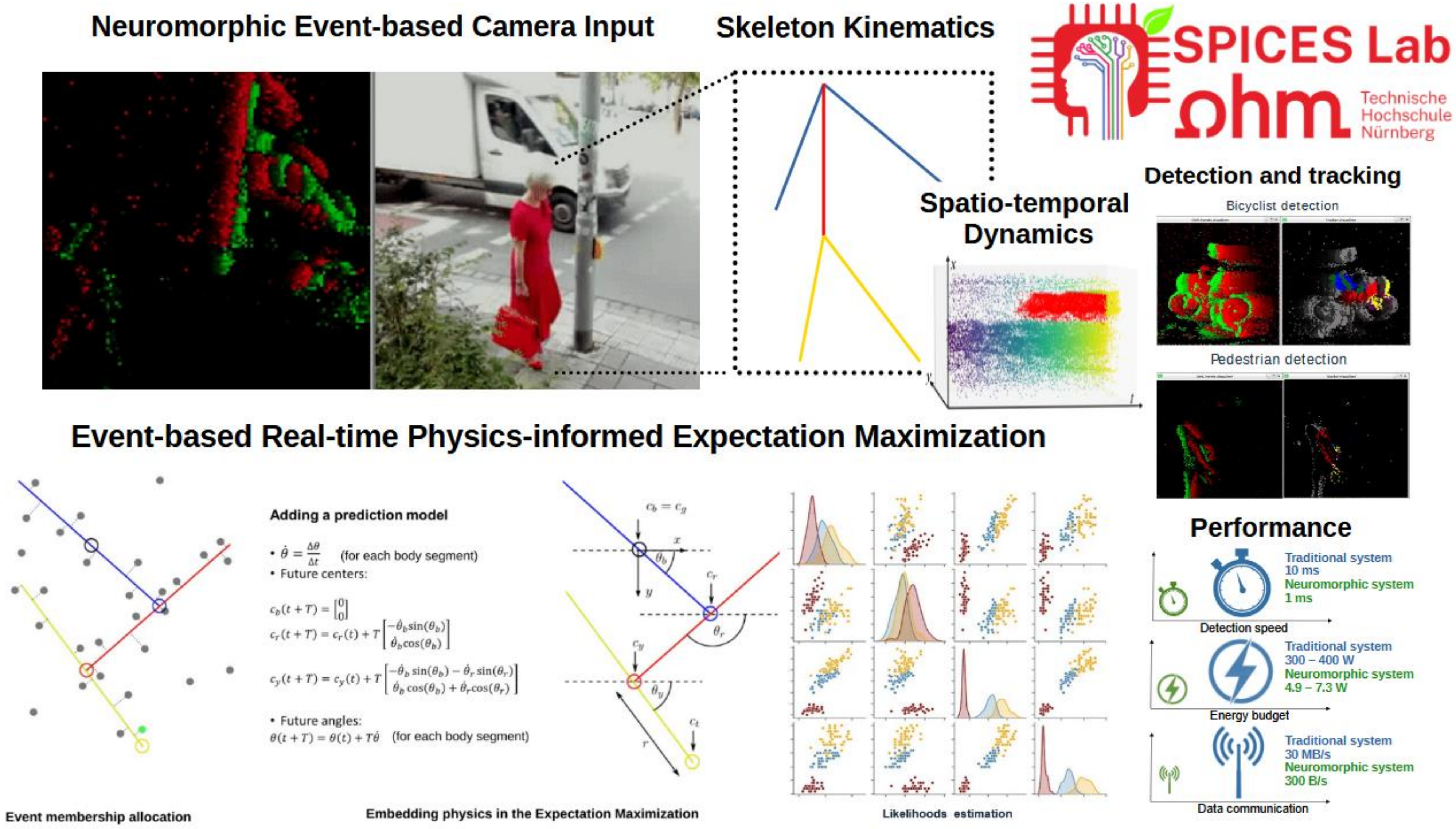
TinyML solution for supporting **VisionZero pedestrian detection**

- uses **low-power neuromorphic sensing** and **processing**
- employs **only local processing** (at the **edge**)
- provides **good accuracy** for **robust visual detection** under varying conditions

- **Current development:**
    - HD Event-based camera (Prophesee EVK3)
    - Explore new computing platforms (Synsense Speck)
    - Deployment at city-scale (Stadt Schwabach)

STADT SCHWABACH

Die Goldschlägerstadt.


Neuromorphic Event-based Camera Input
Skeleton Kinematics
SPICES Lab
ohm
Technische Hochschule Nürnberg
Spatio-temporal Dynamics
Detection and tracking
Bicyclist detection
Pedestrian detection
Event-based Real-time Physics-informed Expectation Maximization
Adding a prediction model
• $\dot{\theta} = \frac{\Delta\theta}{\Delta t}$ (for each body segment)
• Future centers:
$c_b(t+T) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$c_r(t+T) = c_r(t) + T\begin{bmatrix} -\dot{\theta}_b \sin(\theta_b) \\ \dot{\theta}_b \cos(\theta_b) \end{bmatrix}$
$c_y(t+T) = c_y(t) + T\begin{bmatrix} -\dot{\theta}_b \sin(\theta_b) - \dot{\theta}_r \sin(\theta_r) \\ \dot{\theta}_b \cos(\theta_b) + \dot{\theta}_r \cos(\theta_r) \end{bmatrix}$
• Future angles:
$\theta(t+T) = \theta(t) + T\dot{\theta}$ (for each body segment)
$c_b = c_g$
Event membership allocation
Embedding physics in the Expectation Maximization
Likelihoods estimation
Performance
Traditional system
10 ms
Neuromorphic system
1 ms
Detection speed
Traditional system
300 – 400 W
Neuromorphic system
4.9 – 7.3 W
Energy budget
Traditional system
30 MB/s
Neuromorphic system
300 B/s
Data communication