tinyML® EMEA

Enabling Ultra-low Power Machine Learning at the Edge

tinyML EMEA Technical Forum 2021 Proceedings

June 7 – 10, 2021

Virtual Event

www.tinyML.org

Bottom-Up and Top-Down Neural Processing Systems Design:

Unveiling the Road toward Neuromorphic Intelligence

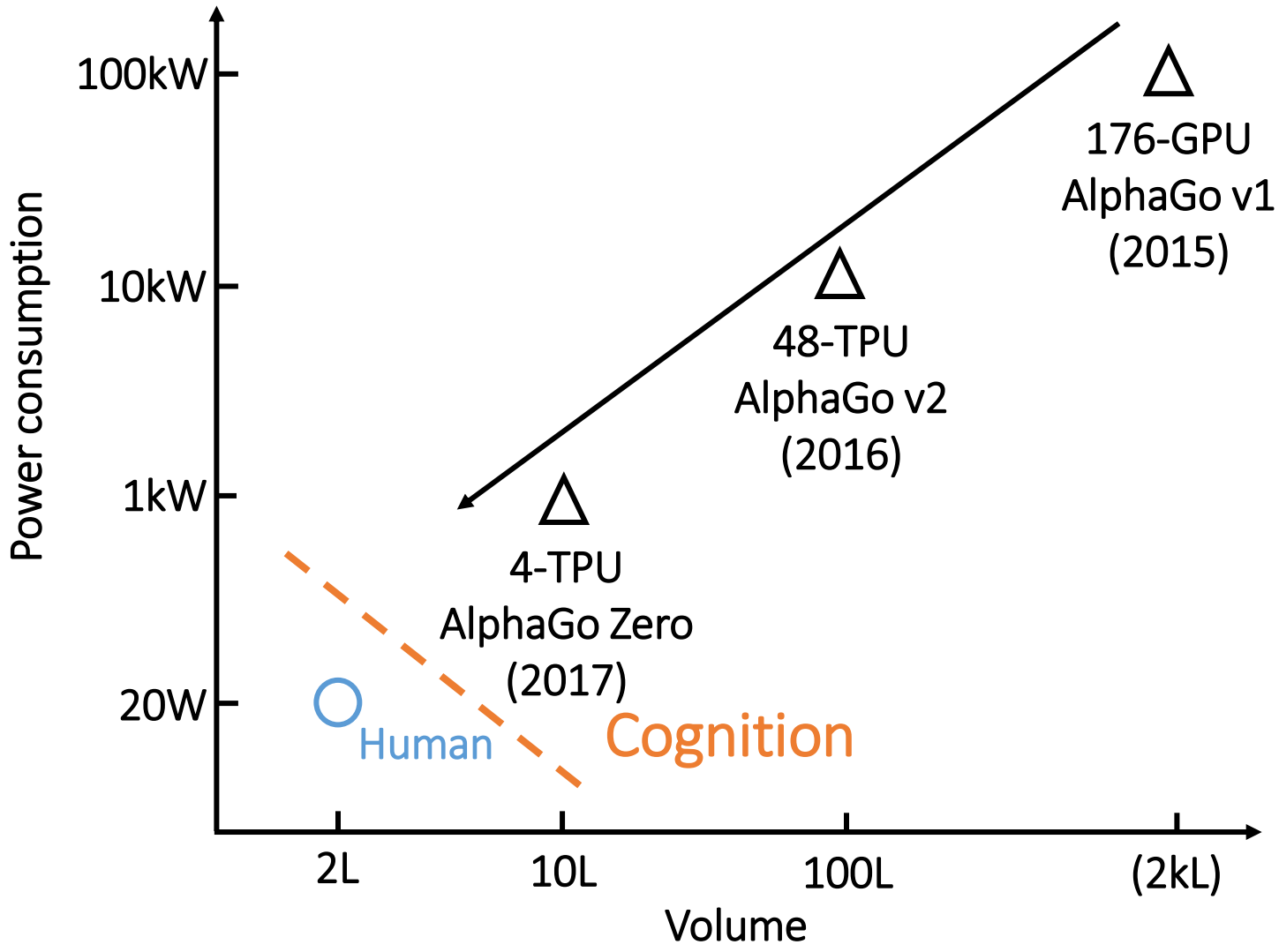
Charlotte Frenkel

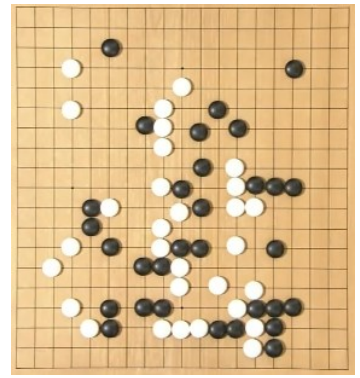
Institute of Neuroinformatics, UZH and ETH Zürich, Switzerland
charlotte@ini.uzh.ch

tinyML EMEA Technical Forum
Online, June 7-10, 2021

Neuromorphic Engineering – Why?

Efficiency of artificial intelligence vs. natural intelligence?

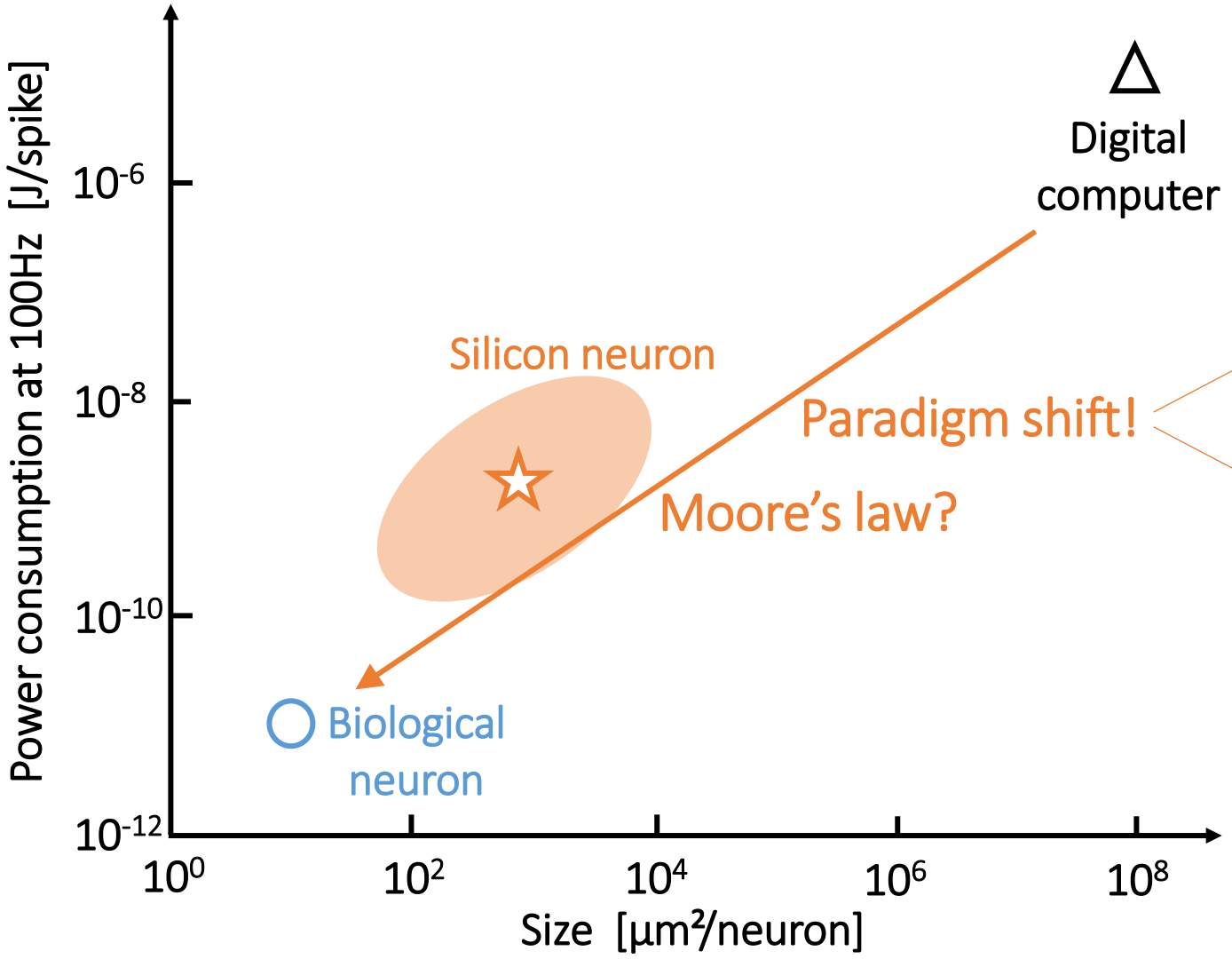




[Silver & Hassabis, *https://deepmind.com/blog/article/alphago-zero-starting-scratch*, 2017]

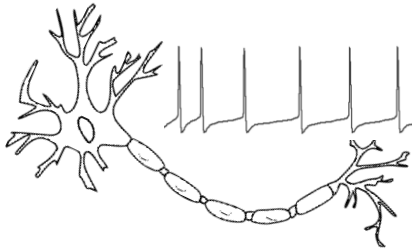
Neuromorphic Engineering – Why?

Efficiency of bio-inspired neuromorphic computing?



Data representation:
sparse, event-driven spike trains

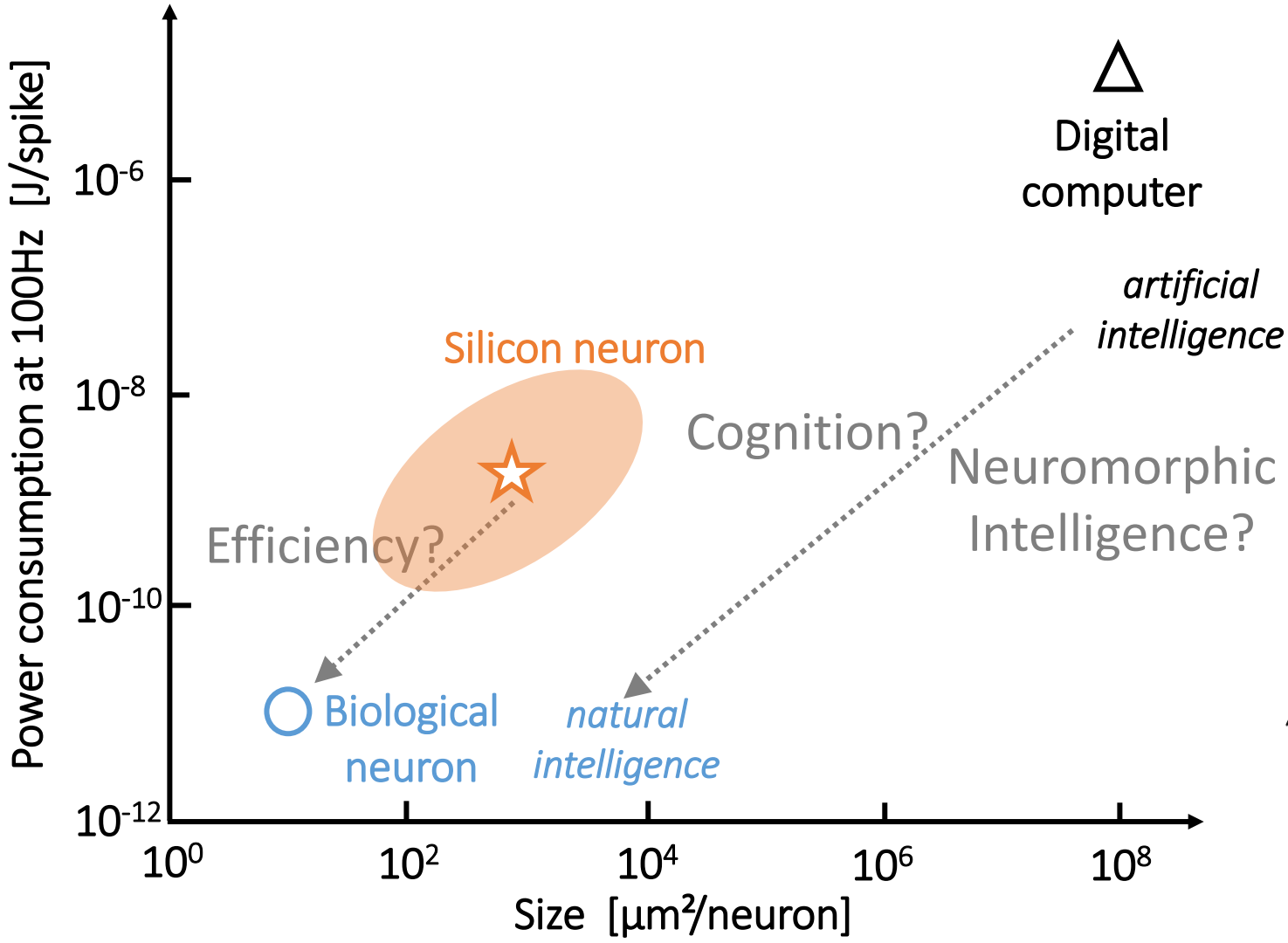
Architecture: distributed processing with co-located neurons and synapses

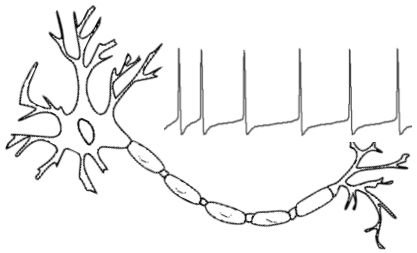


[Poon & Zhou, *Front. Neurosci.*, 2011]

Neuromorphic Engineering – How?

A design strategy toward efficiency and cognition?



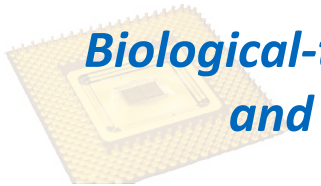


[Poon & Zhou, *Front. Neurosci.*, 2011]

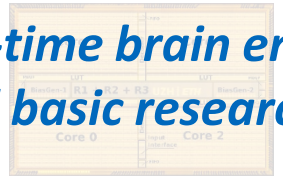
Neuromorphic Engineering – How?

A design strategy toward efficiency and cognition?

Subthreshold analog (mixed-signal)

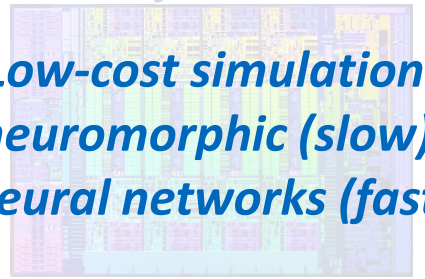


Biological-time brain emulation and basic research



ROLLS (UZH/ETHZ) | DYNAPs (UZH/ETHZ) | NeuroGrid (Stanford)

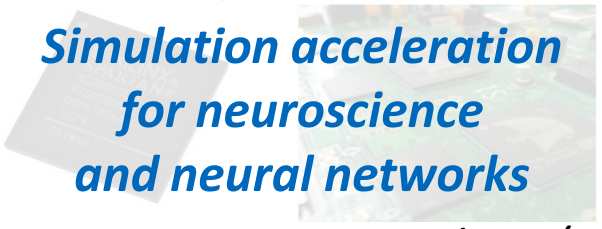
Software



Low-cost simulation: neuromorphic (slow), neural networks (fast)

CPU / GPU

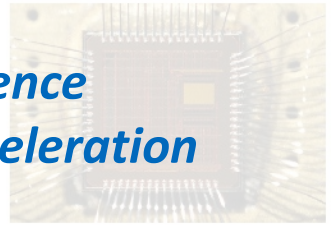
Dedicated/distributed sim.



