# Neuromorphic Computing: An Energy Efficient AI Platform for Edge Computing

Catherine Schuman

Research Scientist

Oak Ridge National Lab

Chesapeake Large-Scale Analytics Conference (CLSAC) 2021

# Why Custom Hardware for Edge-Deployed AI?

Smart Factories

Smart Traffic Control

Smart Transportation

Smart Health

Smart Grid

Smart Buildings

Smart Homes

Smart Satellites

Smart Scientific Instruments

**Power Constraints**

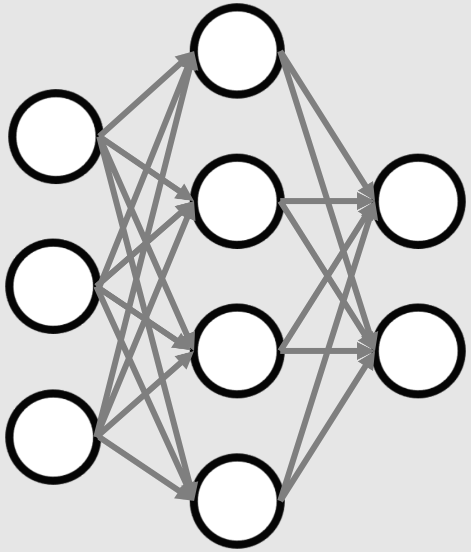
**Time Constraints**

**Size Constraints**

**Privacy and Security**

# Neural Hardware and Neuromorphic Computing

## Neural Hardware

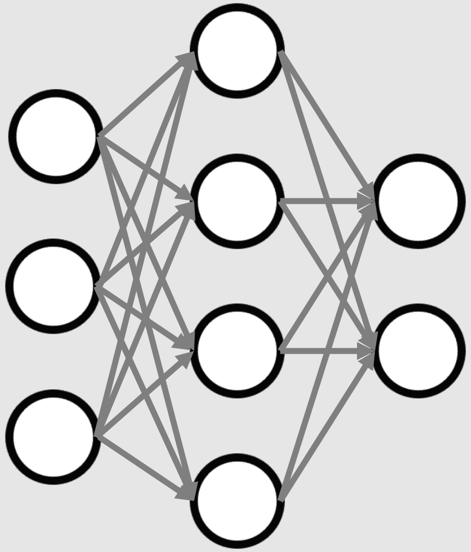


Accelerates traditional neural network and deep learning computation

- Well-suited to existing algorithms
- Fast computation ***or*** low power
- Currently deployed in cloud or mobile devices

# Neural Hardware and Neuromorphic Computing

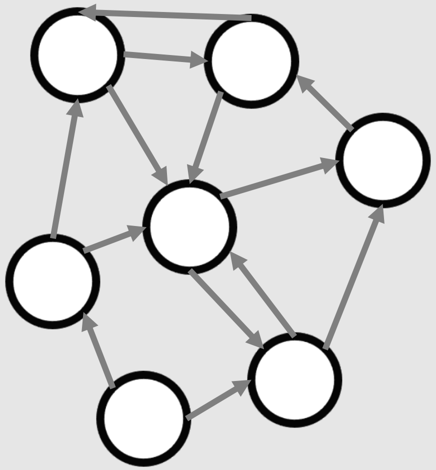
## Neural Hardware



Accelerates traditional neural network and deep learning computation

- Well-suited to existing algorithms
- Fast computation ***or*** low power
- Currently deployed in cloud or mobile devices

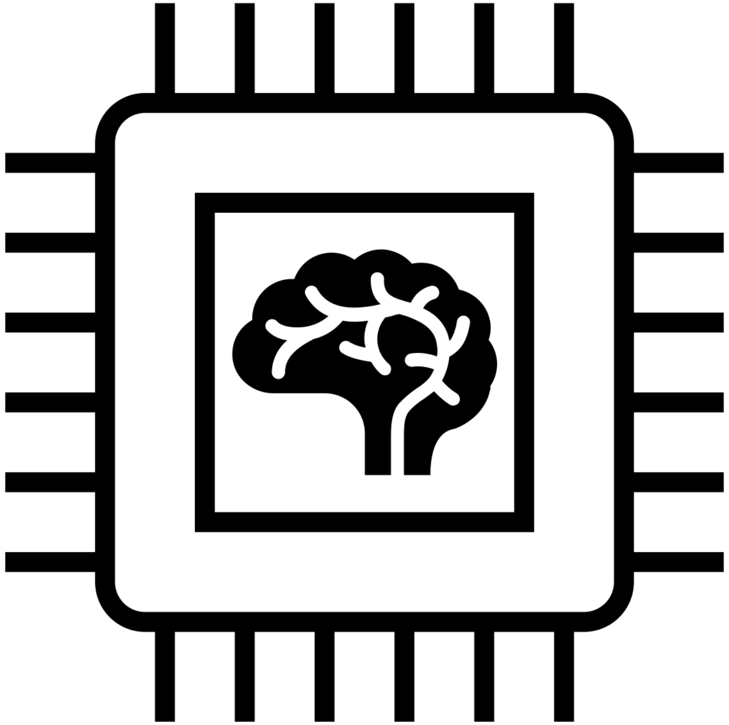
## Neuromorphic Computing



Implements spiking recurrent neural network computation and can be suitable for neuroscience simulation

- Significant promise for future algorithmic development
- Fast computation ***and*** low power
- Still in development

