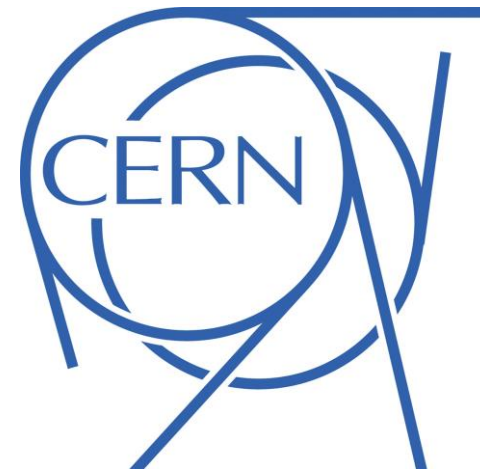# Anomaly Detection with Spiking Neural Networks on Neuromorphic Chips

- Bartlomiej Borzyszkowski *(Gdansk University of Technology, Intel Poland)*

- Eric Moreno *(Caltech)*
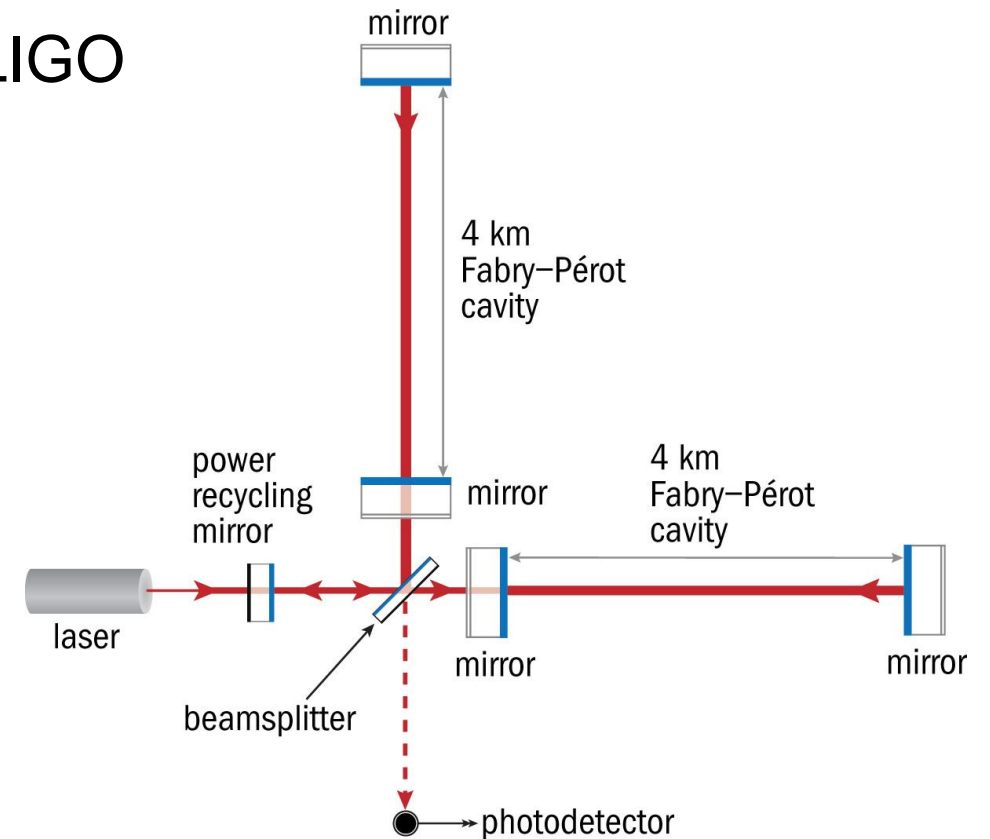
Mentors: Jean-Roch Vlimant *(Caltech)*, Maurizio Pierini *(CERN)*

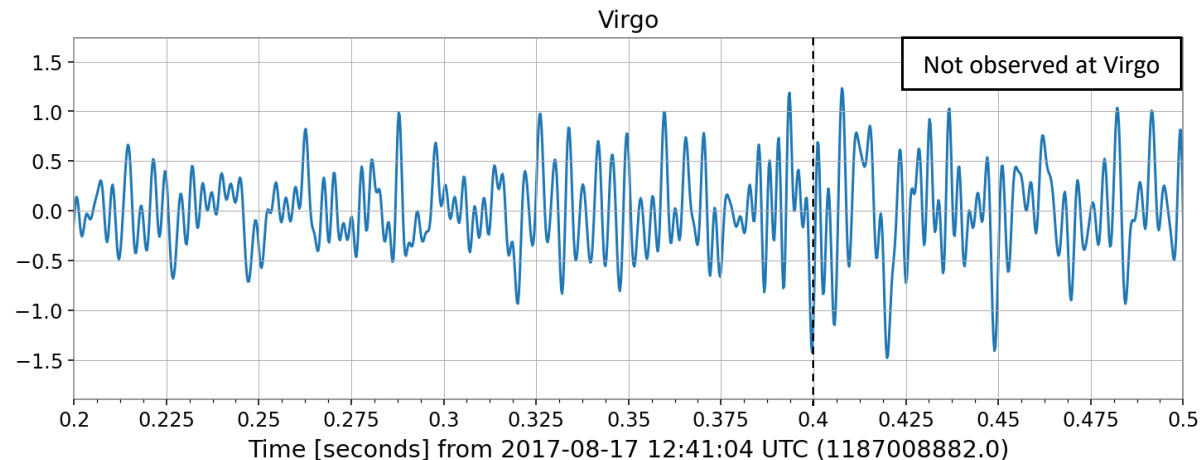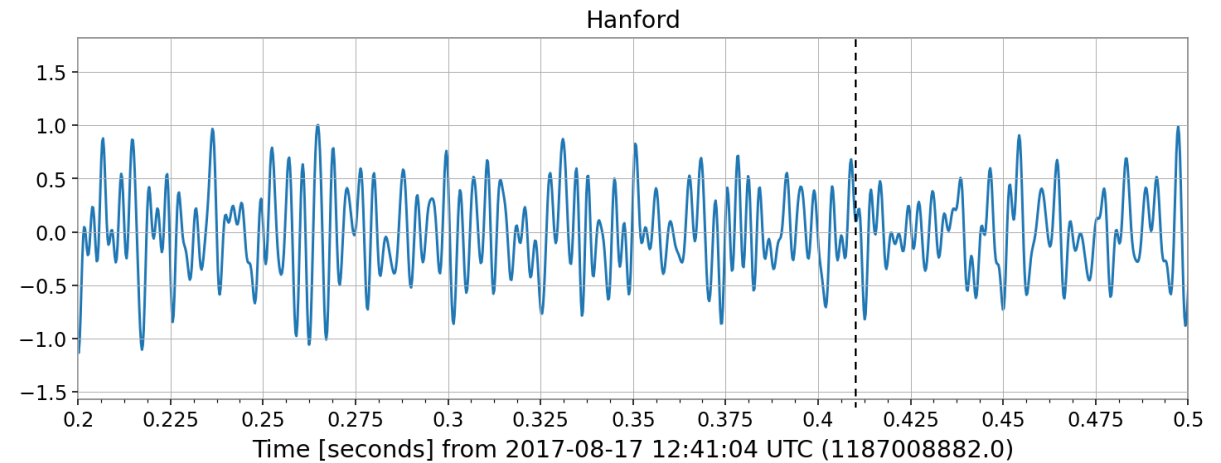*December 2nd, 2020;   Fast Machine Learning for Science*