Simulation acceleration for neuroscience and neural networks

FPGA | SpiNNaker 1/2 (Manchester, TUD)

Above-threshold analog (mixed-signal)



Neuroscience simulation acceleration

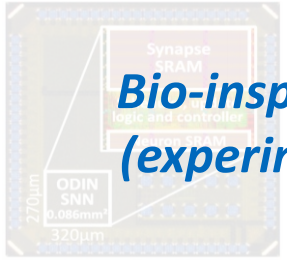
BrainScaleS 1/2 (Heidelberg)

Large-scale full-custom digital designs

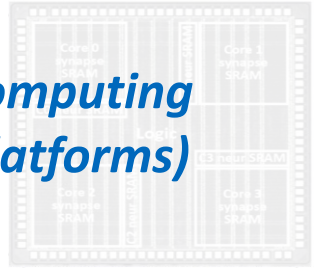
Cognitive computing

TrueNorth (IBM) | Loihi (Intel)

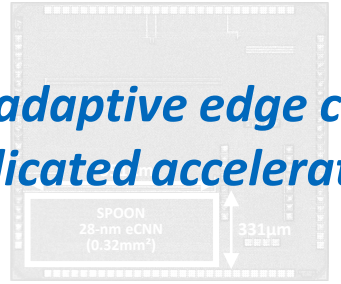
Small-scale full-custom digital designs



Bio-inspired edge computing (experimentation platforms)



Low-cost adaptive edge computing (dedicated accelerators)

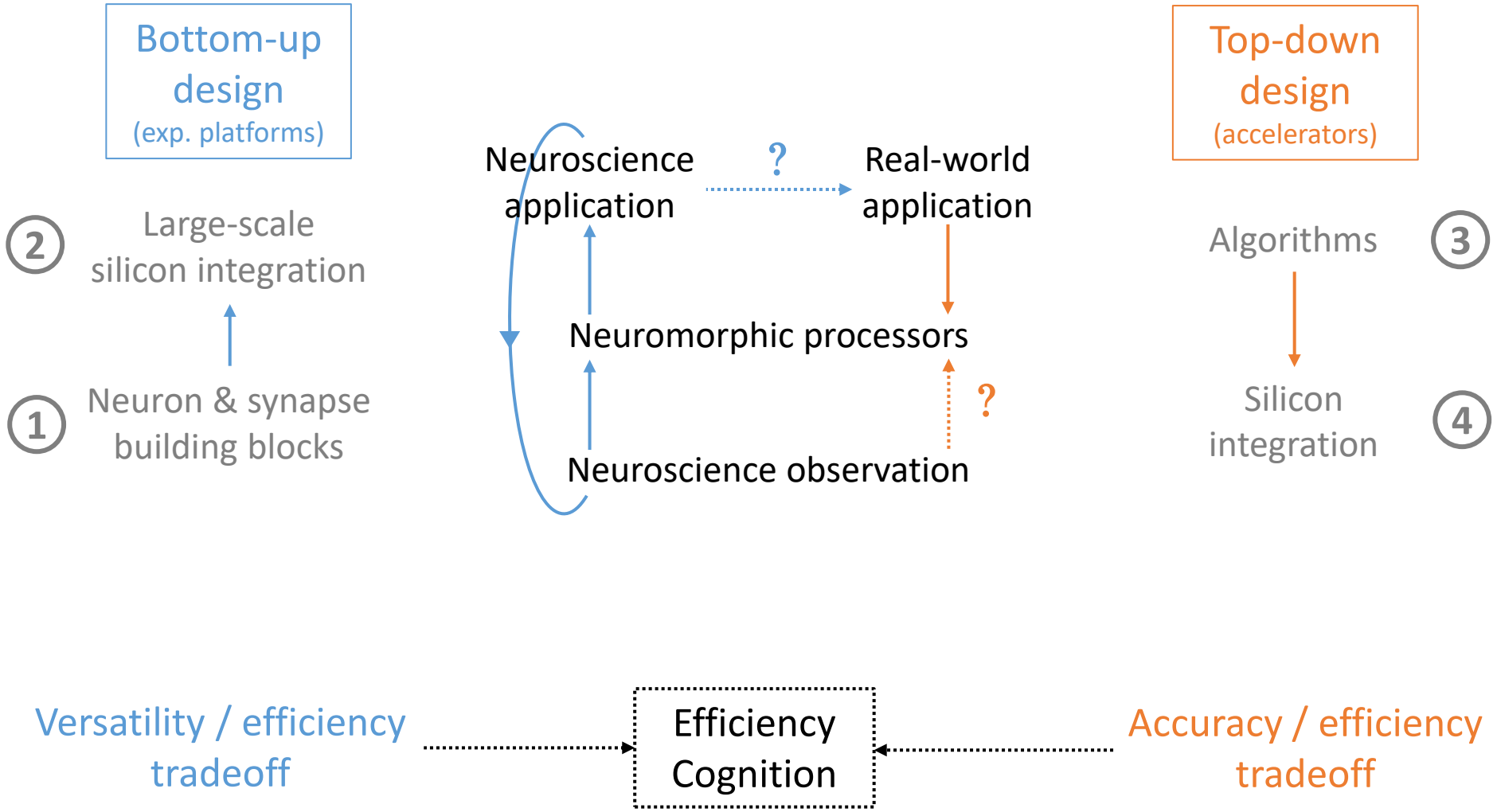


ODIN (UCLouvain) | MorphIC (UCLouvain) | SPOON (UCLouvain)

See also:
[Seo, CICC'11]
[Knag, JSSC'15]
[Park, ISSCC'19]

Neuromorphic Engineering – How?

Unveiling roads to embedded cognition



Outline

Part I – Bottom-up neuromorphic design

- Building blocks
- Integration

Part II – Top-down neuromorphic design

- Algorithms
- Integration

Conclusion and perspectives

Outline

Part I – Bottom-up neuromorphic design

- Building blocks

  Neurons and synapses as adaptive processing and memory elements

  [Frenkel, *ISCAS*, 2017]
  [Frenkel, *BioCAS*, 2017]

- Integration

Part II – Top-down neuromorphic design

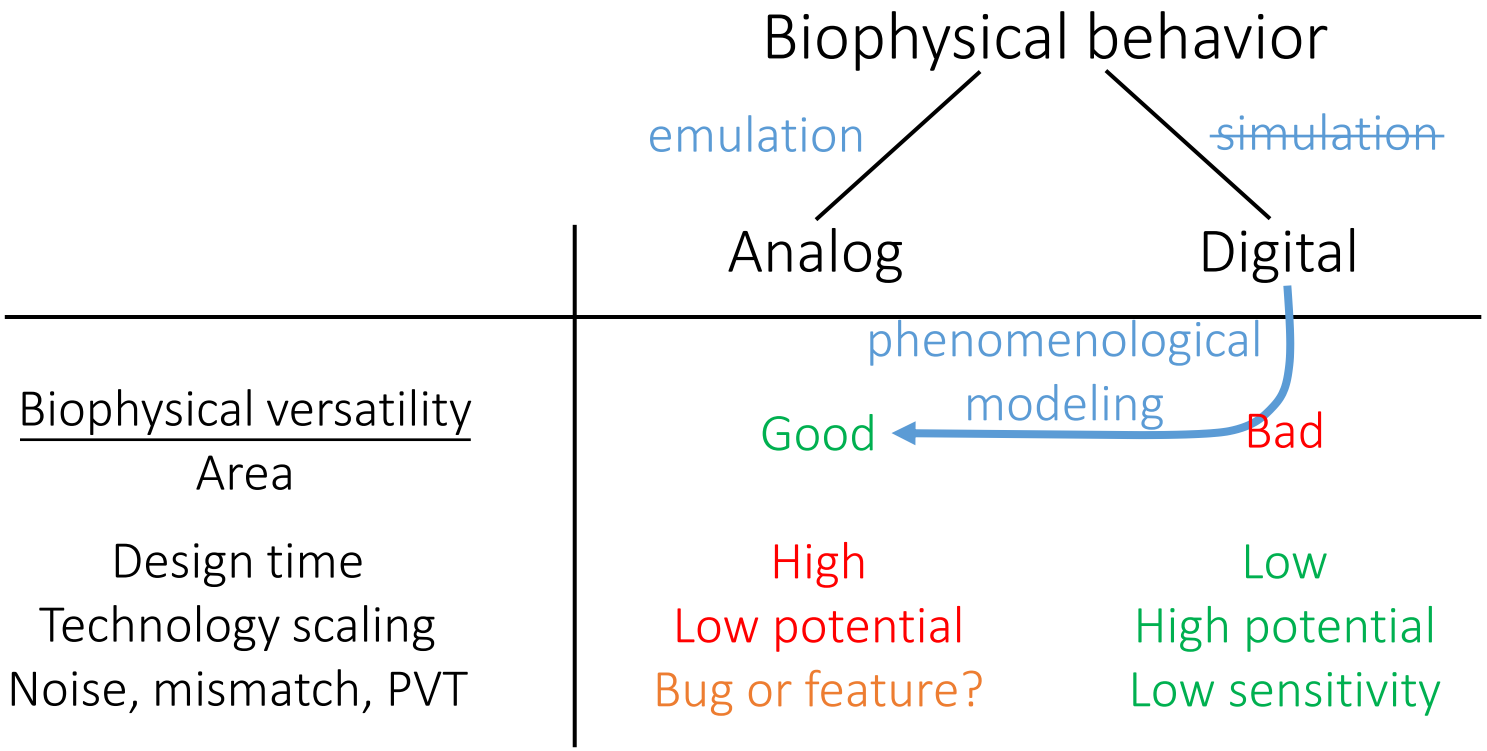
- Algorithms
- Integration

Conclusion and perspectives

Design strategy

Analog or digital?

We'll come back to this. ☺



How can we make the best of both worlds?

Design strategy

What should we aim for and phenomenologically implement?

Neurons

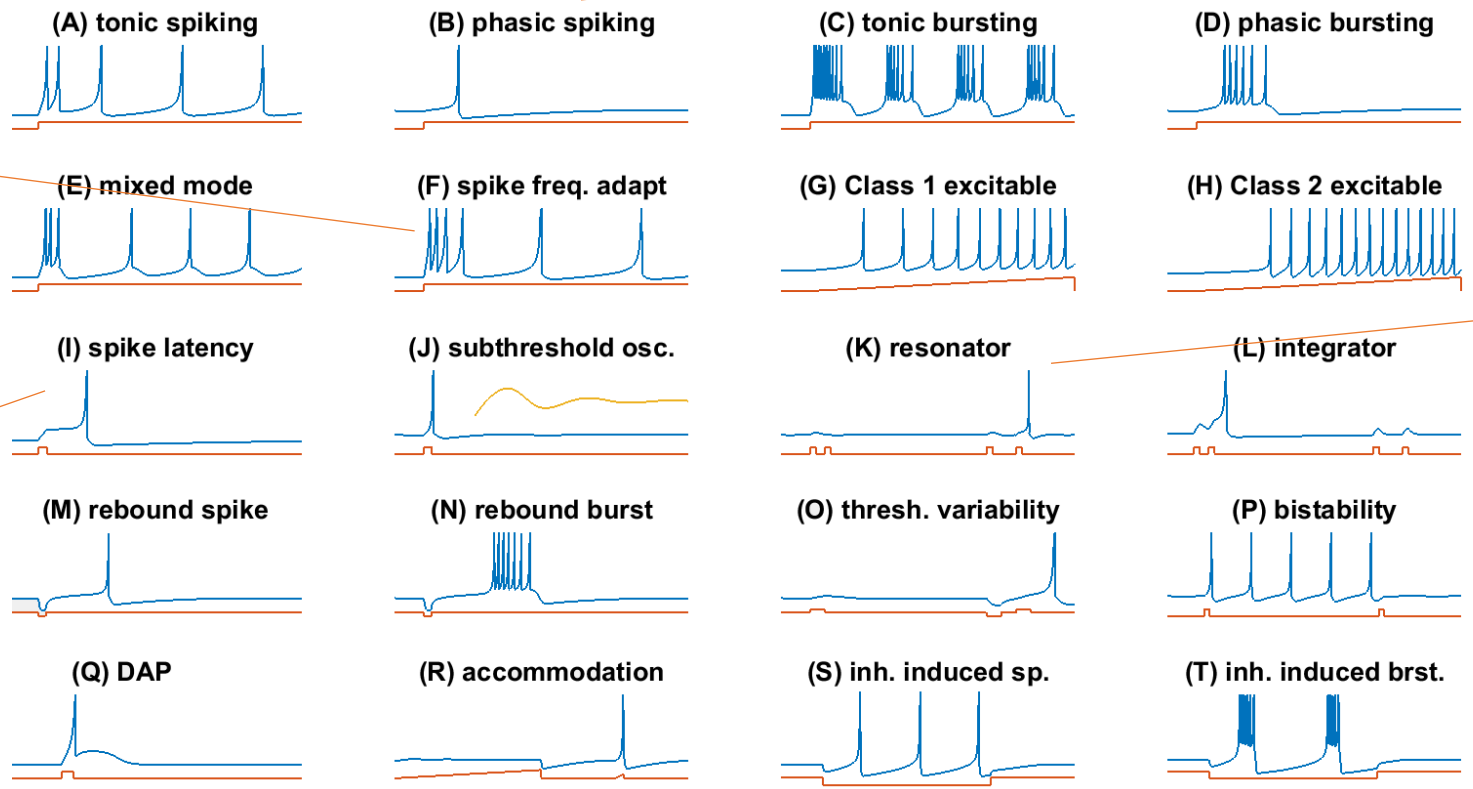
- 20 Izhikevich behaviors of cortical spiking neurons

Useful for time-to-first-spike encodings

Introduce competition for unsupervised learning in winner-take-all networks [Kreiser, BioCAS'17]

Discrimination of specific frequencies

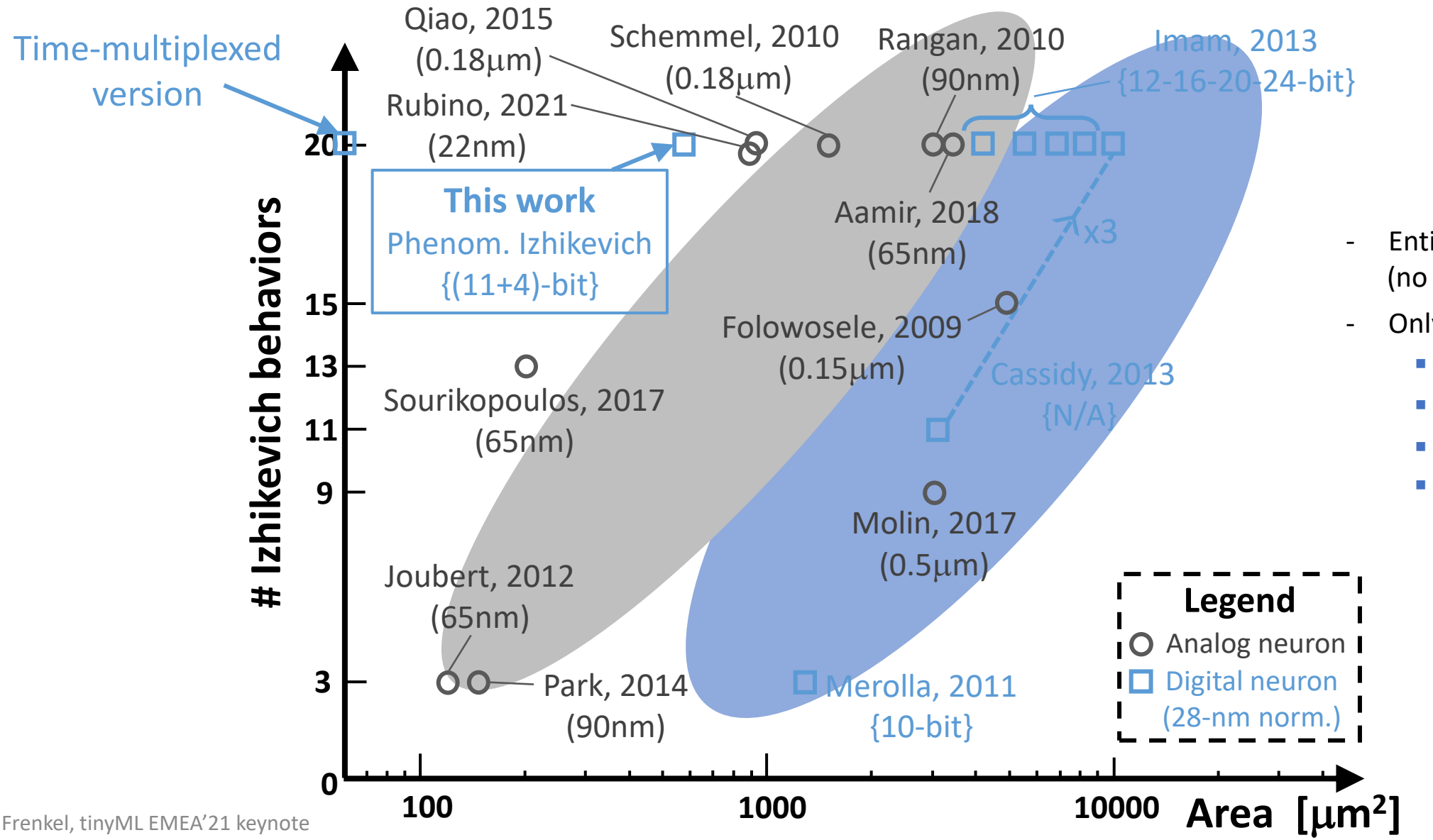
Stereo sound source localization [Schoepe, BioCAS'19]



[Izhikevich, IEEE Trans. NN, 2004]

Proposed phenomenological digital neuron

Tackling the versatility/efficiency tradeoff



Key features:

- Entirely event-driven (no time-stepped integration)
- Only 4 functions necessary:
  - Threshold adaptation
  - Time window generation
  - Simple template matching
  - Membrane potential sign rotation

Design strategy

What should we aim for and phenomenologically implement?