# Neuromorphic Computing

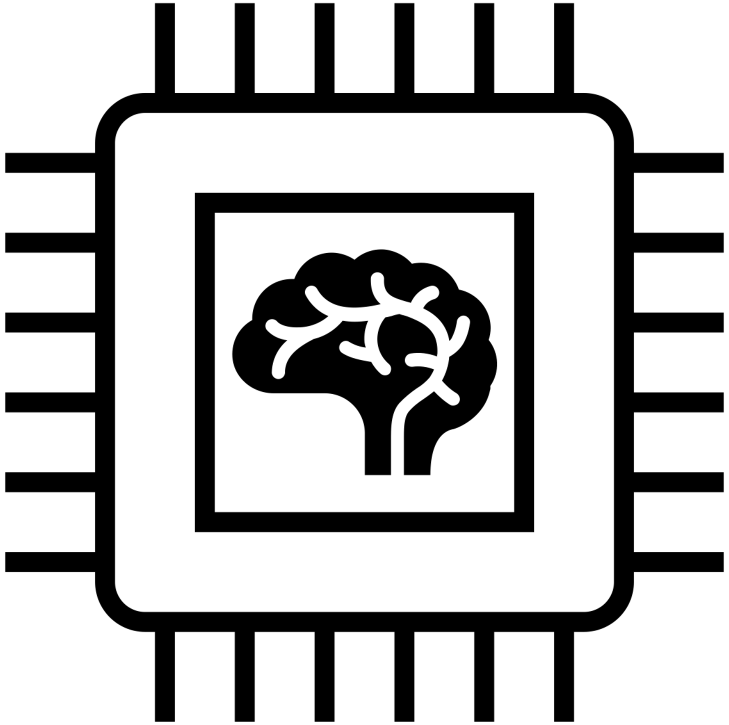


# Neuromorphic Computing

Extremely Low Power

Massively Parallel

Event Driven



Collocated Processing and Memory

Robust and Fault Tolerant

Fast

Adaptability and Plasticity

Training and Learning

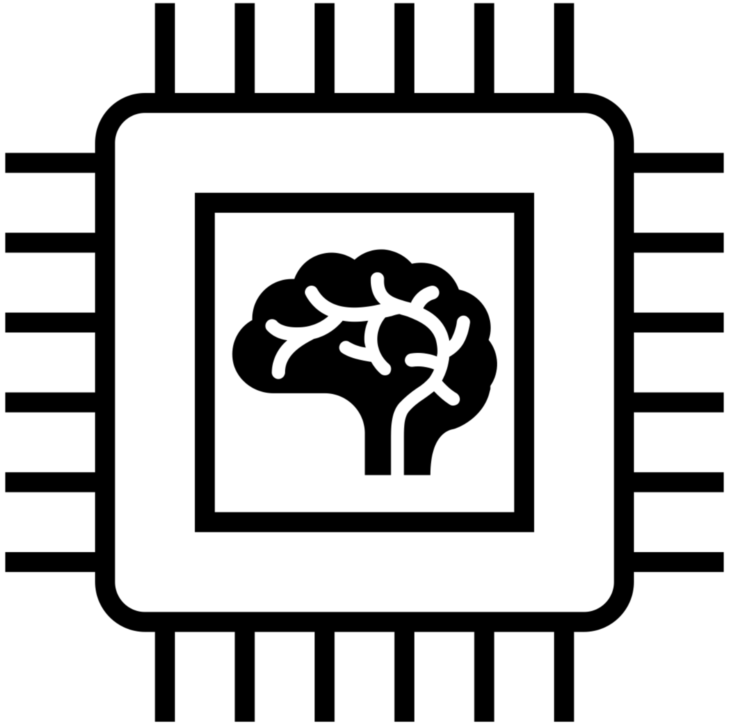
Ability to Generalize

# Neuromorphic Computing

Extremely Low Power

Massively Parallel

Event Driven



Collocated Processing and Memory

Robust and Fault Tolerant

Fast

Adaptability and Plasticity

Training and Learning

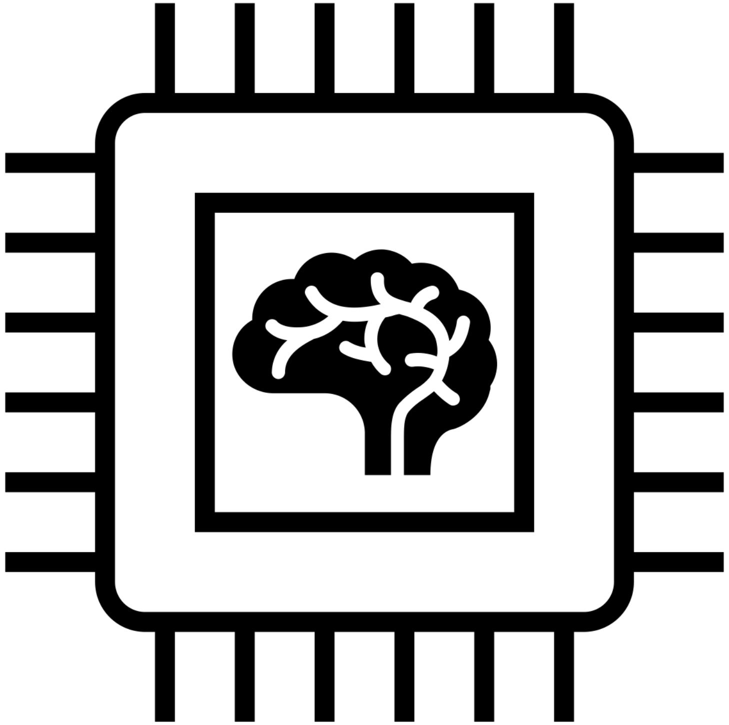
Ability to Generalize

# Neuromorphic Computing

Extremely Low Power

Massively Parallel

Event Driven



Collocated Processing and Memory

Robust and Fault Tolerant

Fast

Adaptability and Plasticity

Training and Learning

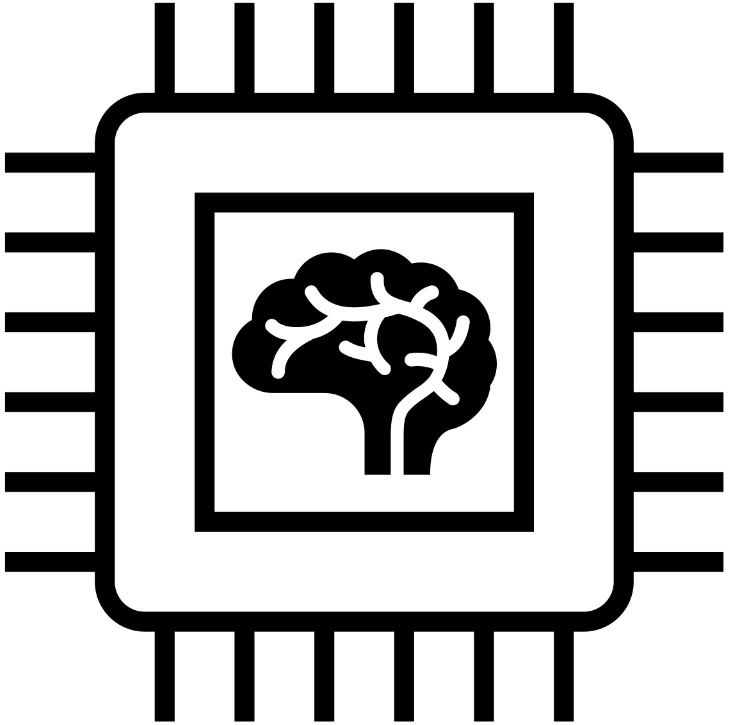