Introduction to the project

- Detection of gravitational waves (GWs) at LIGO

mirror

4 km
Fabry–Pérot
cavity

power
recycling
mirror

mirror

4 km
Fabry–Pérot
cavity

laser

beamsplitter

mirror

mirror

photodetector

Produces: 1-D time-series strain

LIGO Data Generation – GW170817 example
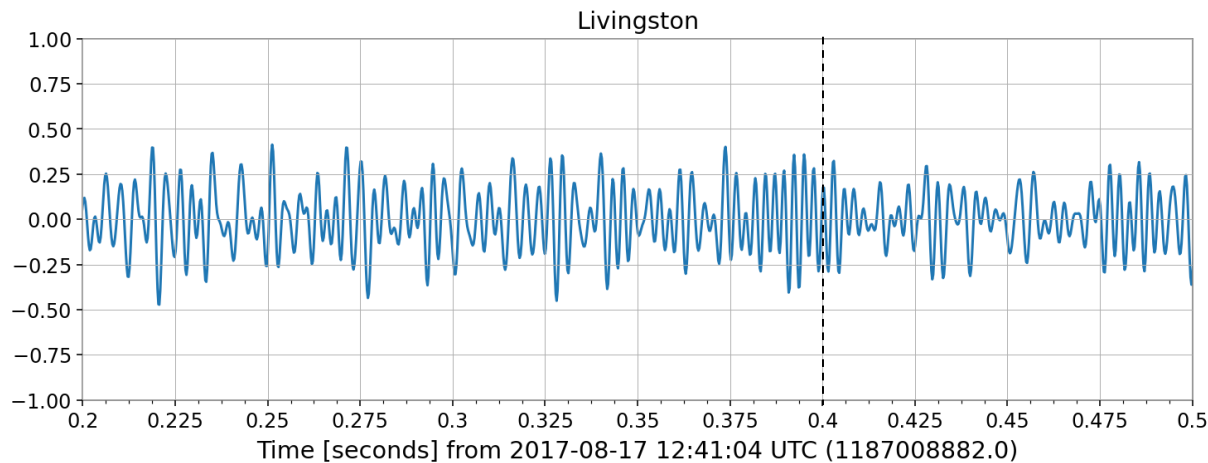
Data is whitened and bandpassed

Source: https://www.gw-openscience.org/data/

Current methods

Matched Filtering (used at LIGO)

- Compares incoming GW data to bank of simulated waveforms

- Can only identify GWs that are available in GW banks (no exotic events)
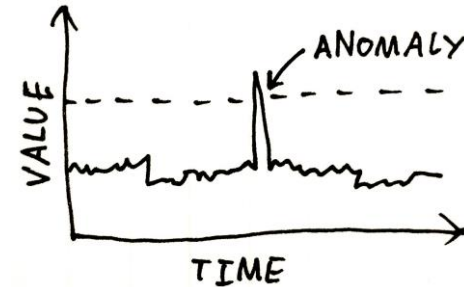
Deep Filtering (CNNs)

- Take time-series inputs, can determine detections and estimate parameters of events

- Still can miss events that aren't included in training set

Source: https://arxiv.org/pdf/1711.03121.pdf, https://arxiv.org/pdf/2005.06534.pdf
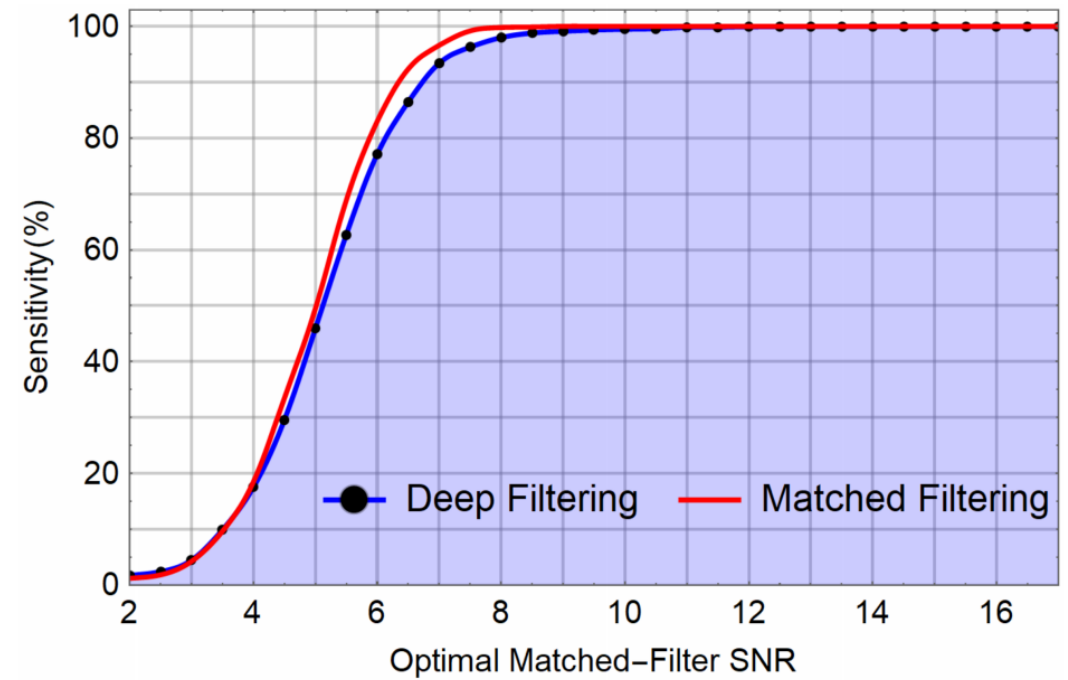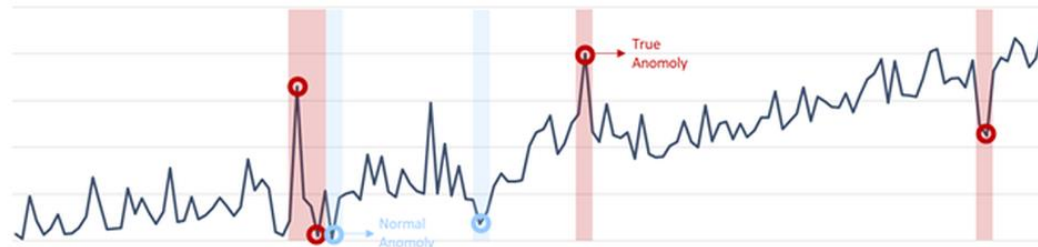
Artificial Neural Networks