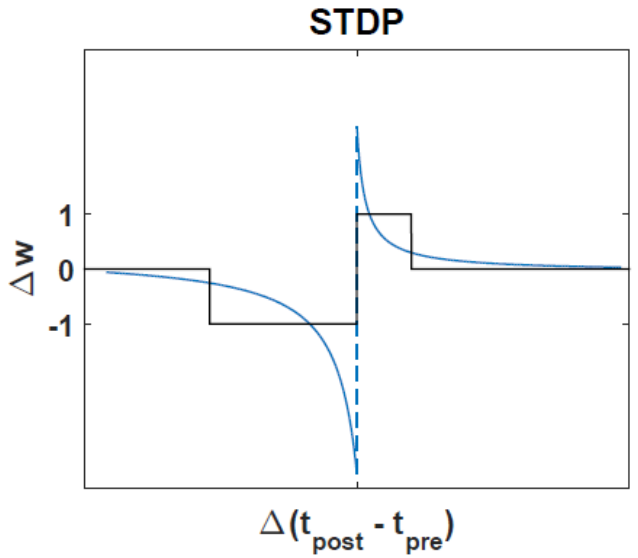
→ *perspectives*

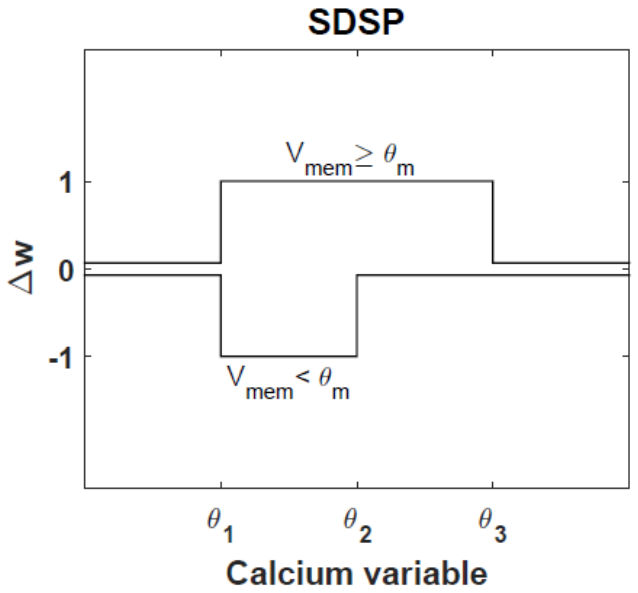
Neurons

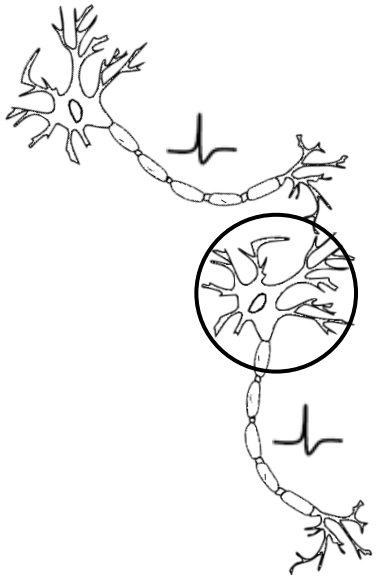
- 20 Izhikevich behaviors of cortical spiking neurons

Synapses

- Spike-based online learning







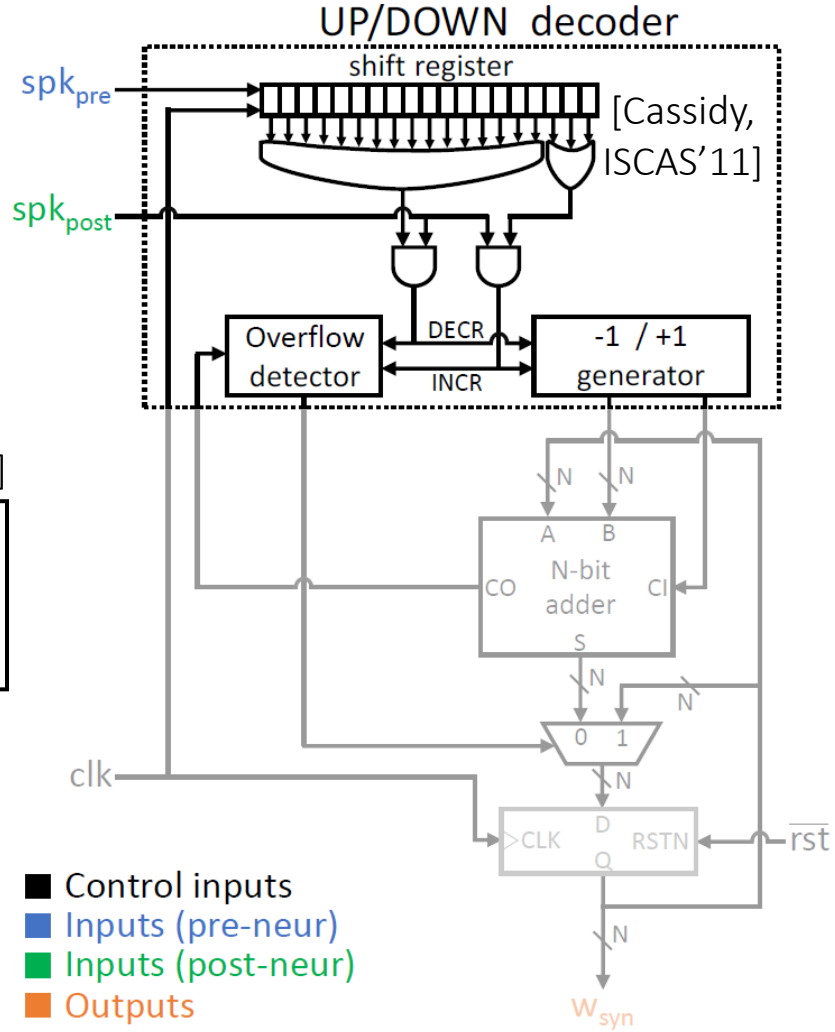
Proposed digital synapse

Tackling the versatility/efficiency tradeoff

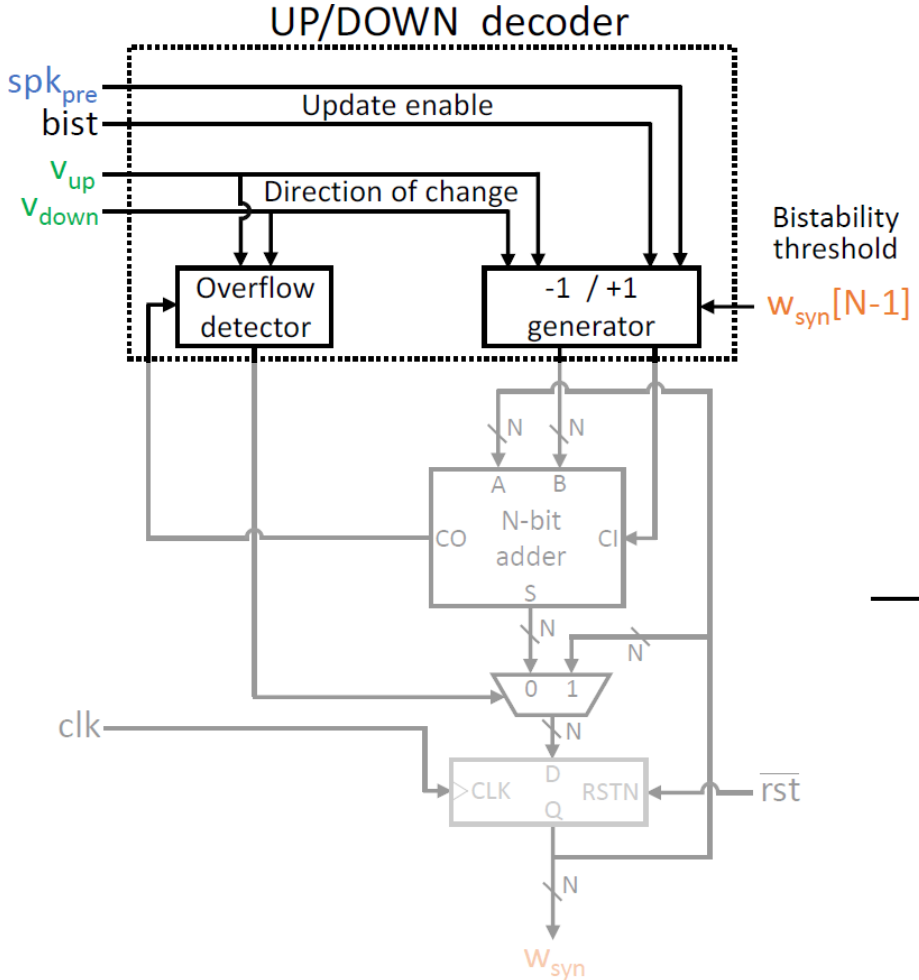
Key challenge – Fan-in = 100-10000 synapses/neuron

[Seo, CICC'11]

Custom dual-port SRAM



STDP



SDSP

Foundry single-port SRAM

Record density

Outline

Part I – Bottom-up neuromorphic design

- Building blocks
- Integration

Proposed neuromorphic experimentation platforms

[Frenkel, *Trans. BioCAS*, 2019a]
[Frenkel, *Trans. BioCAS*, 2019b]