Ability to Generalize

# Neuromorphic Computing

Extremely Low Power

Massively Parallel

Event Driven



Collocated Processing and Memory

Robust and Fault Tolerant

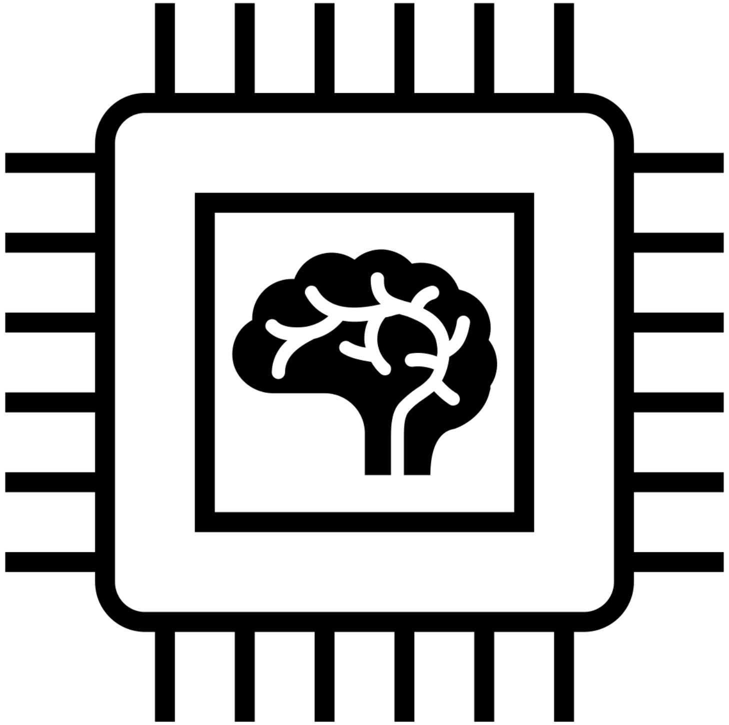
Fast

Adaptability and Plasticity

Training and Learning

Ability to Generalize

# Neuromorphic Computing



Extremely Low Power

Massively Parallel

Event Driven

Collocated Processing and Memory

Robust and Fault Tolerant

Fast

Adaptability and Plasticity

Training and Learning

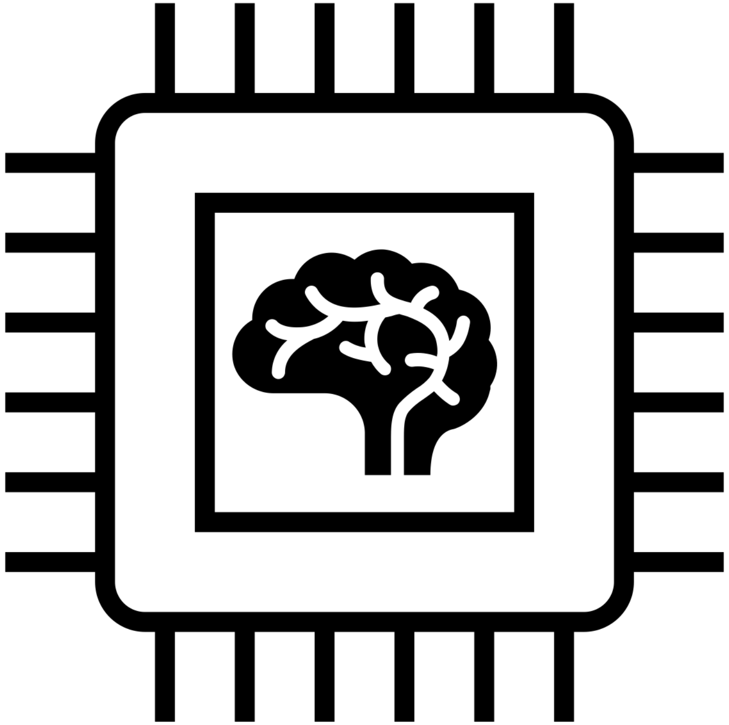
Ability to Generalize

# Neuromorphic Computing

Extremely Low Power

Massively Parallel

Event Driven



Collocated Processing and Memory

Robust and Fault Tolerant

Fast

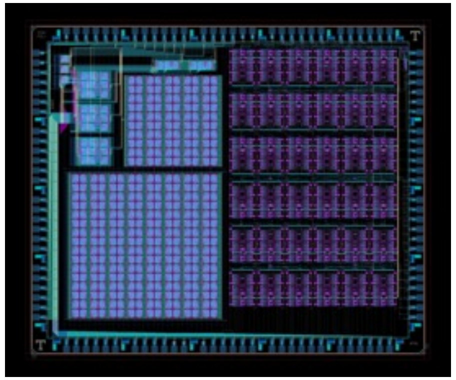
Adaptability and Plasticity

Training and Learning

Ability to Generalize

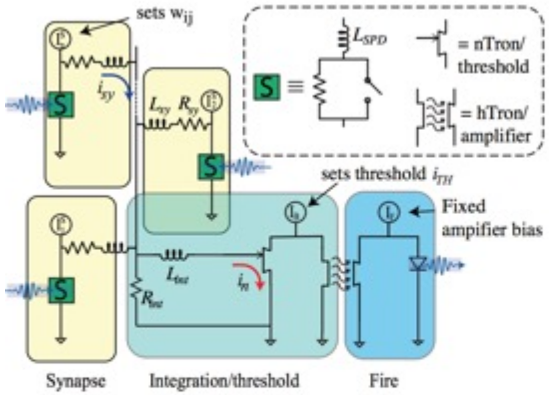