1. *Unsupervised learning*
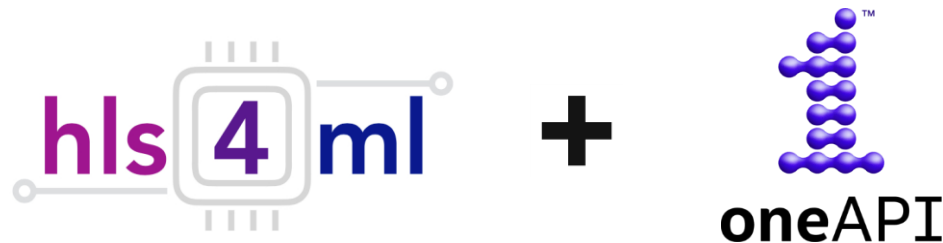   - *unlabeled data*
   - *autoencoders*

2. *Supervised learning*
   - *labels for gravitational waves*
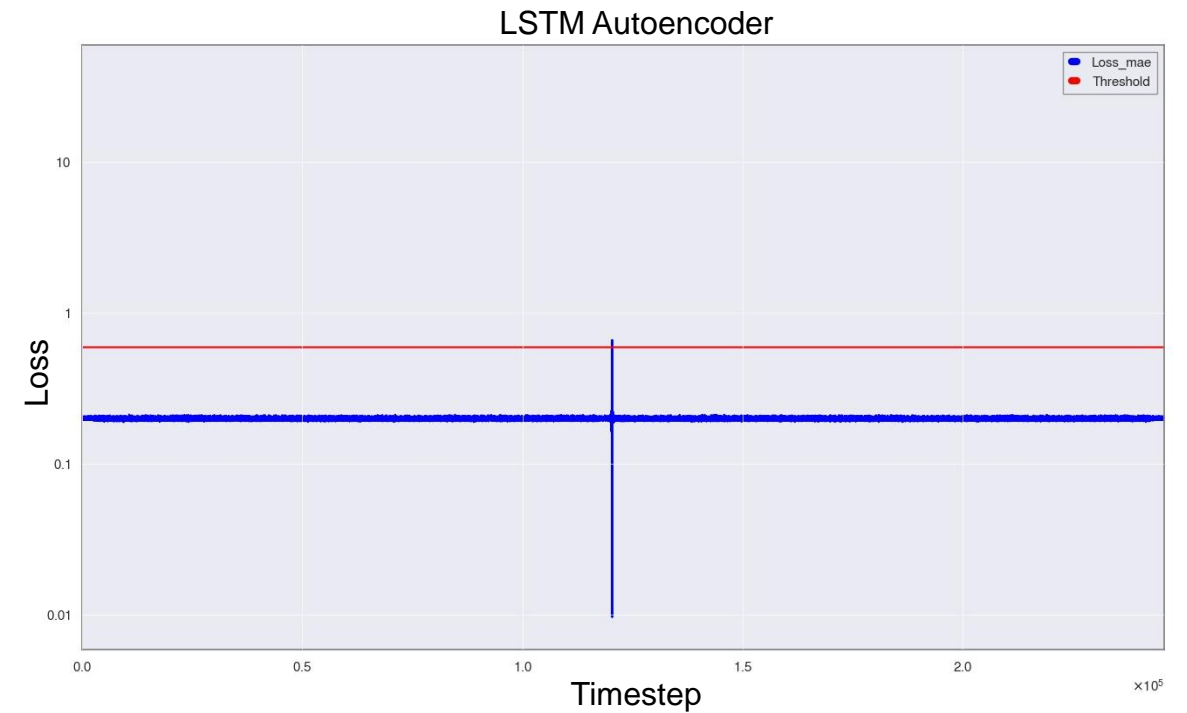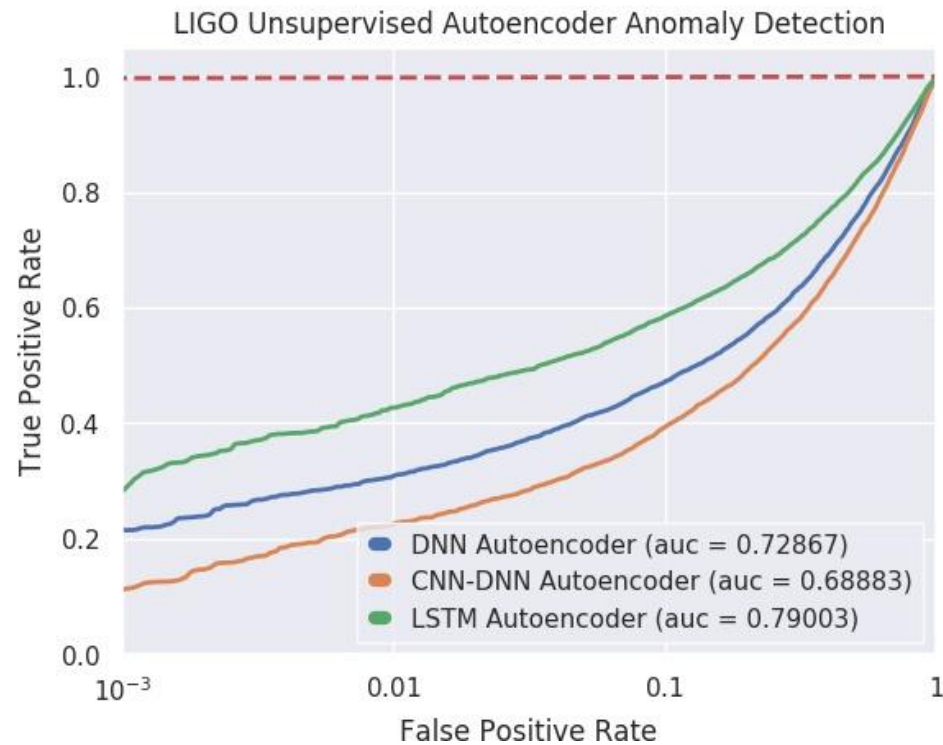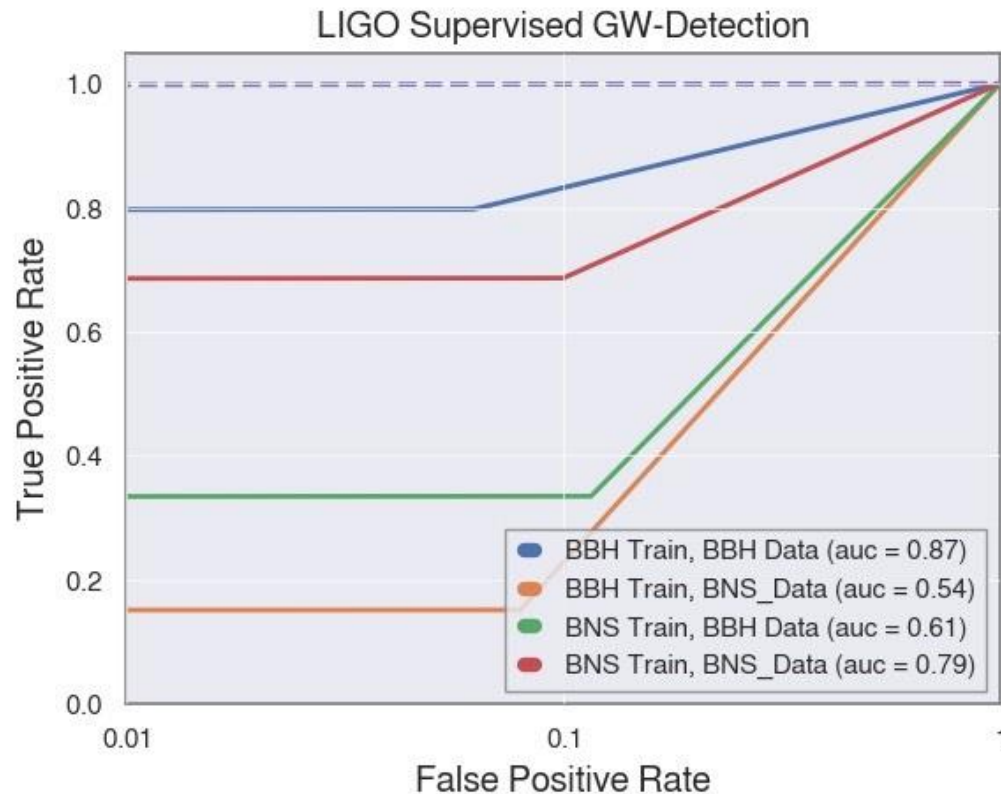   - *deep neural networks*