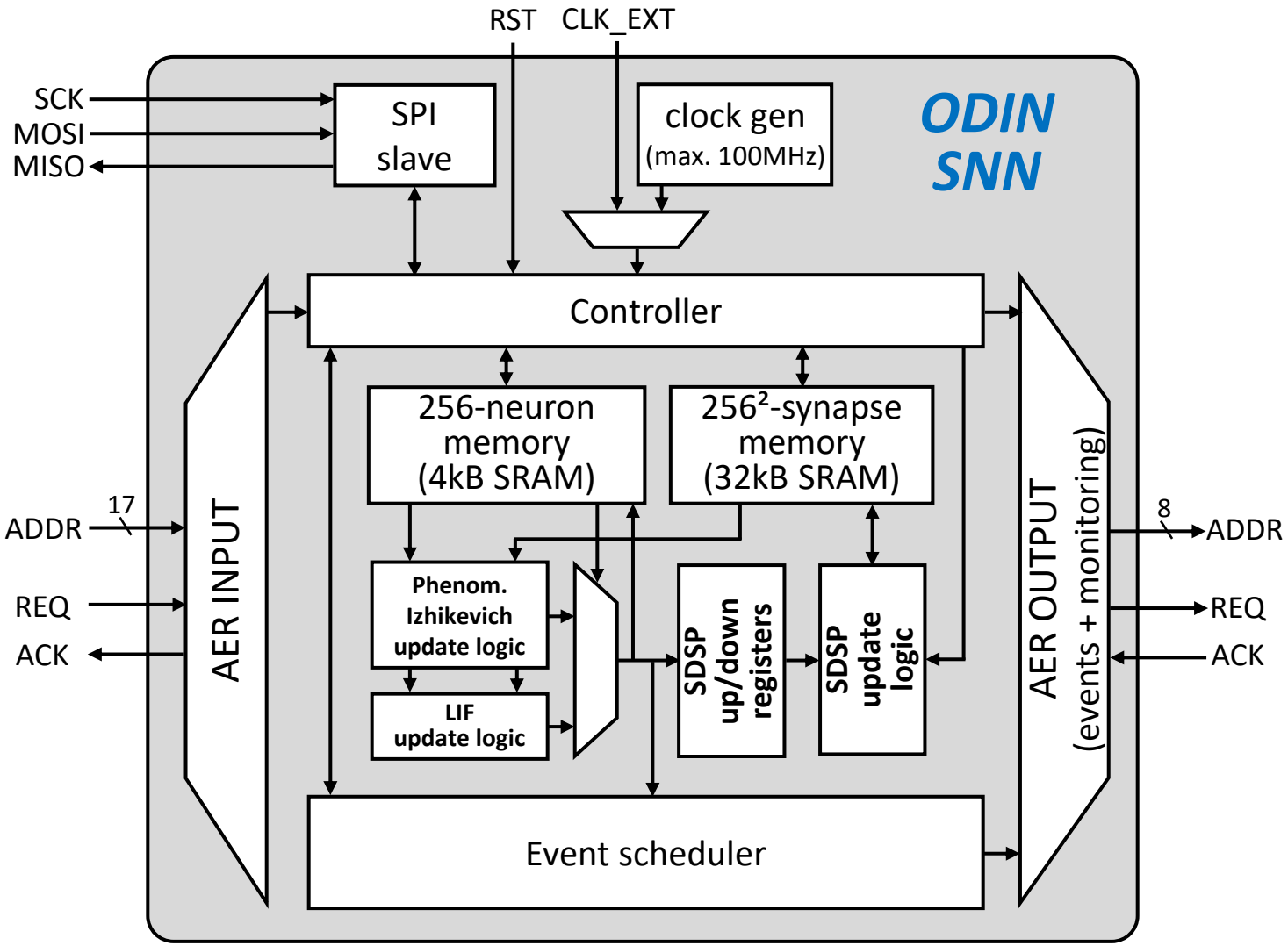
Part II – Top-down neuromorphic design

- Algorithms
- Integration

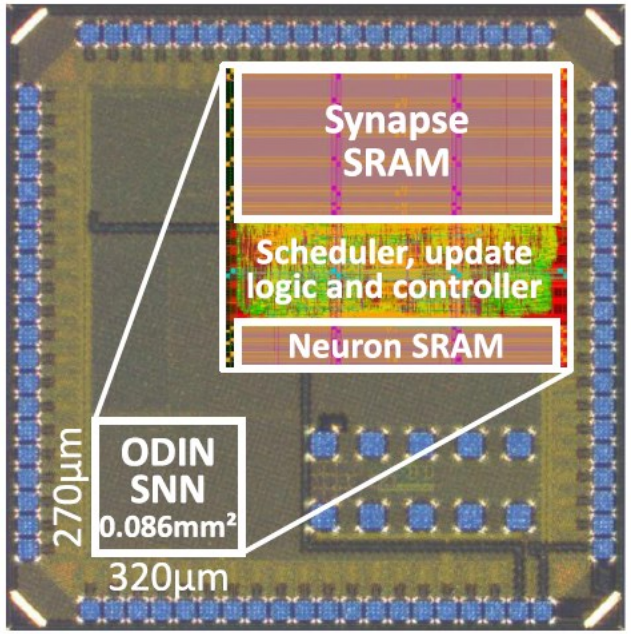
Conclusion and perspectives

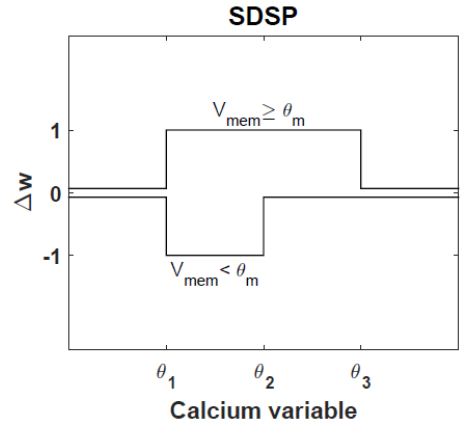
Architecture of ODIN

ODIN – A 256-neuron 64k-synapse Online-learning Digital Neurosynaptic core



ODIN – Chip microphotograph and specifications

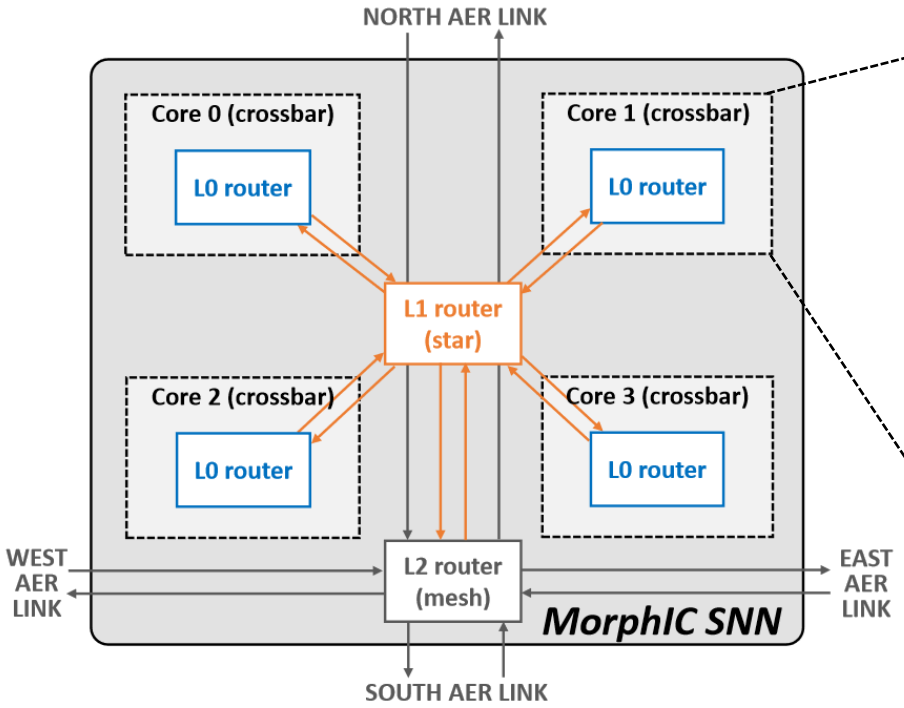




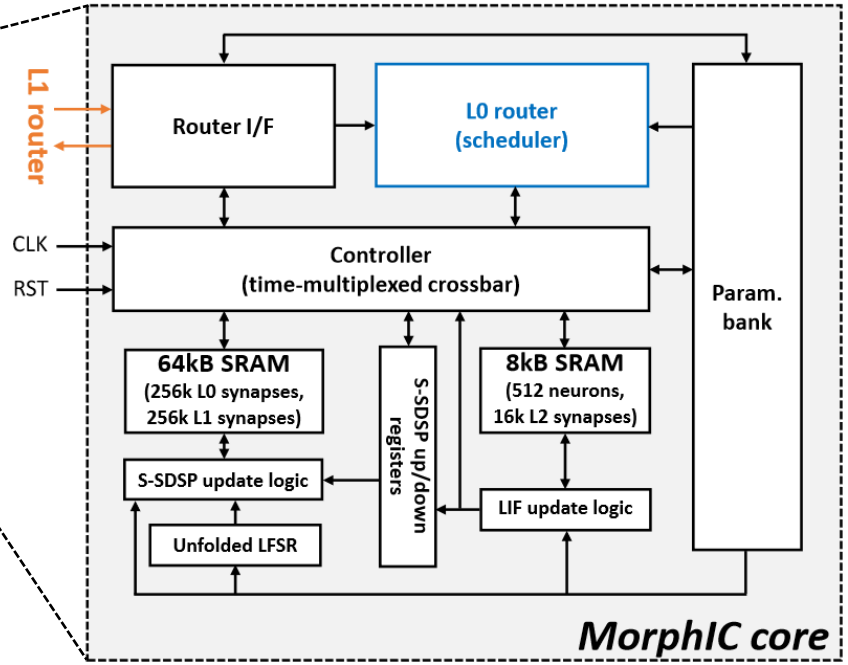
Technology	28nm FDSOI
Implementation	Digital
Area	0.086mm²
# neurons	256
# synapses	64k
# Izhikevich behav.	20
Online learning	SDSP, (3+1)-bit weight
Time constant	Biological to accelerated
Supply voltage	0.55V – 1.0V
Leakage power (P_{leak})	27.3µW @0.55V
Idle power (P_{idle})	1.78µW/MHz @0.55V
Incr. energy/SOP (E_{SOP})	8.43pJ @0.55V
Global energy/SOP ($E_{tot,SOP}$)	>12.7pJ @0.55V
Routing flexibility/efficiency	☹ (AER)
Fan-in	256
Fan-out	256

Architecture of MorphIC

Chip-level architecture



Core architecture

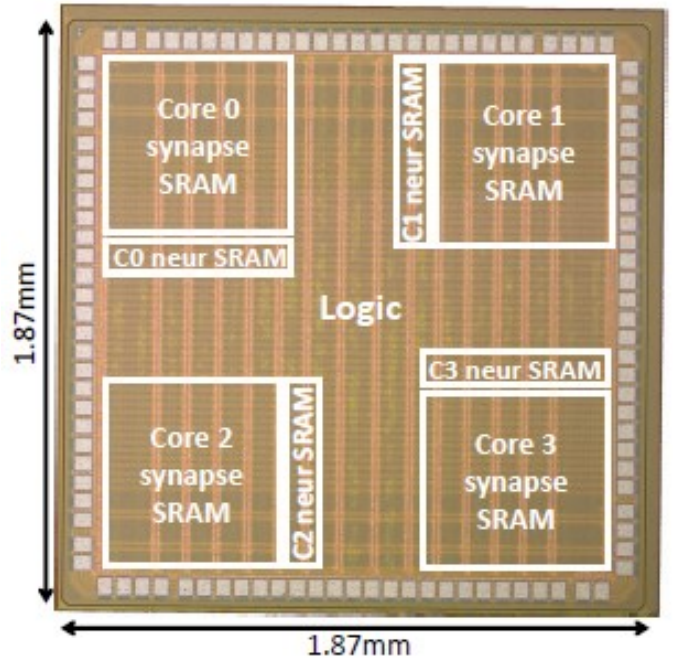


Neurons/core	512
Synapses/core	528k

Fan-in	1k
Fan-out	2k

Stochastic SDSP (S-SDSP) on binary synapses

MorphIC – Chip microphotograph and specifications



Technology	65nm LP CMOS
Implementation	Digital
Area	3.5mm² (incl. pads) 2.86mm² (excl. pads)
Number of cores	4
Total # neurons (type)	2048 (LIF)
Total # synapses (hier.)	1M (L0), 1M (L1), 64k (L2)
Fan-in (hier.)	512 (L0), 512 (L1), 32 (L2)
Fan-out (hier.)	512 (L0), 3x512 (L1), 4 (L2)
Online learning	Stochastic SDSP, 1-bit weight
Time constant	Biological to accelerated
Supply voltage	0.8V – 1.2V
Max. clock frequency	55MHz (0.8V) – 210MHz (1.2V)
Leakage power (P_{leak})	45µW @0.8V
Idle power (P_{idle})	41.3µW/MHz @0.8V
Energy/SOP (E_{SOP})	30pJ @0.8V

Comparison with SoA experimentation platforms

		Mixed-signal				Digital				
Author		Schemmel	Benjamin	Qiao	Moradi	Painkras	Akopyan	Davies	**Frenkel**	**Frenkel**
Publication		ISCAS, 2010	PIEEE, 2014	Front. NS, 2015	TBioCAS, 2017	JSSC, 2013	TCAD, 2015	IEEE Micro, 2018	TBCAS, 2019a	TBCAS, 2019b
Chip name		HICANN	Neurogrid	ROLLS	DYNAPs	SpiNNaker	TrueNorth	Loihi	**ODIN**	**MorphIC**
Implementation		Mixed-signal	Mixed-signal	Mixed-signal	Mixed-signal	Digital	Digital	Digital	Digital	Digital
Technology		0.18μm	0.18μm	0.18μm	0.18μm	0.13μm	28nm	14nm FinFET	28nm FDSOI	65nm LP
# cores		1	16	1	4	18	4096	128	1	4
Neurosynaptic core area [mm^2]		49	168	51.4	7.5	3.75	0.095	0.4	0.086	0.715
# Izhikevich behaviors		(20)	N/A	(20)	(20)	Programmable	11 (3 neur: 20)	(6)	20	3
# neurons per core		512	64k	256	256	max. 1000	256	max. 1024	256	512
Synaptic weight storage		4-bit (SRAM)	Off-chip	Capacitor	12-bit (CAM)	Off-chip	1-bit (SRAM)	1- to 9-bit (SRAM)	(3+1)-bit (SRAM)	1-bit (SRAM)
Embedded online learning		STDP	No	SDSP	No	Programmable	No	Programmable	SDSP	S-SDSP
# synapses per core		112k	–	128k	16k	–	64k	1M to 114k (1-9 bits)	64k	528k
Time constant		Accelerated	Biological	Biological	Biological	Bio. to accel.	Biological	N/A	Bio. to accel.	Bio. to accel.
Flexibility	routing	Medium	Medium	Low	Medium	High	Medium	High	Low	Medium
	learning	Low	–	Low	Low	–	–	High	Low	Low
Neuron core density [neur/mm^2]	raw	10.5	390	5	34	max. 267	2.6k	max. 2.5k	3.0k	716
	norm.	–	–	–	–	max. 5.8k	2.6k	max. 1k	3.0k	3.9k
Synapse core density [syn/mm^2]	raw	2.3k	–	2.5k	2.1k	–	674k	2.5M to 282k	741k	738k
	norm.	–	–	–	–		674k	1M to 113k	741k	4M
Supply voltage		1.8V	3.0V	1.8V	1.3V-1.8V	1.2V	0.7V-1.05V	0.5V-1.25V	0.55V-1.0V	0.8V-1.2V
Energy per SOP	raw	N/A	(941pJ)$^▲$	>77fJ^Δ	134fJ^Δ/30pJ$^▲$ (1.3V)	>11.3nJ^Δ/26.6nJ$^▲$	26pJ$^▲$ (0.775V)	>23.6pJ^Δ (0.75V)	8.4pJ^Δ/12.7pJ$^▲$ (0.55V)	30pJ^Δ/51pJ$^▲$ (0.8V)
	norm.		–	–	–	>2.4nJ^Δ/5.7nJ$^▲$	26pJ$^▲$	(66.1pJ^Δ)	8.4pJ^Δ/12.7pJ$^▲$	12.9pJ^Δ/22pJ$^▲$

Most direct comparison: IBM TrueNorth core vs. ODIN (same technology node, same number of neurons and synapses per neurosynaptic core, same area).

	ODIN	TrueNorth	
Synapses	✔ 4-bit **with** learning	1-bit **without** learning	✘
Neurons	✔ 20 Izh. beh.	11 Izh. beh.	✘
Energy/SOP	✔ 12.7pJ @0.55V	26pJ @0.775V	✘
Connectivity	✘ AER	large-scale mesh	✔ → MorphIC

Comparison with SoA experimentation platforms

		Mixed-signal				Digital				
Author		Schemmel	Benjamin	Qiao	Moradi	Painkras	Akopyan	Davies	**Frenkel**	**Frenkel**
Publication		ISCAS, 2010	PIEEE, 2014	Front. NS, 2015	TBioCAS, 2017	JSSC, 2013	TCAD, 2015	IEEE Micro, 2018	TBCAS, 2019a	TBCAS, 2019b
Chip name		HICANN	Neurogrid	ROLLS	DYNAPs	SpiNNaker	TrueNorth	Loihi	**ODIN**	**MorphIC**
Implementation		Mixed-signal	Mixed-signal	Mixed-signal	Mixed-signal	Digital	Digital	Digital	Digital	Digital
Technology		0.18μm	0.18μm	0.18μm	0.18μm	0.13μm	28nm	14nm FinFET	28nm FDSOI	65nm LP
# cores		1	16	1	4	18	4096	128	1	4
Neurosynaptic core area [mm^2]		49	168	51.4	7.5	3.75	0.095	0.4	0.086	0.715
# Izhikevich behaviors		(20)	N/A	(20)	(20)	Programmable	11 (3 neur: 20)	(6)	20	3
# neurons per core		512	64k	256	256	max. 1000	256	max. 1024	256	512
Synaptic weight storage		4-bit (SRAM)	Off-chip	Capacitor	12-bit (CAM)	Off-chip	1-bit (SRAM)	1- to 9-bit (SRAM)	(3+1)-bit (SRAM)	1-bit (SRAM)
Embedded online learning		STDP	No	SDSP	No	Programmable	No	Programmable	SDSP	S-SDSP
# synapses per core		112k	–	128k	16k	–	64k	1M to 114k (1-9 bits)	64k	528k
Time constant		Accelerated	Biological	Biological	Biological	Bio. to accel.	Biological	N/A	Bio. to accel.	Bio. to accel.
Flexibility	routing	Medium	Medium	Low	Medium	High	Medium	High	Low	Medium
	learning	Low	–	Low	Low	–	–	High	Low	Low
Neuron core density [neur/mm^2]	raw	10.5	390	5	34	max. 267	2.6k	max. 2.5k	3.0k	716
	norm.	–	–	–	–	max. 5.8k	2.6k	max. 1k	3.0k	3.9k
Synapse core density [syn/mm^2]	raw	2.3k	–	2.5k	2.1k	–	674k	2.5M to 282k	741k	738k
	norm.	–	–	–	–		674k	1M to 113k	741k	4M
Supply voltage		1.8V	3.0V	1.8V	1.3V-1.8V	1.2V	0.7V-1.05V	0.5V-1.25V	0.55V-1.0V	0.8V-1.2V
Energy per SOP	raw	N/A	(941pJ)$^▲$	>77fJ^Δ	134fJ^Δ/30pJ$^▲$ (1.3V)	>11.3nJ^Δ/26.6nJ$^▲$	26pJ$^▲$ (0.775V)	>23.6pJ^Δ (0.75V)	8.4pJ^Δ/12.7pJ$^▲$ (0.55V)	30pJ^Δ/51pJ$^▲$ (0.8V)
	norm.		–	–	–	>2.4nJ^Δ/5.7nJ$^▲$	26pJ$^▲$	(66.1pJ^Δ)	8.4pJ^Δ/12.7pJ$^▲$	12.9pJ^Δ/22pJ$^▲$

Area

ODIN and MorphIC have the highest neuron and synapse densities among all SNNs with embedded synaptic weight storage