# Neuromorphic Hardware at the Edge is NOT One-Size-Fits-All

## Mixed Analog Digital Memristive Neuromorphic



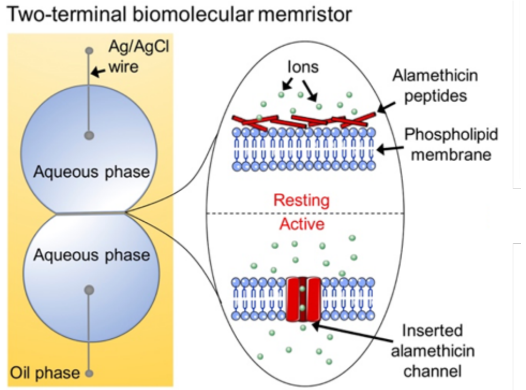
Chakma, Gangotree, et al. "Memristive mixed-signal neuromorphic systems: Energy-efficient learning at the circuit-level." *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 8.1 (2018): 125-136.

## Superconducting Optoelectronic Neuromorphic

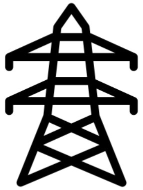



Buckley, Sonia, et al. "Design of superconducting optoelectronic networks for neuromorphic computing." *2018 IEEE International Conference on Rebooting Computing (ICRC)*. IEEE, 2018.

## Biomimetic Neuromorphic

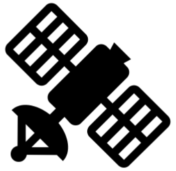



Najem, Joseph S., et al. "Memristive ion channel-doped biomembranes as synaptic mimics." *ACS nano* 12.5 (2018): 4702-4711.