3. *Fast inference*
   - *hls4ml with oneAPI backend*
   - *hardware acceleration*

ANNs: Unsupervised learning

ANNs: Supervised learning

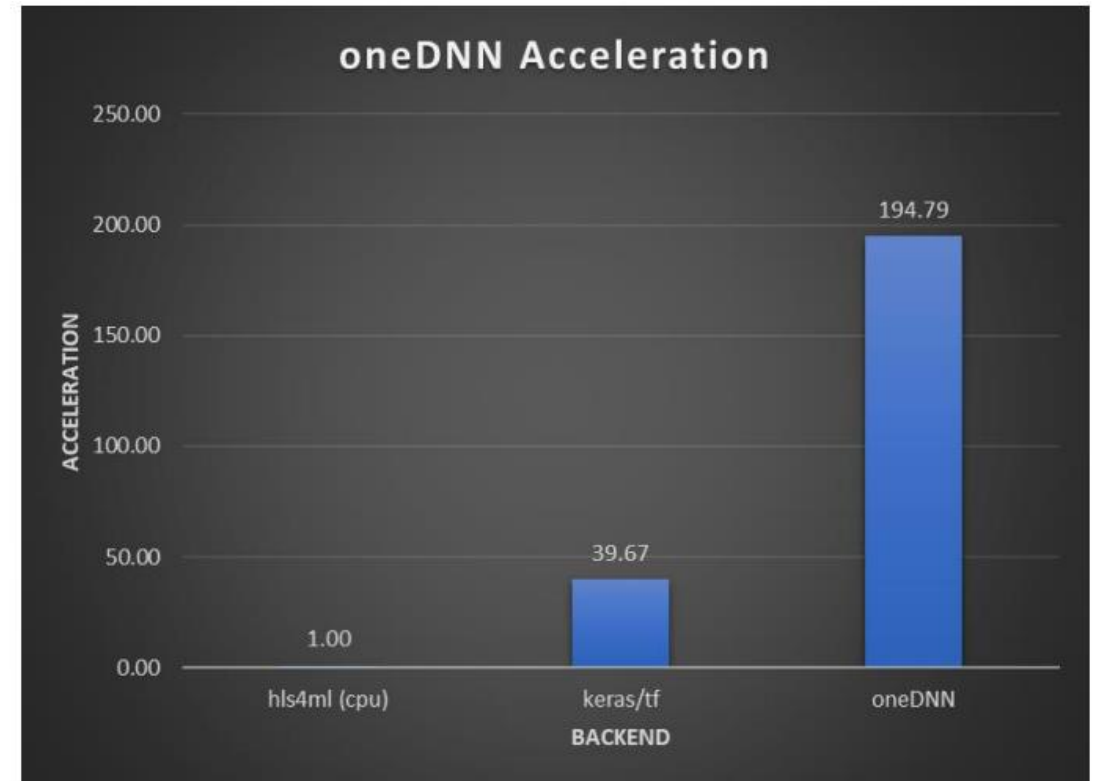LIGO Supervised GW-Detection

BBH Train, BBH Data (auc = 0.87)
BBH Train, BNS_Data (auc = 0.54)
BNS Train, BBH Data (auc = 0.61)
BNS Train, BNS_Data (auc = 0.79)