Comparison with SoA experimentation platforms

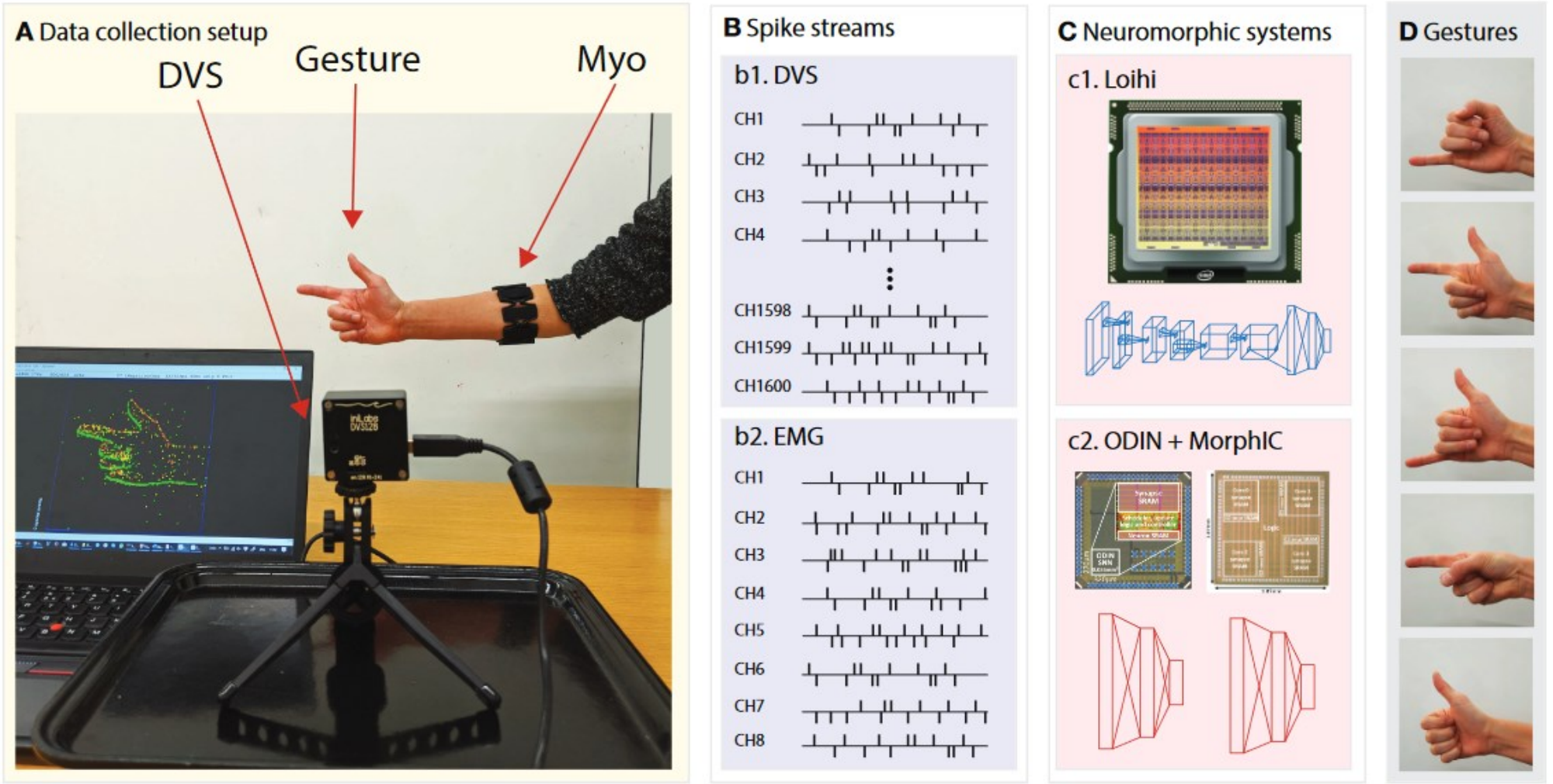
		Mixed-signal				Digital				
Author		Schemmel	Benjamin	Qiao	Moradi	Painkras	Akopyan	Davies	**Frenkel**	**Frenkel**
Publication		ISCAS, 2010	PIEEE, 2014	Front. NS, 2015	TBioCAS, 2017	JSSC, 2013	TCAD, 2015	IEEE Micro, 2018	TBCAS, 2019a	TBCAS, 2019b
Chip name		HICANN	Neurogrid	ROLLS	DYNAPs	SpiNNaker	TrueNorth	Loihi	**ODIN**	**MorphIC**
Implementation		Mixed-signal	Mixed-signal	Mixed-signal	Mixed-signal	Digital	Digital	Digital	Digital	Digital
Technology		0.18μm	0.18μm	0.18μm	0.18μm	0.13μm	28nm	14nm FinFET	28nm FDSOI	65nm LP
# cores		1	16	1	4	18	4096	128	1	4
Neurosynaptic core area [mm^2]		49	168	51.4	7.5	3.75	0.095	0.4	0.086	0.715
# Izhikevich behaviors		(20)	N/A	(20)	(20)	Programmable	11 (3 neur: 20)	(6)	20	3
# neurons per core		512	64k	256	256	max. 1000	256	max. 1024	256	512
Synaptic weight storage		4-bit (SRAM)	Off-chip	Capacitor	12-bit (CAM)	Off-chip	1-bit (SRAM)	1- to 9-bit (SRAM)	(3+1)-bit (SRAM)	1-bit (SRAM)
Embedded online learning		STDP	No	SDSP	No	Programmable	No	Programmable	SDSP	S-SDSP
# synapses per core		112k	–	128k	16k	–	64k	1M to 114k (1-9 bits)	64k	528k
Time constant		Accelerated	Biological	Biological	Biological	Bio. to accel.	Biological	N/A	Bio. to accel.	Bio. to accel.
Flexibility	routing	Medium	Medium	Low	Medium	High	Medium	High	Low	Medium
	learning	Low	–	Low	Low	–	–	High	Low	Low
Neuron core density [neur/mm^2]	raw	10.5	390	5	34	max. 267	2.6k	max. 2.5k	3.0k	716
	norm.	–	–	–	–	max. 5.8k	2.6k	max. 1k	3.0k	3.9k
Synapse core density [syn/mm^2]	raw	2.3k	–	2.5k	2.1k	–	674k	2.5M to 282k	741k	738k
	norm.	–	–	–	–		674k	1M to 113k	741k	4M
Supply voltage		1.8V	3.0V	1.8V	1.3V-1.8V	1.2V	0.7V-1.05V	0.5V-1.25V	0.55V-1.0V	0.8V-1.2V
Energy per SOP	raw	N/A	(941pJ)$^{\blacktriangle}$	>77fJ$^{\Delta}$	134fJ$^{\Delta}$/30pJ$^{\blacktriangle}$ (1.3V)	>11.3nJ$^{\Delta}$/26.6nJ$^{\blacktriangle}$	26pJ$^{\blacktriangle}$ (0.775V)	>23.6pJ$^{\Delta}$ (0.75V)	8.4pJ$^{\Delta}$/12.7pJ$^{\blacktriangle}$ (0.55V)	30pJ$^{\Delta}$/51pJ$^{\blacktriangle}$ (0.8V)
	norm.		–	–	–	>2.4nJ$^{\Delta}$/5.7nJ$^{\blacktriangle}$	26pJ$^{\blacktriangle}$	(66.1pJ$^{\Delta}$)	8.4pJ$^{\Delta}$/12.7pJ$^{\blacktriangle}$	12.9pJ$^{\Delta}$/22pJ$^{\blacktriangle}$

Power

ODIN has the lowest energy per synaptic event among all digital SNNs,
MorphIC keeps a competitive energy efficiency.
They outperform subthreshold analog SNNs in accelerated time,
but not for biological-time processing.

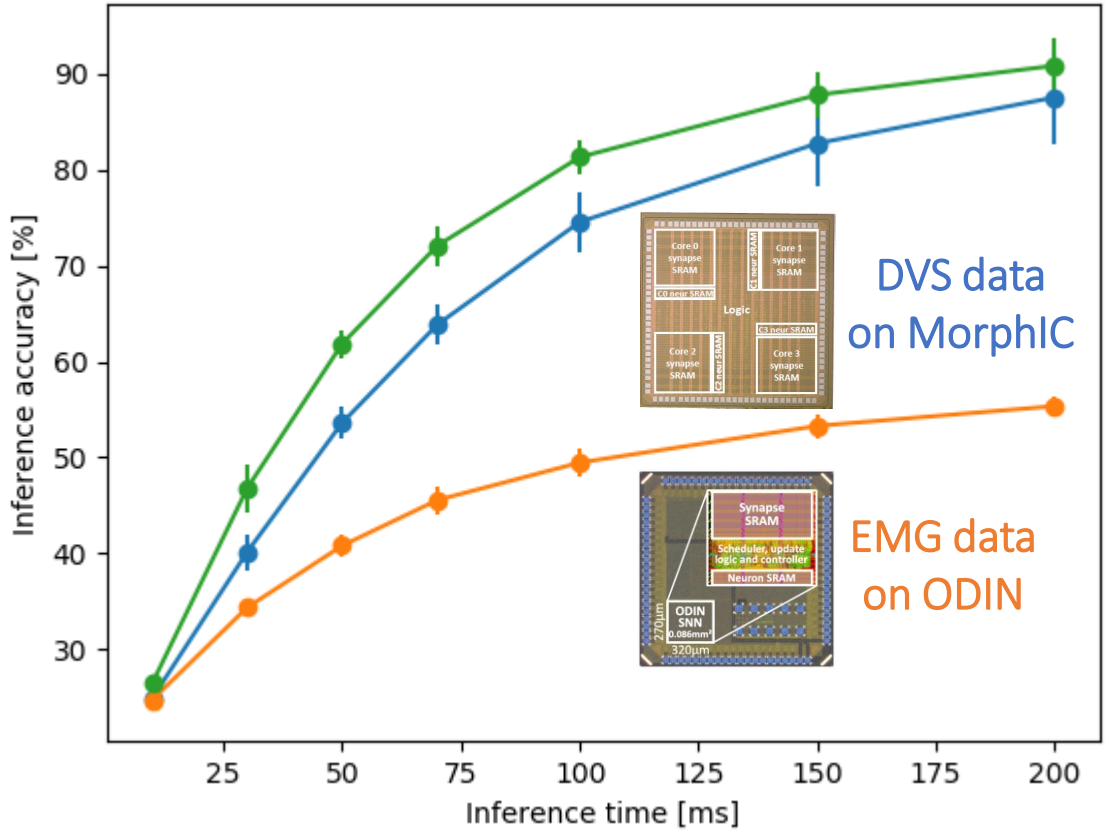
Results on the spiking EMG/DVS sensor fusion benchmark

[Ceolini, Frenkel, Shrestha et al., *Front. Neurosci.*, 2020]



Results on the spiking EMG/DVS sensor fusion benchmark

[Ceolini, Frenkel, Shrestha et al., *Front. Neurosci.*, 2020]



Sensor fusion

ODIN+MorphIC	89.4% /	37.4µJ
Loihi	96% /	1105µJ
Software	95.4% /	32100µJ

Accuracy / Energy tradeoff

Neuromorphic designs are more efficient than GPUs, as would be expected from dedicated hardware. But are they more efficient than conventional accelerators?

→ *perspectives*

See the ODIN and MorphIC papers for more benchmarking, incl. online- and offline-trained MNIST.

Outline

Part I – Bottom-up neuromorphic design

- Building blocks
- Integration

Part II – Top-down neuromorphic design

- Algorithms

  Minimizing the training cost of neural networks for adaptive edge computing

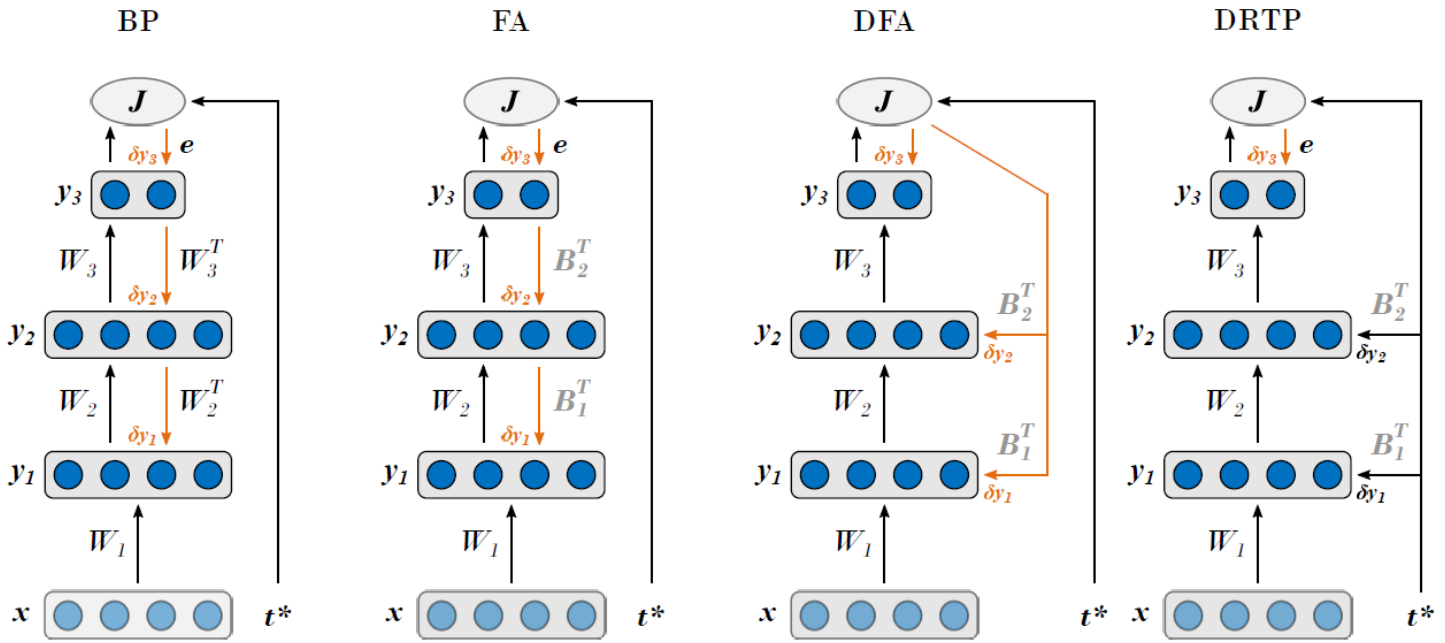
  [Frenkel & Lefebvre, *Front. Neurosci.*, 2021]

- Integration

Conclusion and perspectives

Learning without feedback

Releasing the weight transport and update locking of backprop



	δy_k	$\frac{\partial J}{\partial y_k} = W_{k+1}^T \delta z_{k+1}$	$B_k^T \delta z_{k+1}$	$B_k^T e$	$B_k^T t^*$
Weight-transport-free		×	✓	✓	✓
Update-unlocked		×	×	×	✓

Feedforward local training

↘ Computational and memory cost ↘

[Lillicrap, *Nat. Comms.*, 2016] [Nokland, *NeurIPS*, 2016]

Direct Random Target Projection (DRTP)

Ideal use cases?

Adaptive edge computing

- Very low power and area overheads can be expected for an on-chip implementation.
- Datasets representative of the complexity associated to autonomous smart sensors: MNIST or CIFAR-10.

→ We'll verify these claims *in silico*.

Disclaimer: whether DRTP scales to ImageNET is probably **not** the right question. ☺

Neuroscience

DRTP could come in line with recent findings in cortical areas that reveal the existence of output-independent target signals in the dendritic instructive pathways of intermediate-layer neurons.