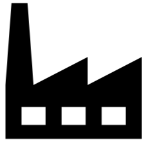
Smart Grid

Smart Transportation



Smart Satellites

Smart Scientific Instruments

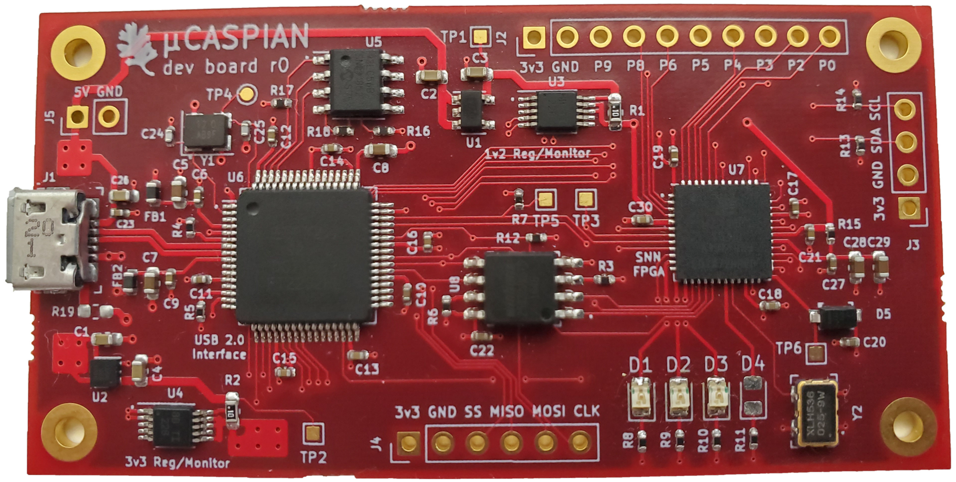


Smart Factories

Smart Health

# µCaspian: A Development Board for SNNs at the Edge

- Edge deployable neuromorphic system
- 256 neurons, 4096 synapses
- Integrated power monitoring for quantifying network efficiency
- FPGA used to implement the µCaspian architecture is the Lattice iCE40 UP5K1.
- Hardware is runtime reconfigurable for different networks and applications
- Development board has both USB and direct I/O interfaces
- Hardware accurate software simulator in C++




Power:

- Using the USB interface: 500 mW
- Without USB interface: 10-20 mW

Size and weight:

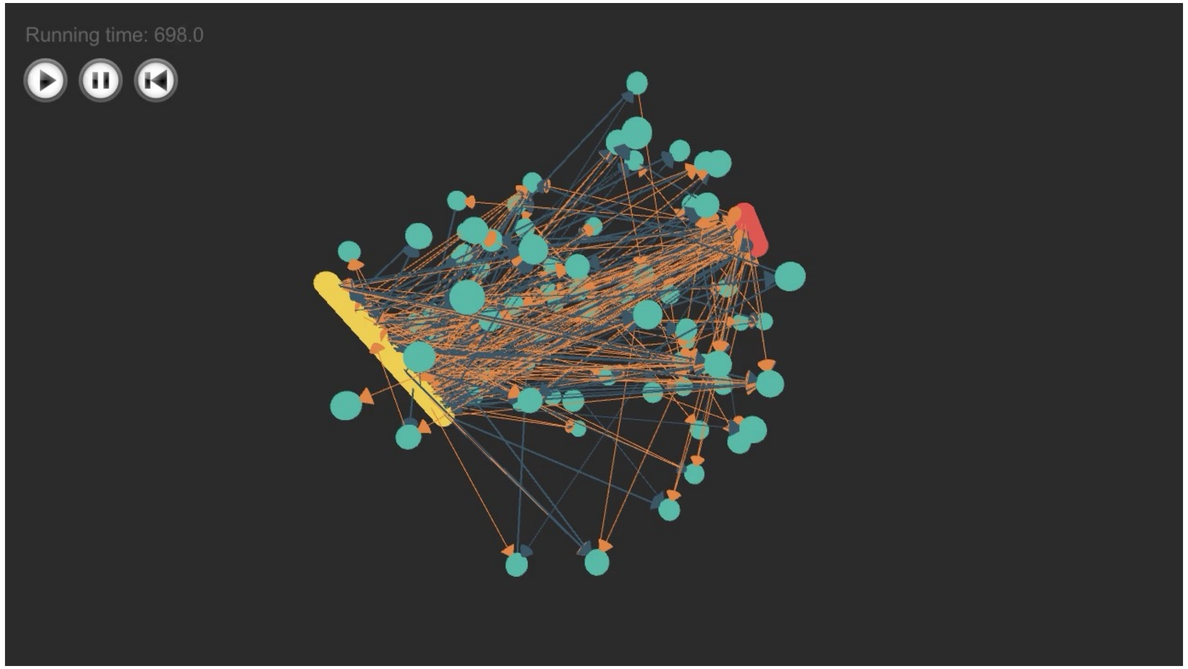
- The iCE40 UP5K is available in a 2.15×2.55 mm packaging, with a weight of approximately 0.0505 oz.

# How do you "program" a neuromorphic computer?

- Neuromorphic computers are programmed with spiking neural networks

- Spiking neural networks take additional inspiration from biological brains by including time as a fundamental part of computation