	Input	vector (size: 8192)
1	Reshape	matrix (size: 1 × 8192)
2	Convolution	matrix (size: 64 × 8177)
3	Pooling	matrix (size: 64 × 2044)
4	ReLU	matrix (size: 64 × 2044)
5	Convolution	matrix (size: 128 × 2014)
6	Pooling	matrix (size: 128 × 503)
7	ReLU	matrix (size: 128 × 503)
8	Convolution	matrix (size: 256 × 473)
9	Pooling	matrix (size: 256 × 118)
10	ReLU	matrix (size: 256 × 118)
11	Convolution	matrix (size: 512 × 56)
12	Pooling	matrix (size: 512 × 14)
13	ReLU	matrix (size: 512 × 14)
14	Flatten	vector (size: 7168)
15	Linear Layer	vector (size: 128)
16	ReLU	vector (size: 128)
17	Linear Layer	vector (size: 64)
18	ReLU	vector (size: 64)
19	Linear Layer	vector (size: 2)
	Output	vector (size: 2)

Inference engine with hls4ml and Intel® oneAPI

Anomaly Detection
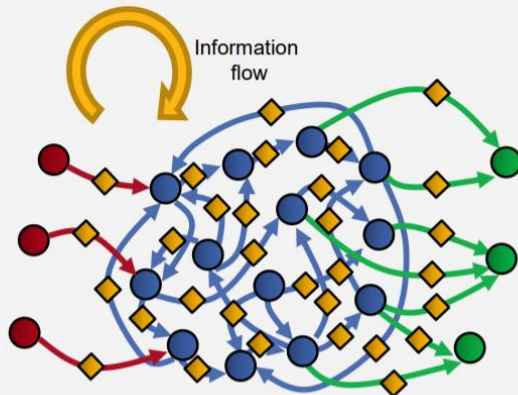runs in hls4ml with oneAPI backend

8
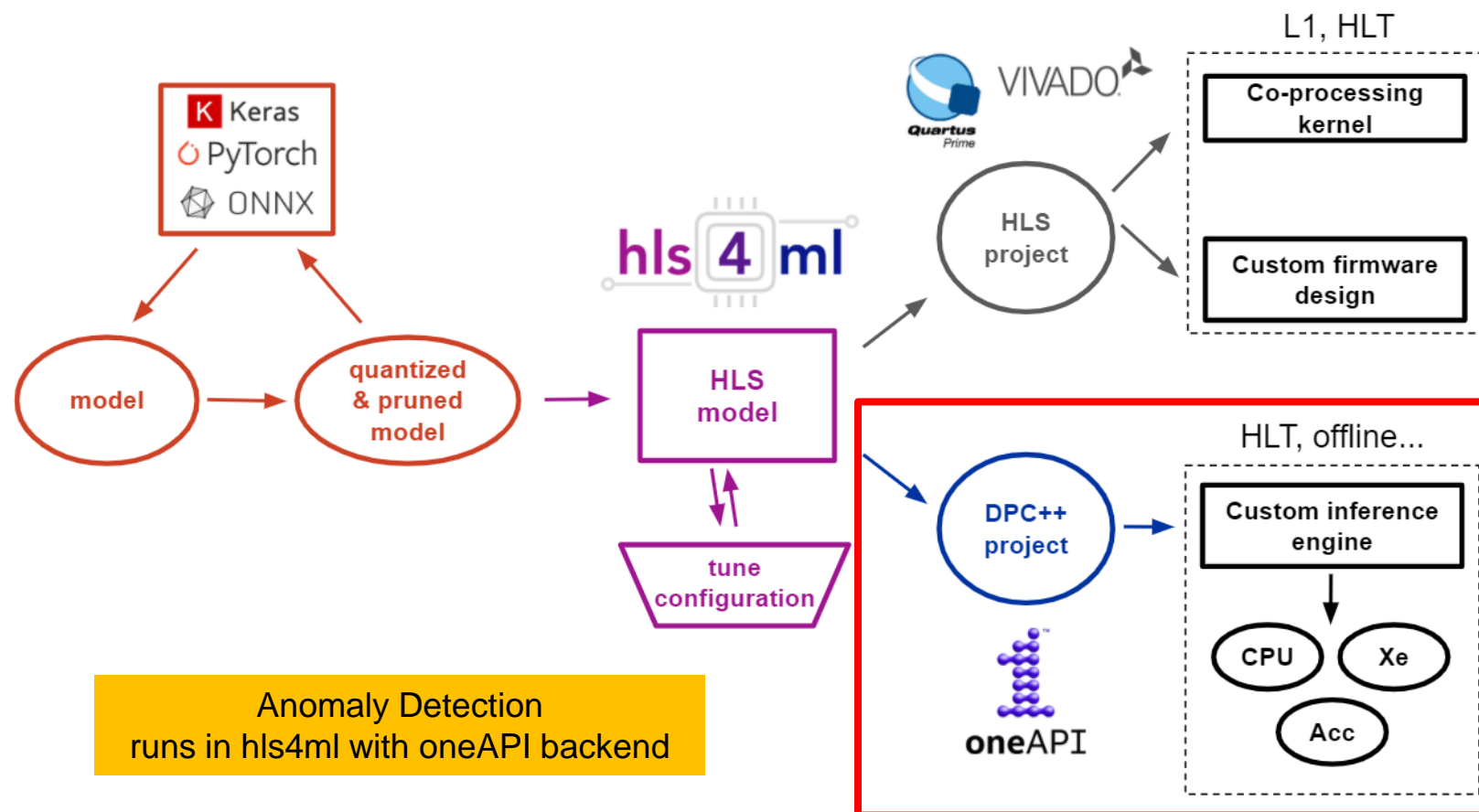
Intel® oneAPI baseline model performance