[Magee & Grienberger, *Annual Review of Neuroscience*, 2020]

Outline

Part I – Bottom-up neuromorphic design

- Building blocks
- Integration

Part II – Top-down neuromorphic design

- Algorithms
- Integration

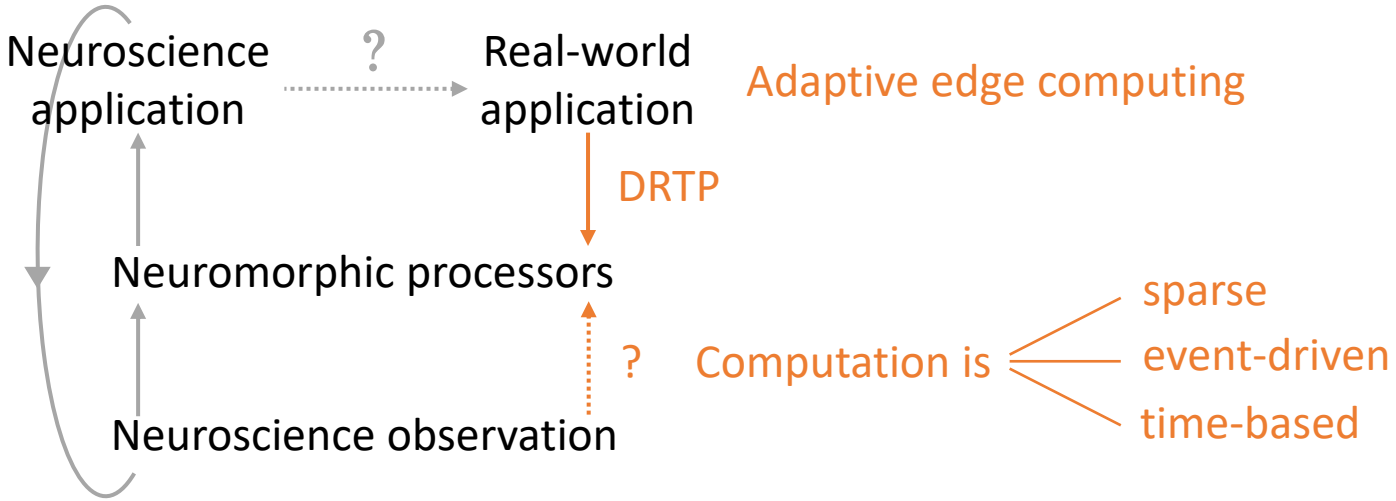
Neuromorphic accelerators

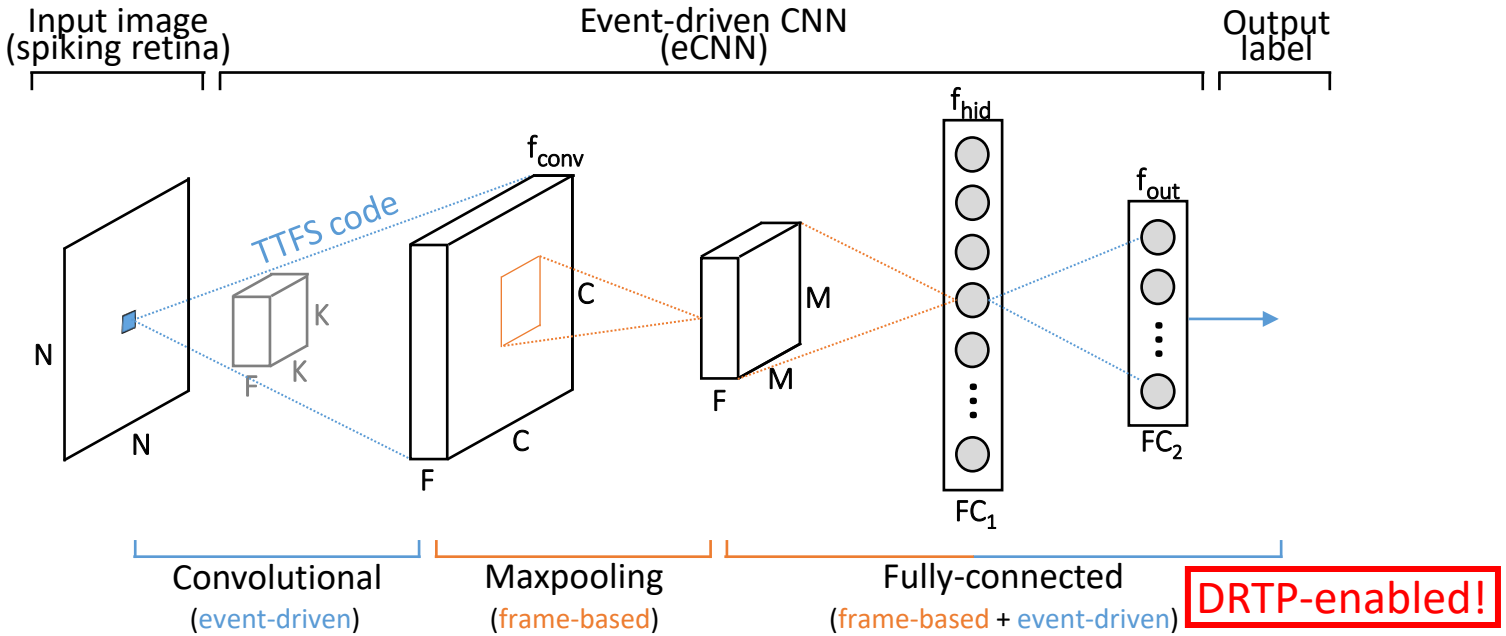
[Frenkel, *ISCAS*, 2020]
(*Best paper award* 🏆)

Conclusion and perspectives

Which bio-inspired elements?

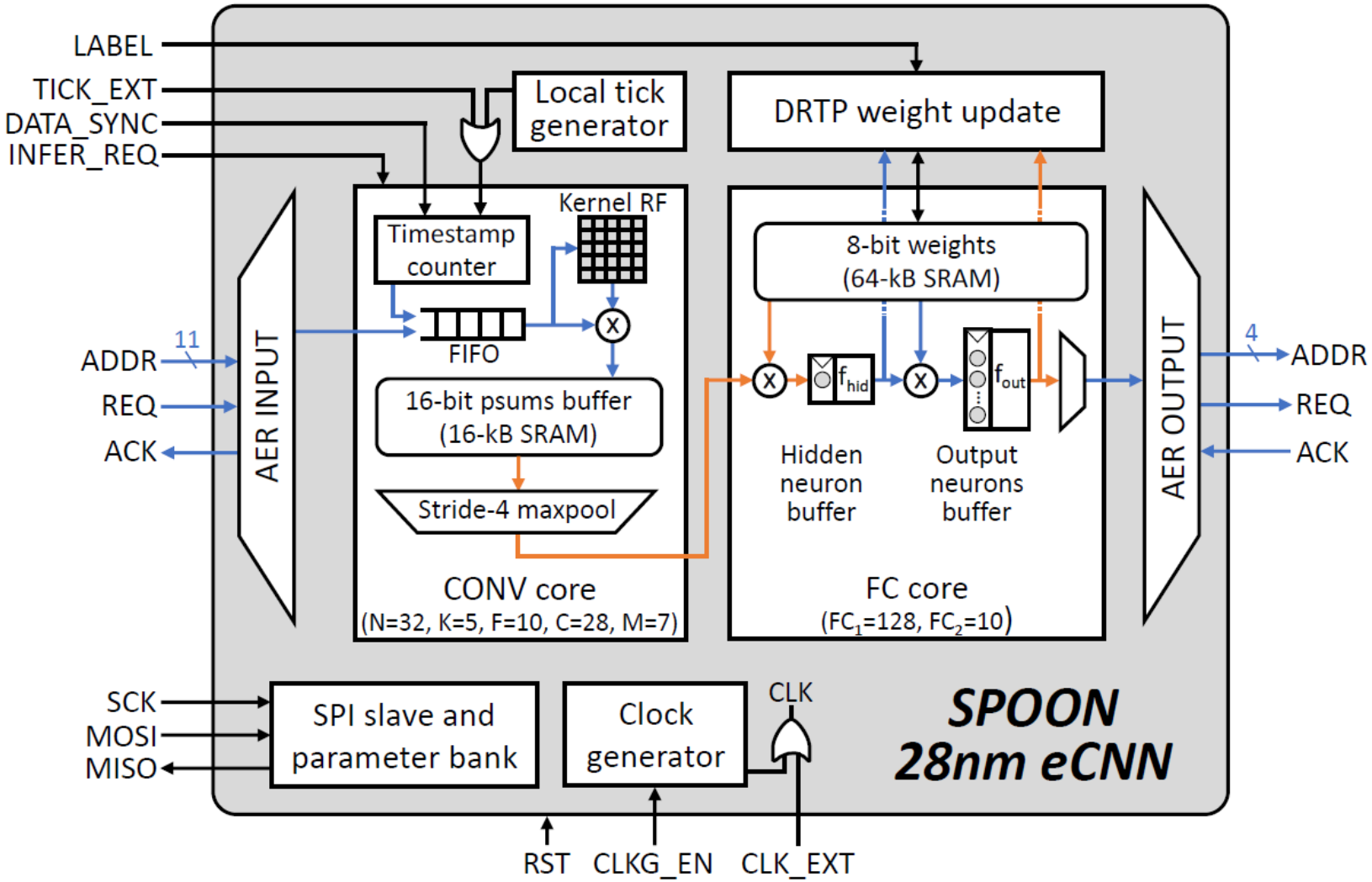
Taking a step back with the top-down design strategy



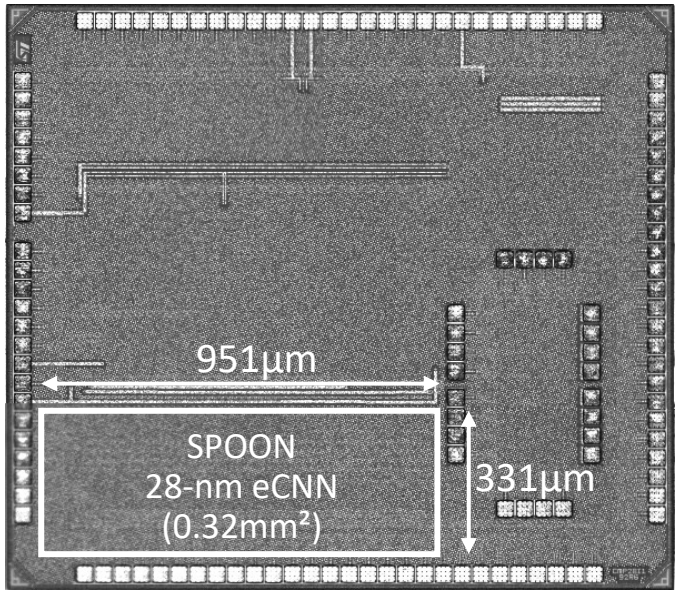


Architecture of SPOON

SPOON – A Spiking Online-Learning Convolutional Neuromorphic Processor



SPOON – Chip microphotograph and specifications



(pre-silicon numbers, not yet updated)

Technology	28nm FDSOI CMOS
Implementation	Digital
Area	0.32mm^2 (0.26mm^2 excl. rails)
Topology	C5×5@10–FC128–FC10
Online learning	Stochastic DRTP, 8-bit weights
Time constant	Biological to accelerated
Supply voltage	0.6V – 1.0V
Max. clock frequency	150MHz
Leakage power	61μW at 0.6V
Energy for CONV core	1.7nJ/event at 0.6V
Energy for FC core	55nJ/inference at 0.6V
Online learning overhead	16.8% in power, 11.8% in area

→ *perspectives*

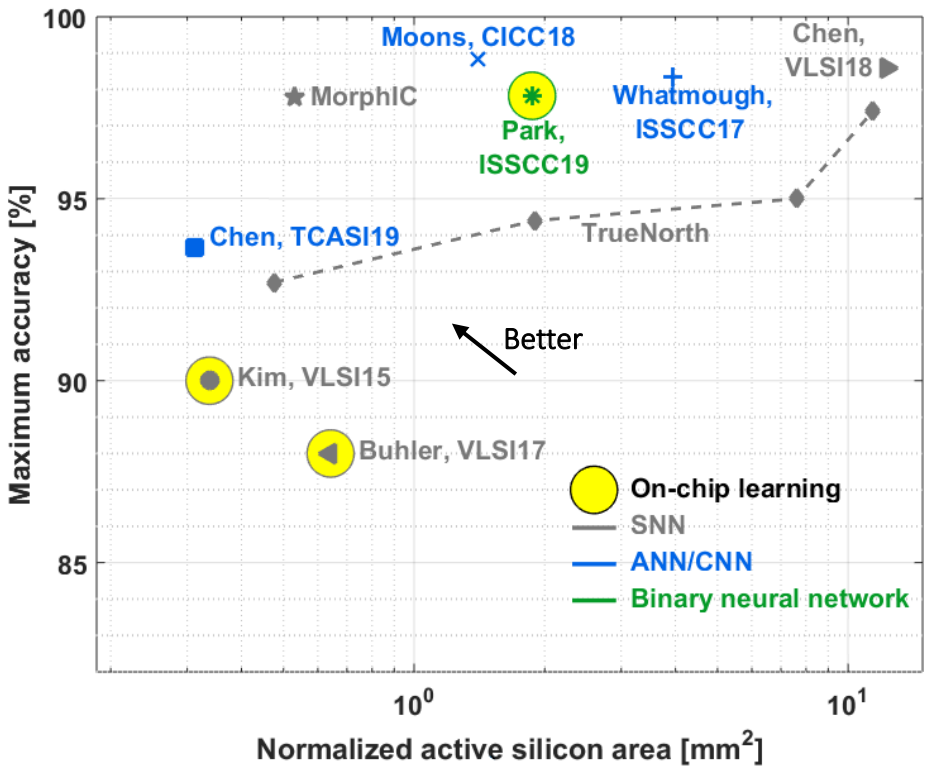
Stay tuned for the journal extension!

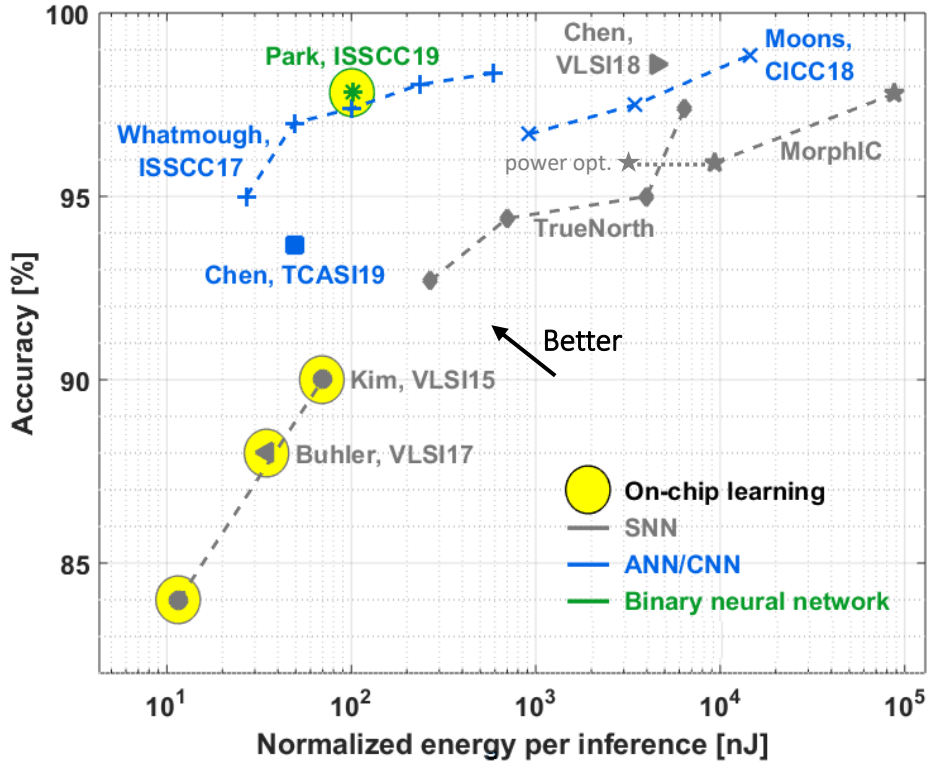
DRTP can be implemented on-chip at a very low cost!