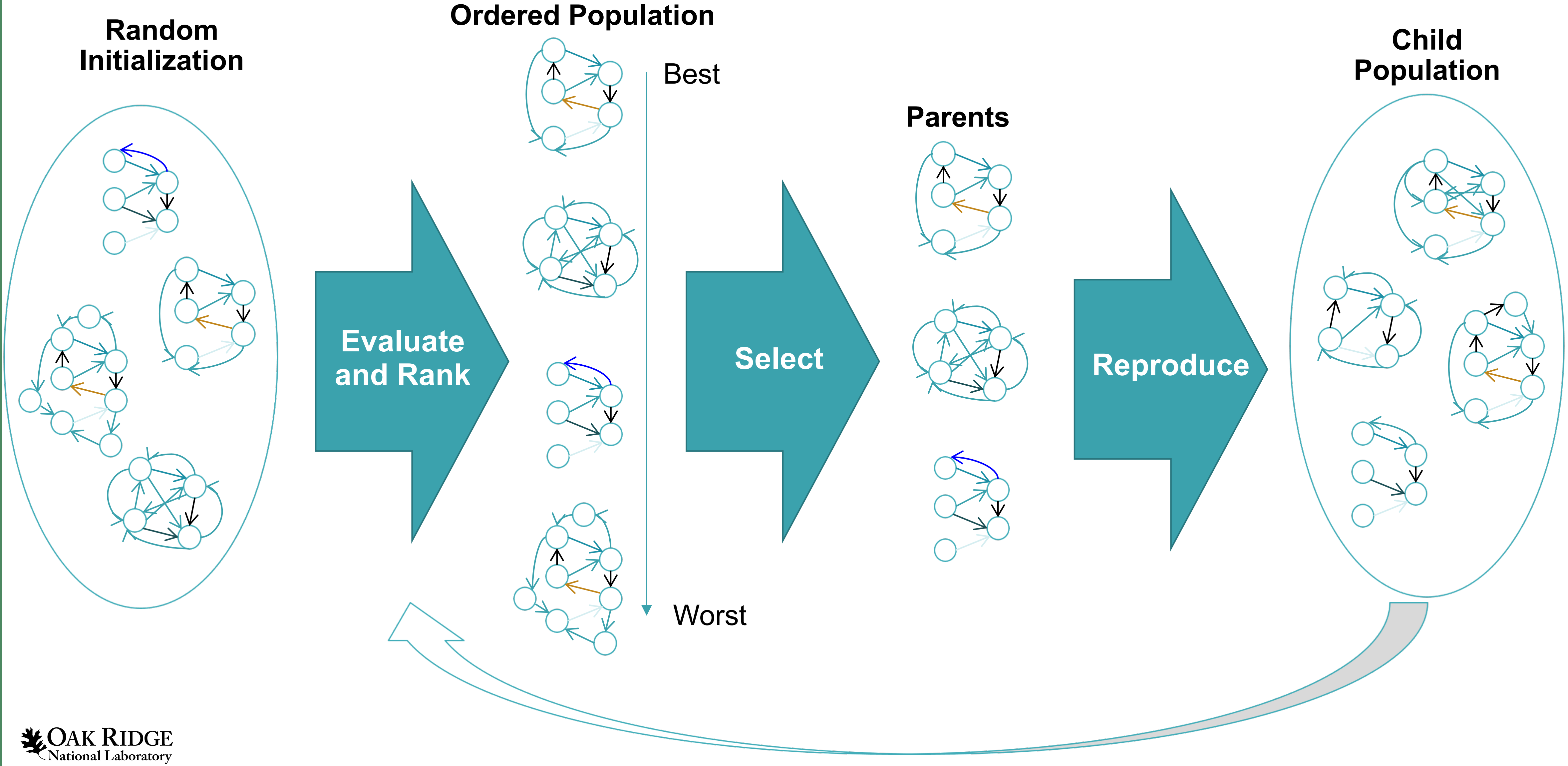
- Spiking neural networks have a lot of computational power, but unlike traditional neural networks, there is not one "best" training or learning approach





Oak Ridge National Laboratory

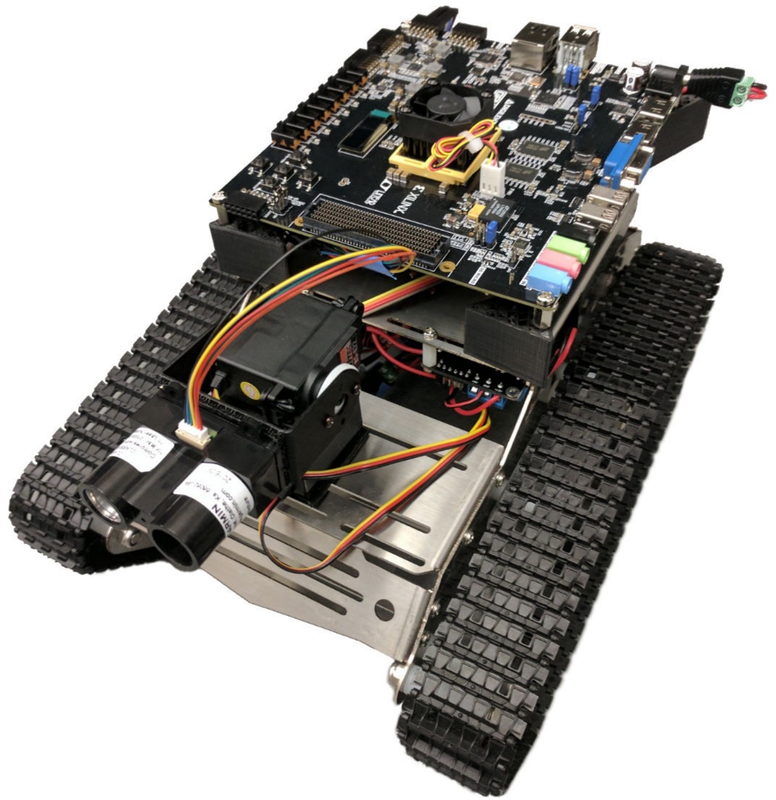
# Evolutionary Optimization for Neuromorphic Systems (EONS)