Three hidden-layer neural network on Xeon Gold 6128 3.4Ghz
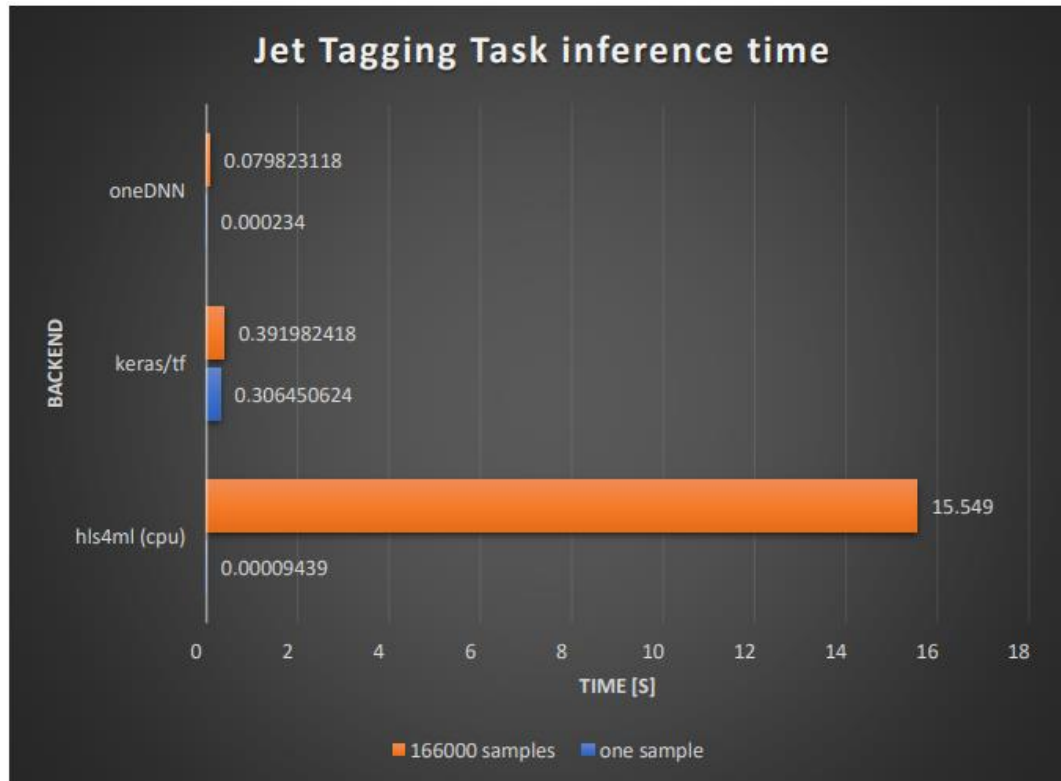
Spiking Neural Networks (SNNs)

Brain-like Neural network:

Deep Artificial Neural Network:

Information flow

Information flow

- Input Neurons
- Hidden Neurons
- Output Nodes
- Synaptic weights

- Omni-directional signal flow
- Asynchronous pulse signals
- Information encoded in signal timing
- Inspired by biology
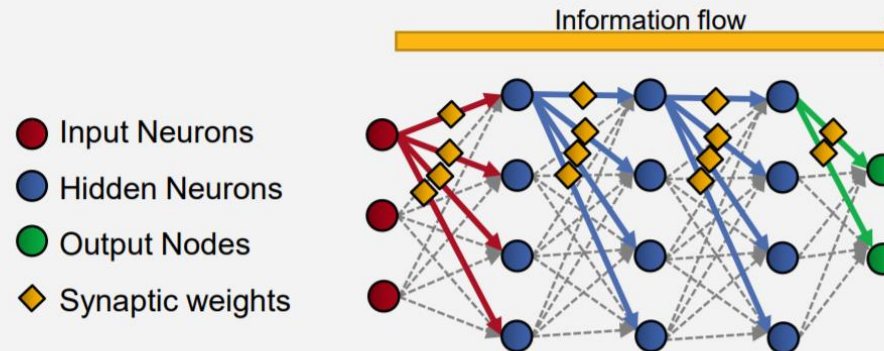
- Feed-forward sequential processing
- Use activation functions
- Information encoded in signal amplitude
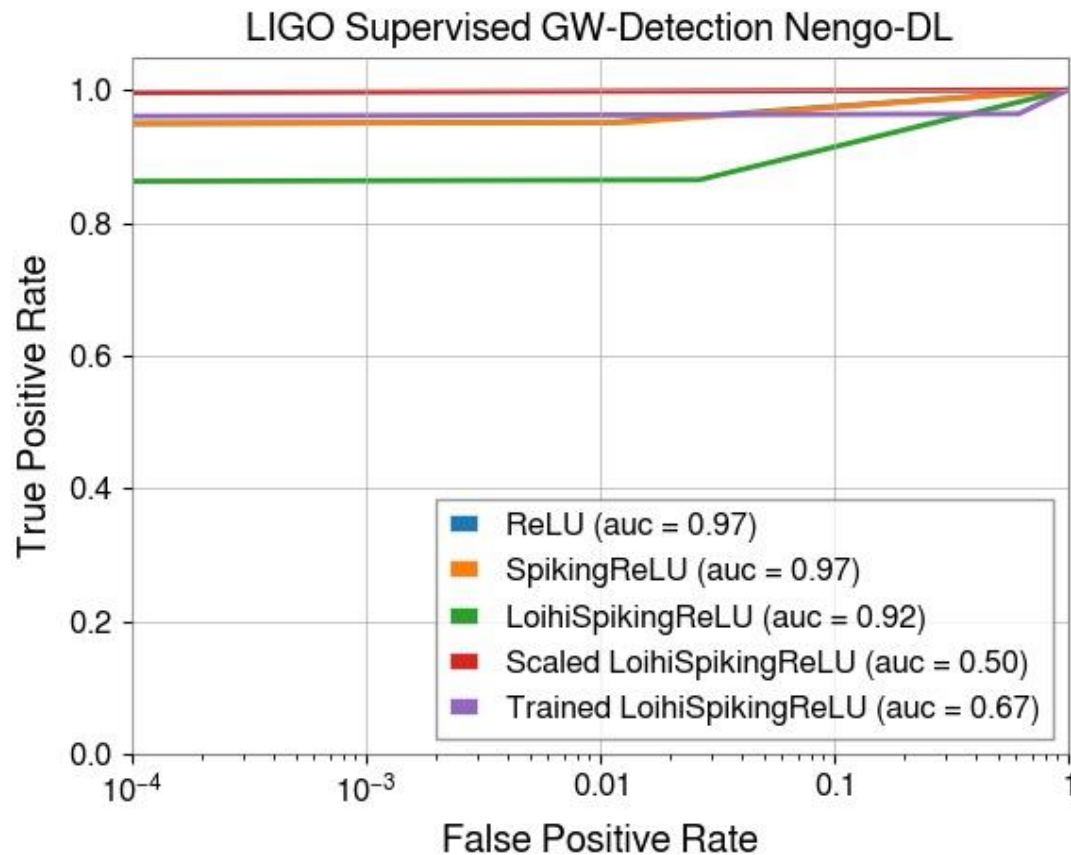- Trained with backpropagation algorithms