Benchmarking: MNIST and N-MNIST

SPOON benchmarking

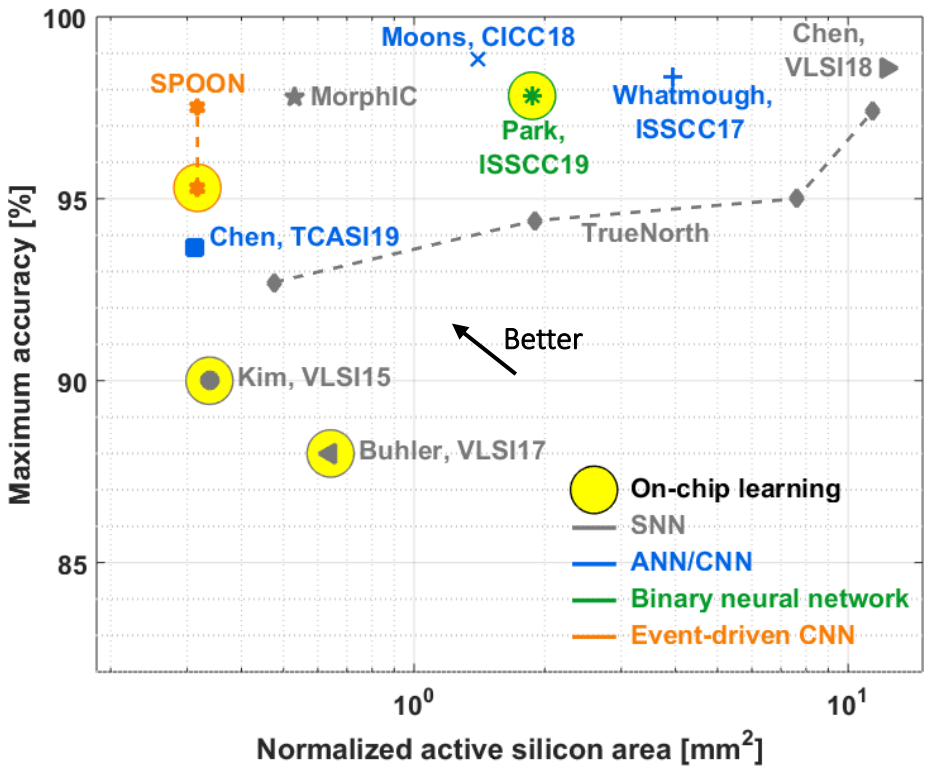
Against SoA spiking neural networks on MNIST

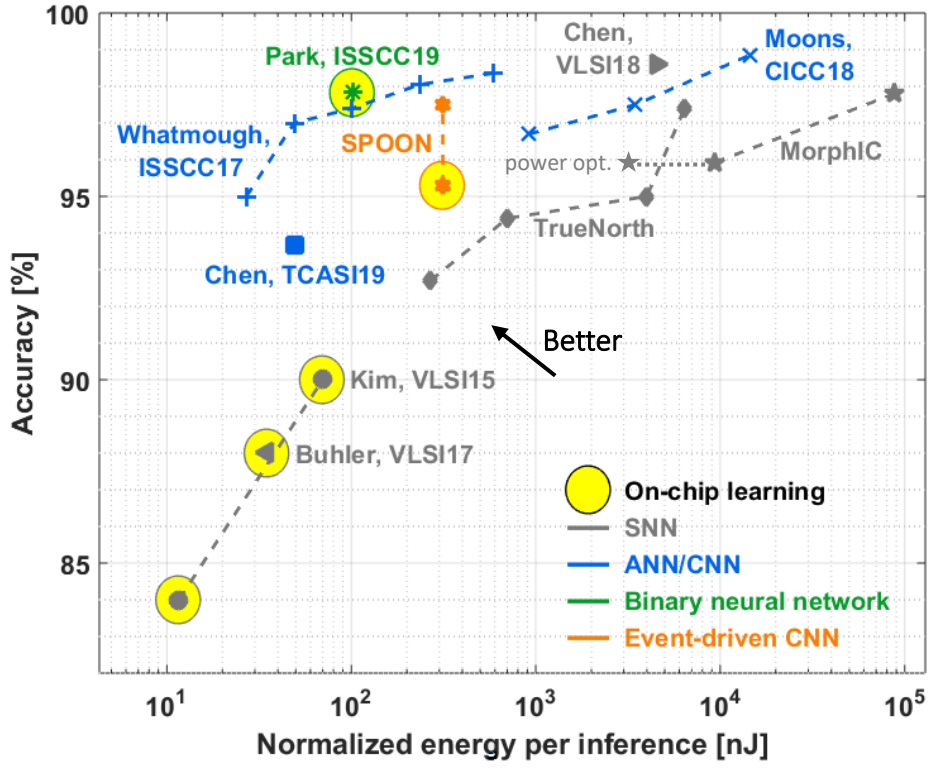




SPOON benchmarking

Against SoA spiking neural networks on MNIST





Only SPOON allows reaching the efficiency of ANN/CNN/BNN accelerators while enabling online learning with event-based sensors.

Outline

Part I – Bottom-up neuromorphic design

- Building blocks
- Integration

Part II – Top-down neuromorphic design

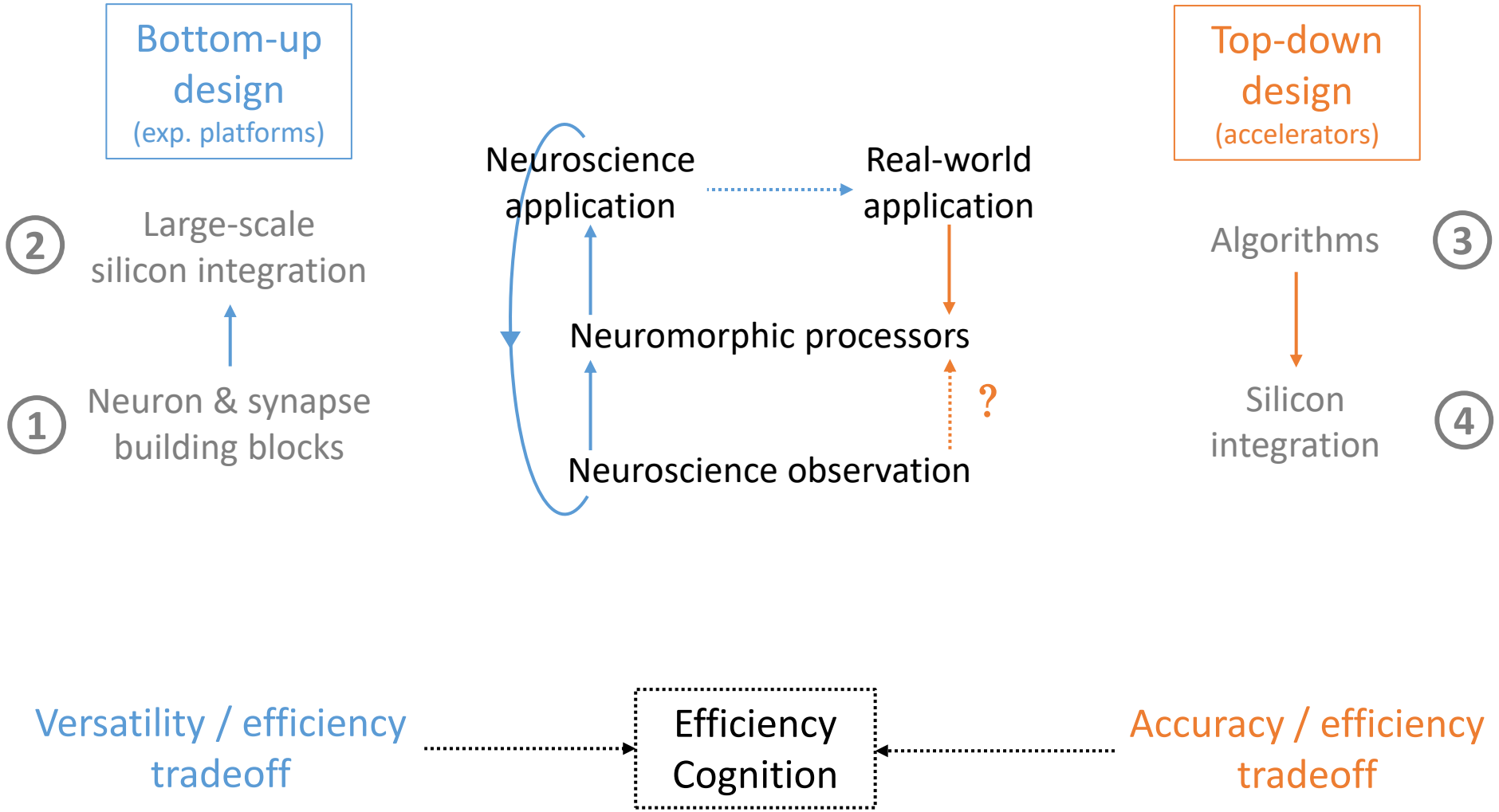
- Algorithms
- Integration

Conclusion and perspectives

Summary of the key messages, next directions

Neuromorphic Engineering – Key Claims

Unveiling roads to embedded cognition



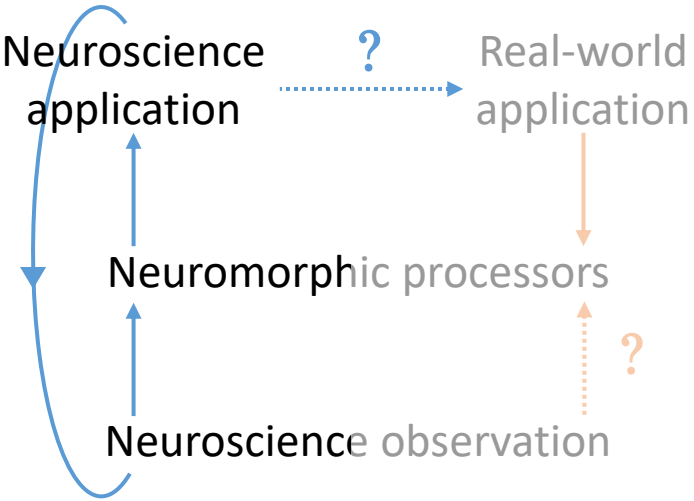
Neuromorphic Engineering – Key Claims

Unveiling roads to embedded cognition

Bottom-up design
(exp. platforms)

② Large-scale silicon integration

① Neuron & synapse building blocks



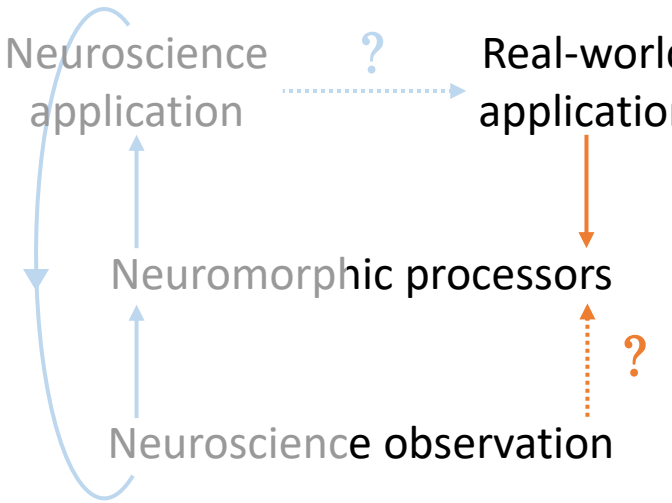
Versatility / efficiency tradeoff

Claim 1

Hardware-aware neuroscience model design and selection allows reaching record neuron and synapse densities with low-power operation for large-scale integration *in silico*.

Neuromorphic Engineering – Key Claims

Unveiling roads to embedded cognition



Top-down design
(accelerators)

Algorithms ③

Silicon integration ④

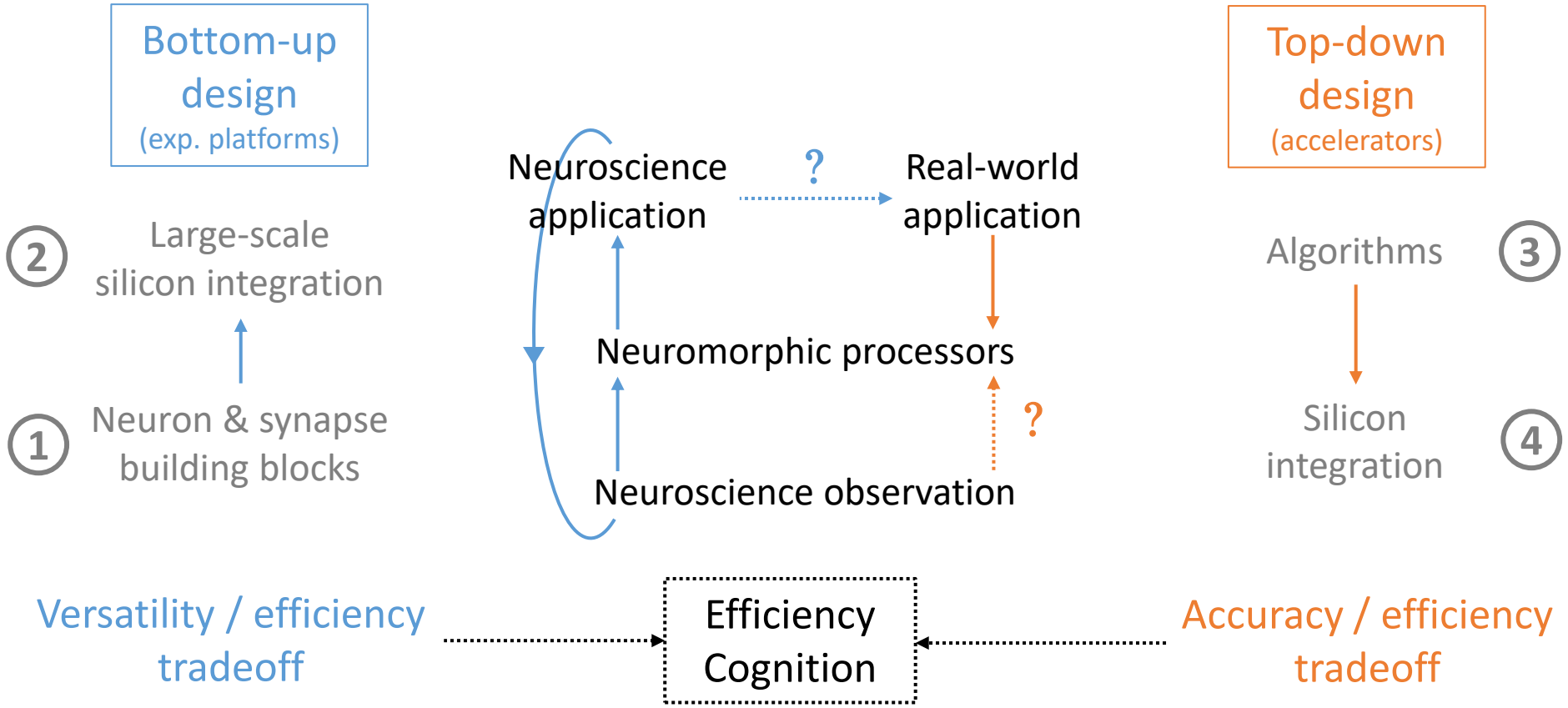
Claim 2

Combining event-driven and frame-based processing with weight-transport-free update-unlocked training supports low-cost adaptive edge computing with spike-based sensors.

Accuracy / efficiency tradeoff

Neuromorphic Engineering – Key Claims

Unveiling roads to embedded cognition



Claim 3

Top-down guidance helps pushing bottom-up neuron and synapse integration beyond the purpose of neuroscience experimentation platforms, while bottom-up guidance supports top-down design toward brain reverse-engineering.