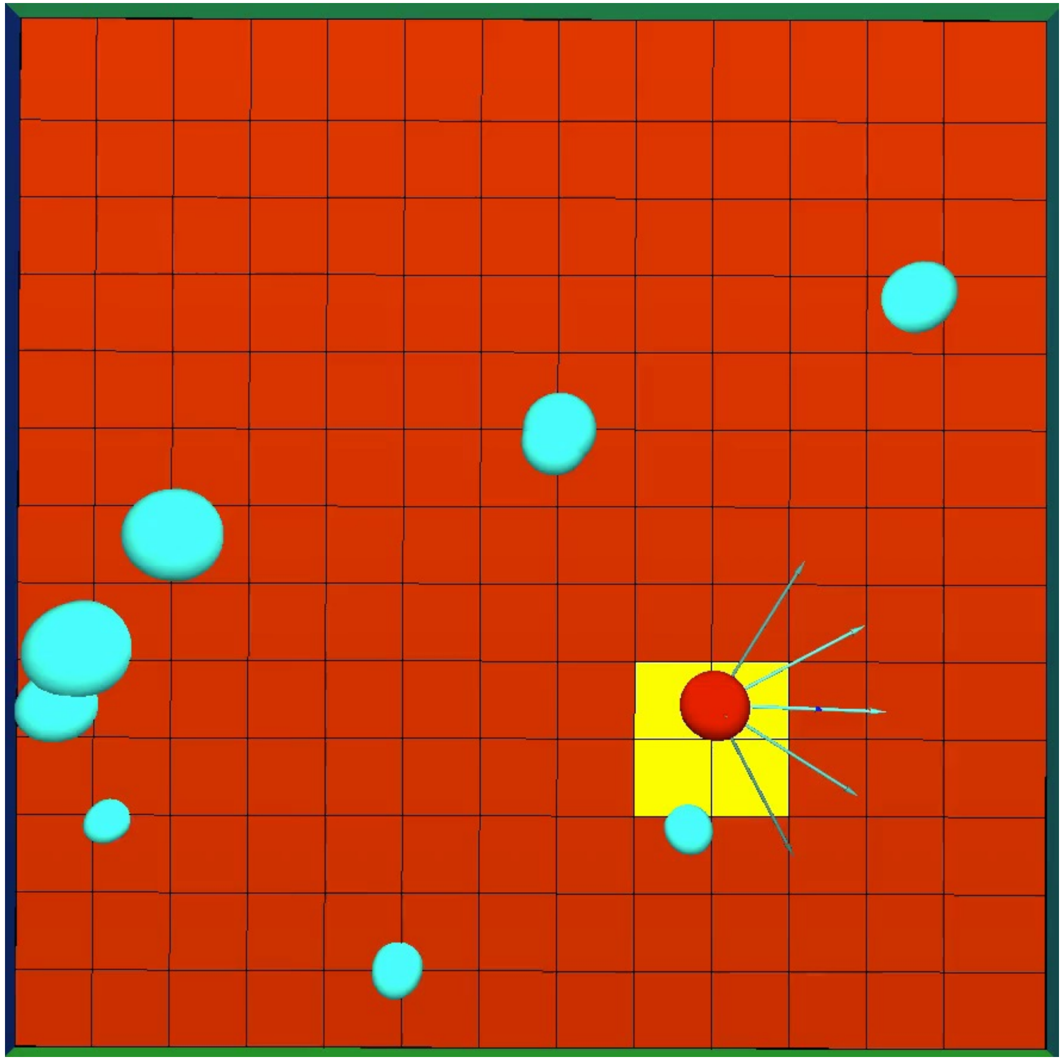
# Applications at the Edge

# Example Application: Autonomous Robot Navigation

- Task: Navigate and explore an unfamiliar environment while avoiding obstacles
- Challenges:
  - Process all inputs and make control decisions on-board the robot (no communication to/from the robot to another computer system)
  - Limited input resolution (LIDAR sensors)
  - No explicit instructions on how to operate
  - No prior knowledge about the environment
  - Train only in simulation