Intel Loihi – Dedicated neuromorphic chip

Embedded ⟷ Datacenter

Bare Loihi Chip   Kapoho Bay   Oheo Gulch   Nahuku   Pohoiki Springs

Source: https://intel-ncl.atlassian.net/wiki/spaces/INRC/pages/2752539/Loihi+Systems+Overview

SNNs supervised learning - Nengo

LIGO Supervised GW-Detection Nengo-DL

Legend:
- ReLU (auc = 0.97)
- SpikingReLU (auc = 0.97)
- LoihiSpikingReLU (auc = 0.92)
- Scaled LoihiSpikingReLU (auc = 0.50)
- Trained LoihiSpikingReLU (auc = 0.67)

```
Layer (type)              Output Shape            Param #
=================================================================
input (InputLayer)        [(None, 2048)]          0

reshape (Reshape)         (None, 2048, 1, 1)      0

conv2d (Conv2D)           (None, 2045, 1, 16)     64        ──→ Off-chip

conv2d_1 (Conv2D)         (None, 511, 1, 16)      1024      ⎤

conv2d_2 (Conv2D)         (None, 127, 1, 32)      2048      ⎥

conv2d_3 (Conv2D)         (None, 31, 1, 64)       8192      ⎥

conv2d_4 (Conv2D)         (None, 6, 1, 128)       65536     ⎥ ON-chip

flatten (Flatten)         (None, 768)             0         ⎥

dense (Dense)             (None, 128)             98304     ⎥

dense_1 (Dense)           (None, 64)              8192      ⎦

output (Dense)            (None, 2)               130
=================================================================
Total params: 183,490
Trainable params: 183,490
Non-trainable params: 0
```
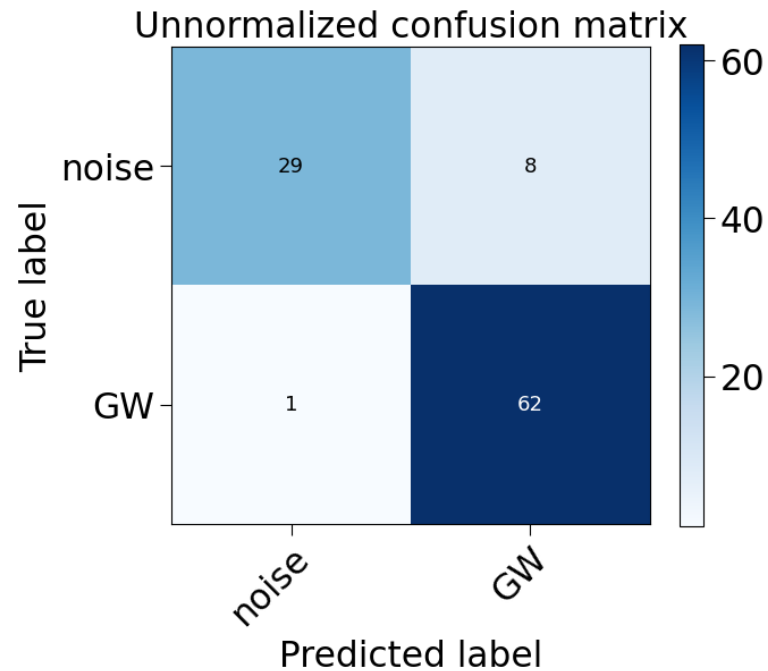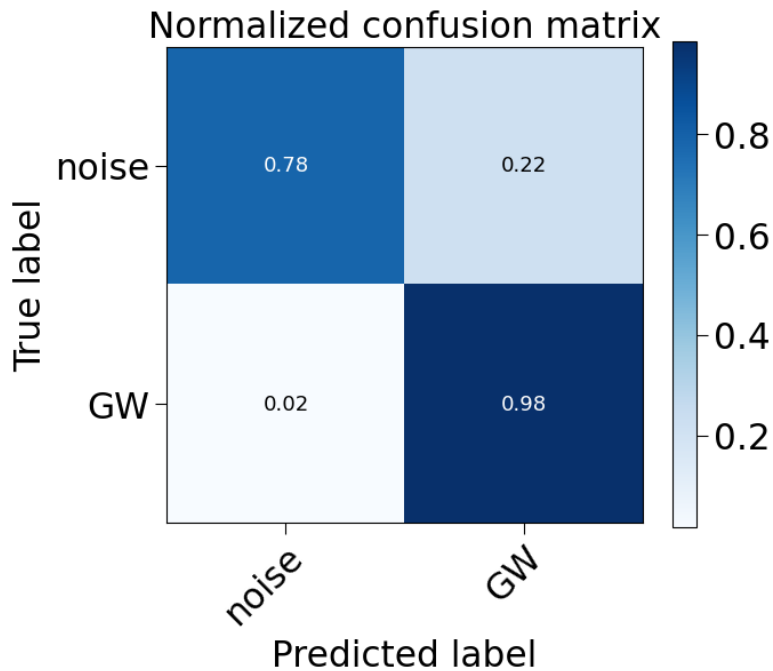
Executed on Loihi chip

SNNs supervised learning - SNN-TB

Normalized confusion matrix

Unnormalized confusion matrix

Accuracy results in simualtion:
ANN: 94%
SNN: 91%

Executed on Loihi chip

Selected performance benchmarks
(keyword spotting task in Nengo for DNN-SNN conversion)

Average Inference Speed

Loihi provides extremely good scaling vs conventional architectures as network size grows by 50x