Perspectives

- Neuromorphic engineering and spiking neural networks:
  "Can we make it work?" → "Will it bring a competitive advantage?" (not only against GPUs)
  Need something better than MNIST → Audio (KWS) and bio-signal processing (time, biological-time)
  [Davies, Nat. Mach. Intel., 2019]

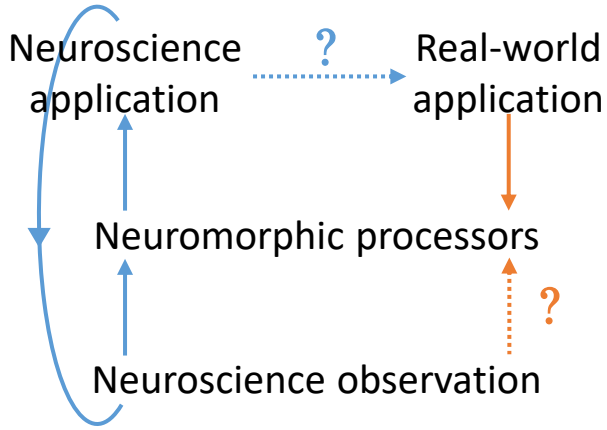
- Phenomenological digital design: pragmatic short-to-midterm approach.
  Promising avenues: leveraging the variability of subthreshold analog design; fine-grained mixed-signal design.

- Bottom-up trend: dendrites

- Top-down trend: new wave of training algorithms mapping onto bio-plausible primitives

[Sacramento, NeurIPS'18]
[Payeur, bioRxiv, 2020]
[Bellec, Nat. Comms., 2020]

- Cognition: a case for neuromorphic robots?
  [Man & Damasio, Nat. Mach. Intel., 2019]



Acknowledgments

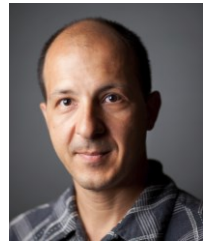
PhD → Postdoc

Institution:

Funding:

Supervisors:

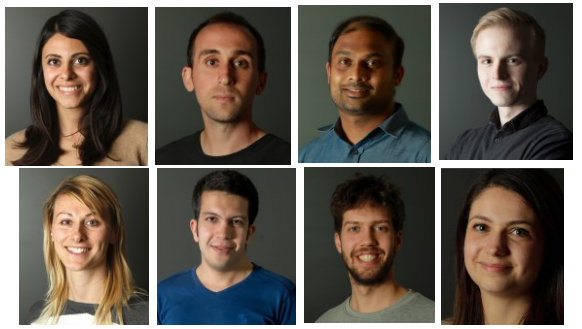

Profs. David Bol, Jean-Didier Legat



Prof. Giacomo Indiveri

Key colleagues:

Martin Lefebvre



Questions?

@C_Frenkel
cfrenkel
Charlotte-Frenkel
ChFrenkel
charlotte@ini.uzh.ch

Main references:

- ODIN: [C. Frenkel et al., "A 0.086-mm² 12.7-pJ/SOP 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28nm CMOS," *IEEE Trans. BioCAS*, 2019]
- MorphIC: [C. Frenkel et al. "MorphIC: A 65-nm 738k-synapse/mm² quad-core binary-weight digital neuromorphic processor with stochastic spike-driven online learning," *IEEE Trans. BioCAS*, 2019]
- DRTP: [C. Frenkel, M. Lefebvre et al., "Learning without feedback: Fixed random learning signals allow for feedforward training of deep neural networks," Frontiers in Neuroscience, 2021]
- SPOON: [C. Frenkel et al., "A 28-nm convolutional neuromorphic processor enabling online learning with spike-based retinas," *IEEE ISCAS*, 2020]
- Review: [C. Frenkel, D. Bol and G. Indiveri, "Bottom-up and top-down neural processing systems design: Neuromorphic intelligence as the convergence of natural and artificial intelligence", *arXiv preprint arXiv:2106.01288, 2021]*

Open-sourced!
github.com/ChFrenkel/ODIN

Open-sourced!
github.com/ChFrenkel/DirectRandomTargetProjection

Journal extension coming soon

Just released!

TINY ML EMEA
Virtual
JUNE 7-10,2021

Premier Sponsor

Intelligent Agent
Neuton

Automated TinyML

Zero-code SaaS solution

Create tiny models, ready for embedding, in just a few clicks!

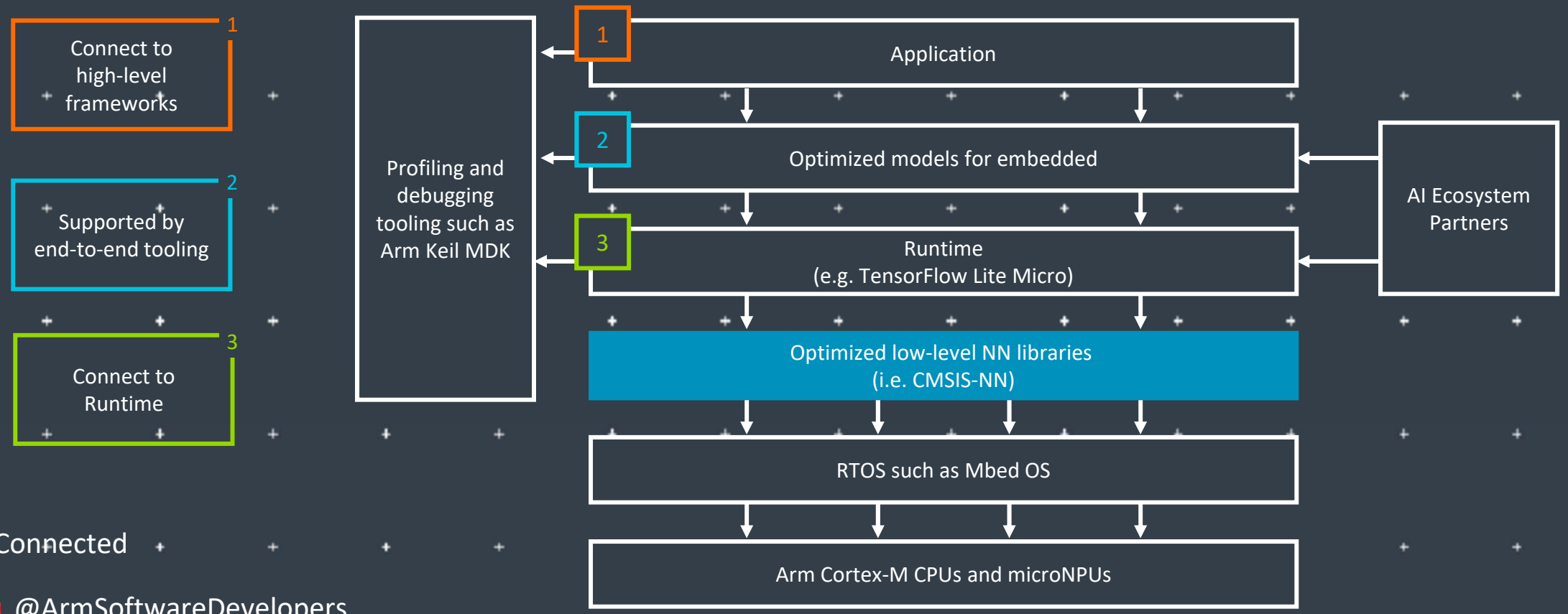
Compare the benchmarks of our compact models to those of TensorFlow and other leading neural network frameworks.

Build Fast. Build Once. Never Compromise.

TINY ML EMEA
Virtual
JUNE 7-10,2021

Executive Sponsors

Arm: The Software and Hardware Foundation for tinyML



Stay Connected

@ArmSoftwareDevelopers

@ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm

arm

TinyML for all developers

C++ library

Arduino library

WebAssembly

Dataset

Acquire valuable training data securely

Enrich data and train ML algorithms

Edge Device

Real sensors in real time

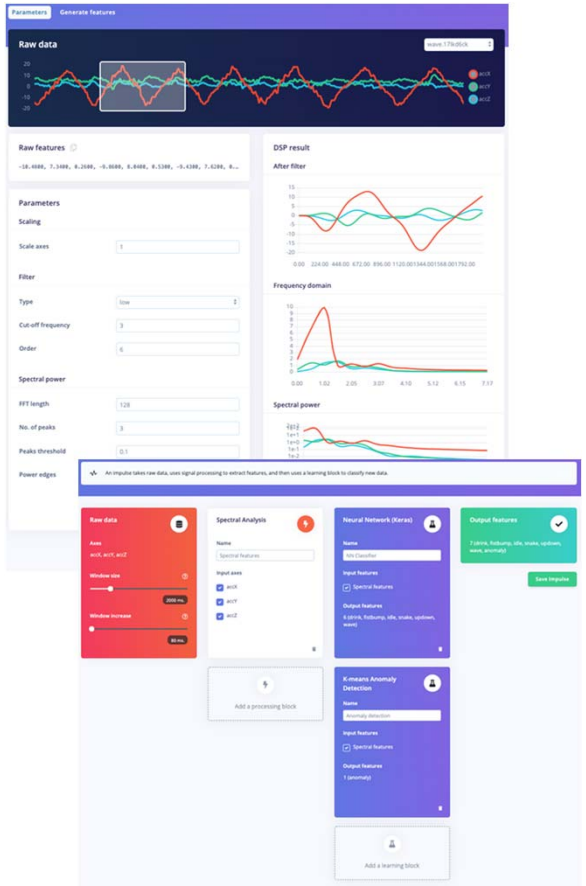
Open source SDK

Impulse

Embedded and edge compute deployment options

Test impulse with real-time device data flows

Test



www.edgeimpulse.com

Qualcomm
AI research

Advancing AI research to make efficient AI ubiquitous

Power efficiency

Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization

Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning

Robust learning through minimal data, unsupervised learning, on-device learning

A platform to scale AI across the industry

Perception

Object detection, speech recognition, contextual fusion

Reasoning

Scene understanding, language understanding, behavior prediction

Action

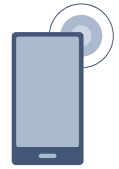
Reinforcement learning for decision making

Edge cloud

Cloud

IoT/IIoT

Automotive



Mobile

SYNTIANT

Syntiant Corp. is moving artificial intelligence and machine learning from the cloud to edge devices. Syntiant's chip solutions merge deep learning with semiconductor design to produce ultra-low-power, high performance, deep neural network processors. These network processors enable always-on applications in battery-powered devices, such as smartphones, smart speakers, earbuds, hearing aids, and laptops. Syntiant's Neural Decision Processors™ offer wake word, command word, and event detection in a chip for always-on voice and sensor applications.

Founded in 2017 and headquartered in Irvine, California, the company is backed by Amazon, Applied Materials, Atlantic Bridge Capital, Bosch, Intel Capital, Microsoft, Motorola, and others. Syntiant was recently named a CES® 2021 Best of Innovation Awards Honoree, shipped over 10M units worldwide, and unveiled the NDP120 part of the NDP10x family of inference engines for low-power applications.

www.syntiant.com

@Syntiantcorp

TINY ML EMEA
Virtual
JUNE 7-10,2021

Platinum Sponsors

Part of your life. Part of tomorrow.

www.infineon.com

Reality AI®

Add Advanced Sensing to your Product with Edge AI / TinyML

https://reality.ai info@reality.ai @SensorAI Reality AI

Pre-built Edge AI sensing modules, plus tools to build your own

Reality AI solutions

Prebuilt sound recognition models for indoor and outdoor use cases

Solution for industrial anomaly detection

Pre-built automotive solution that lets cars "see with sound"

Reality AI Tools® software

Build prototypes, then turn them into real products

Explain ML models and relate the function to the physics

Optimize the hardware, including sensor selection and placement

TINY ML EMEA
Virtual
JUNE 7-10,2021

Gold Sponsors

LatentAI

Adaptive AI for the Intelligent Edge

Latentai.com

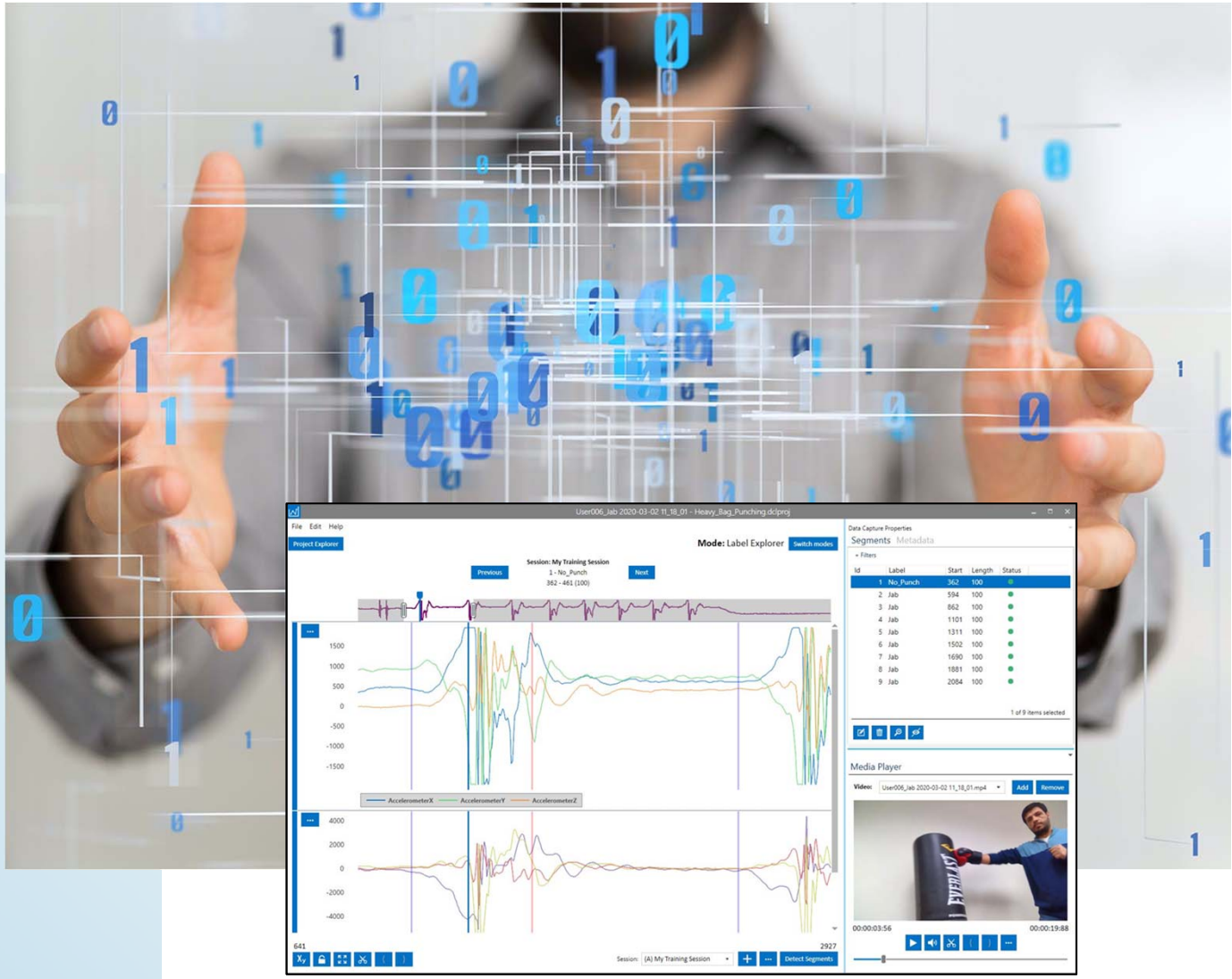
SensiML™

Build Smart IoT Sensor Devices From Data

SensiML pioneered TinyML software tools that auto generate AI code for the intelligent edge.

- End-to-end AI workflow
- Multi-user auto-labeling of time-series data
- Code transparency and customization at each step in the pipeline

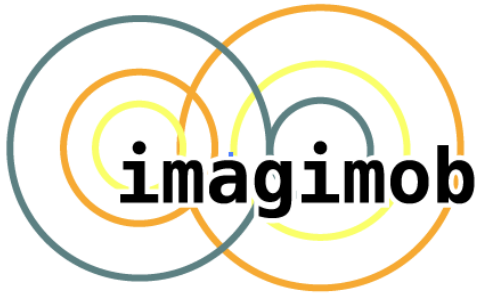
We enable the creation of production-grade smart sensor devices.



sensiml.com

TINY ML EMEA Virtual JUNE 7-10,2021

Silver Sponsors



Copyright Notice