Dynamic Energy Cost Per Inference (batchsize = 1)

|	LOIHI	MOVIDIUS	JETSON	CPU	GPU
	1x	5.3x	20.5x	23.2x	109.1x
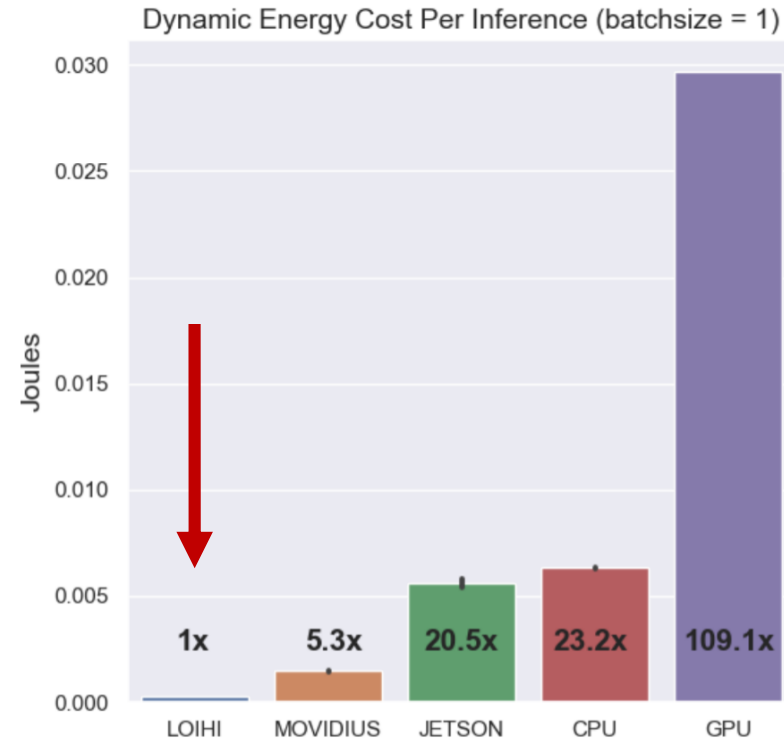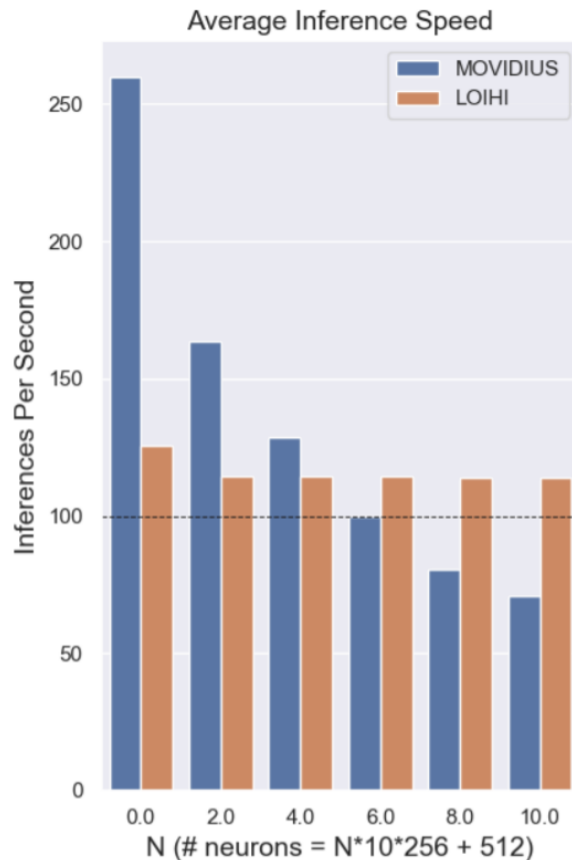
Accuracy results:
90.6% SNN
92.7% DNN

Loihi is the most energy-efficient architecture for real-time inference (batchsize=1 case)

Future opportunities for SNNs

- Unsupervised learning
  (Recurrent SNNs: LSTMs, LMUs)

- Backpropagation algorithm for direct training (SLAYER)

- Evolutionary algorithims to adjust the parameters of spikes
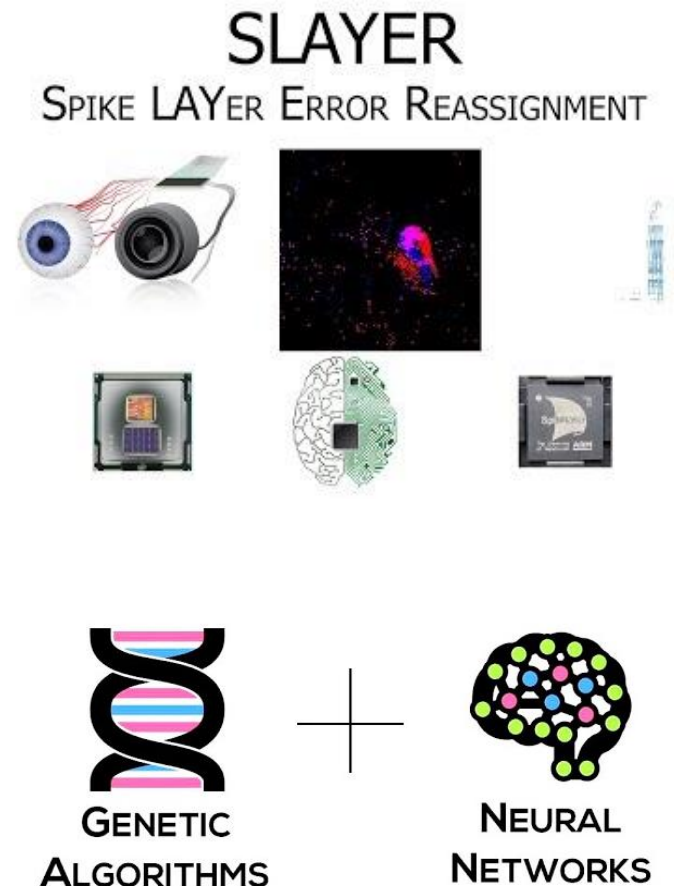
Thank you for your attention!

Do you have any questions?

bartek_borzyszkowski@wp.pl

emoreno@caltech.edu

github.com/eric-moreno/Anomaly-Detection-Autoencoder