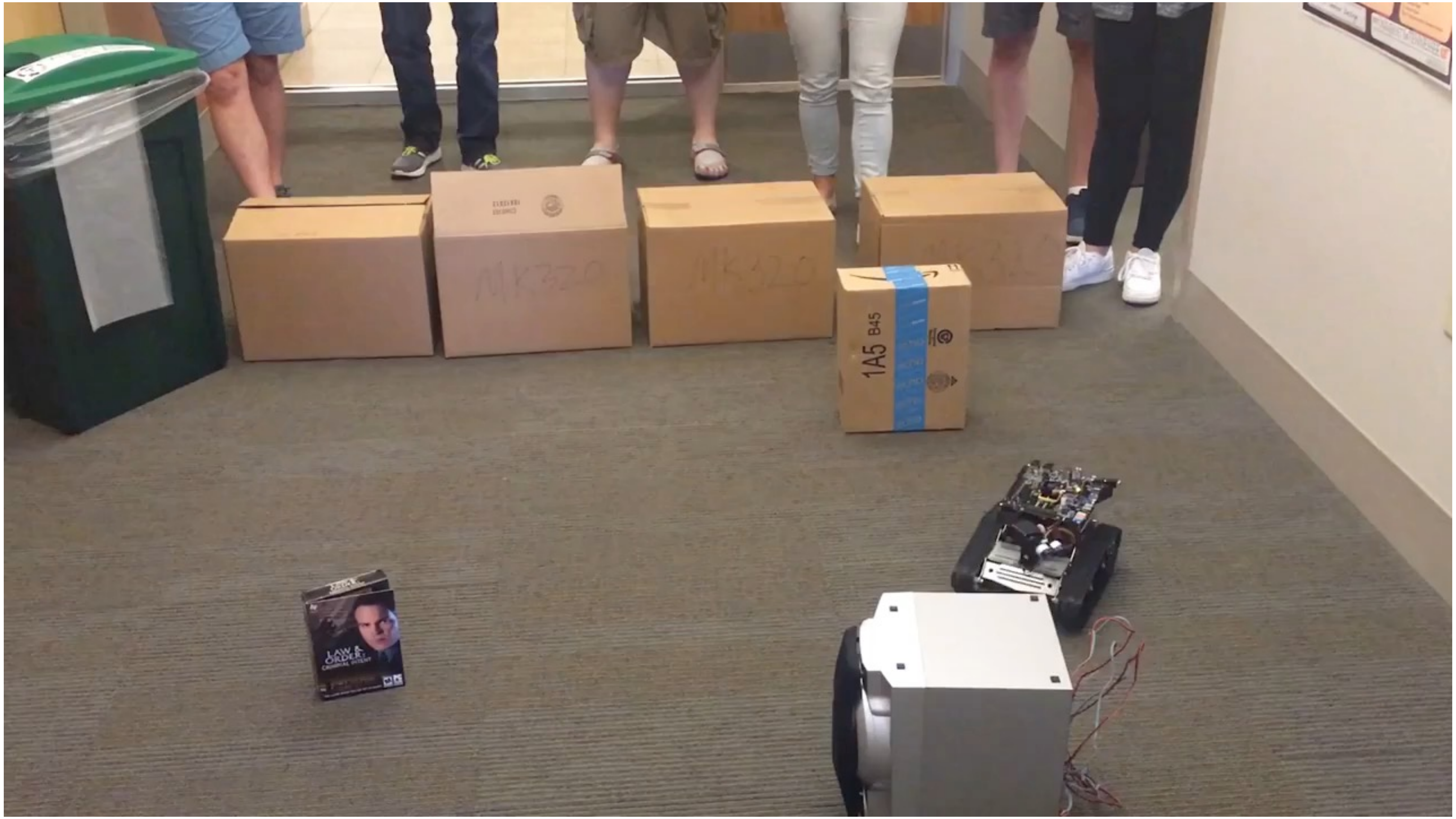


# Application: Robotics Control Results



Student Application: Parker Mitchell and Grant Bruer

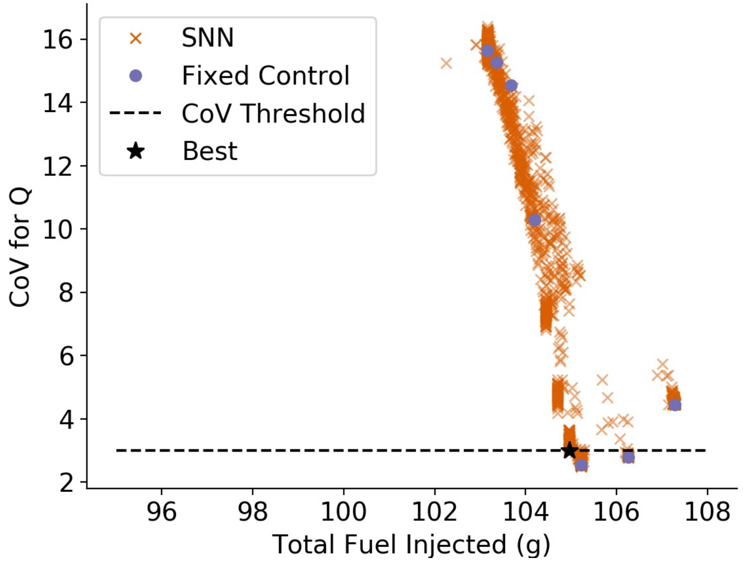
# Application: Robotics Control Results

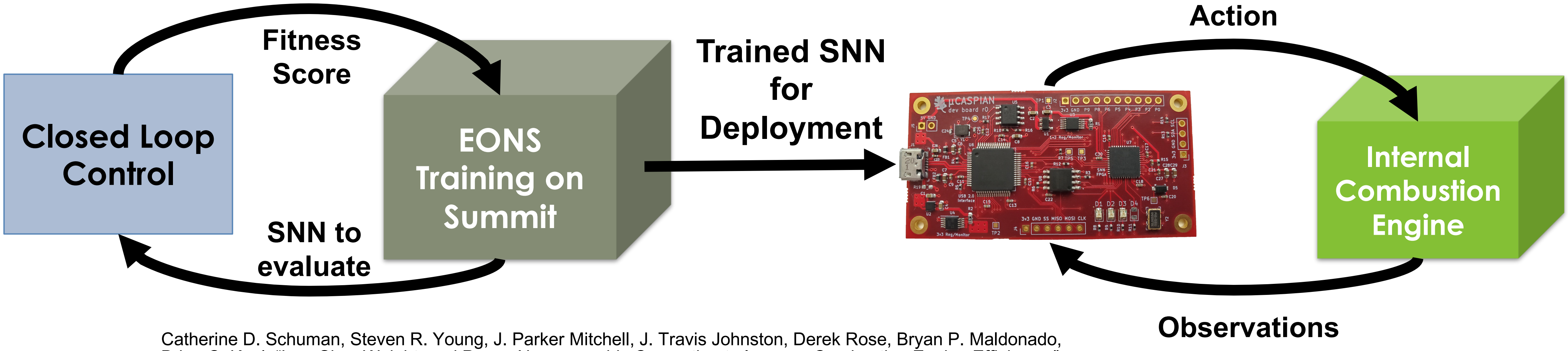



Mitchell, J. Parker, Grant Bruer, Mark E. Dean, James S. Plank, Garrett S. Rose, and Catherine D. Schuman. "NeoN: Neuromorphic control for autonomous robotic navigation." In *2017 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, pp. 136-142. IEEE, 2017.

# Neuromorphic Engine Control for Fuel Efficiency

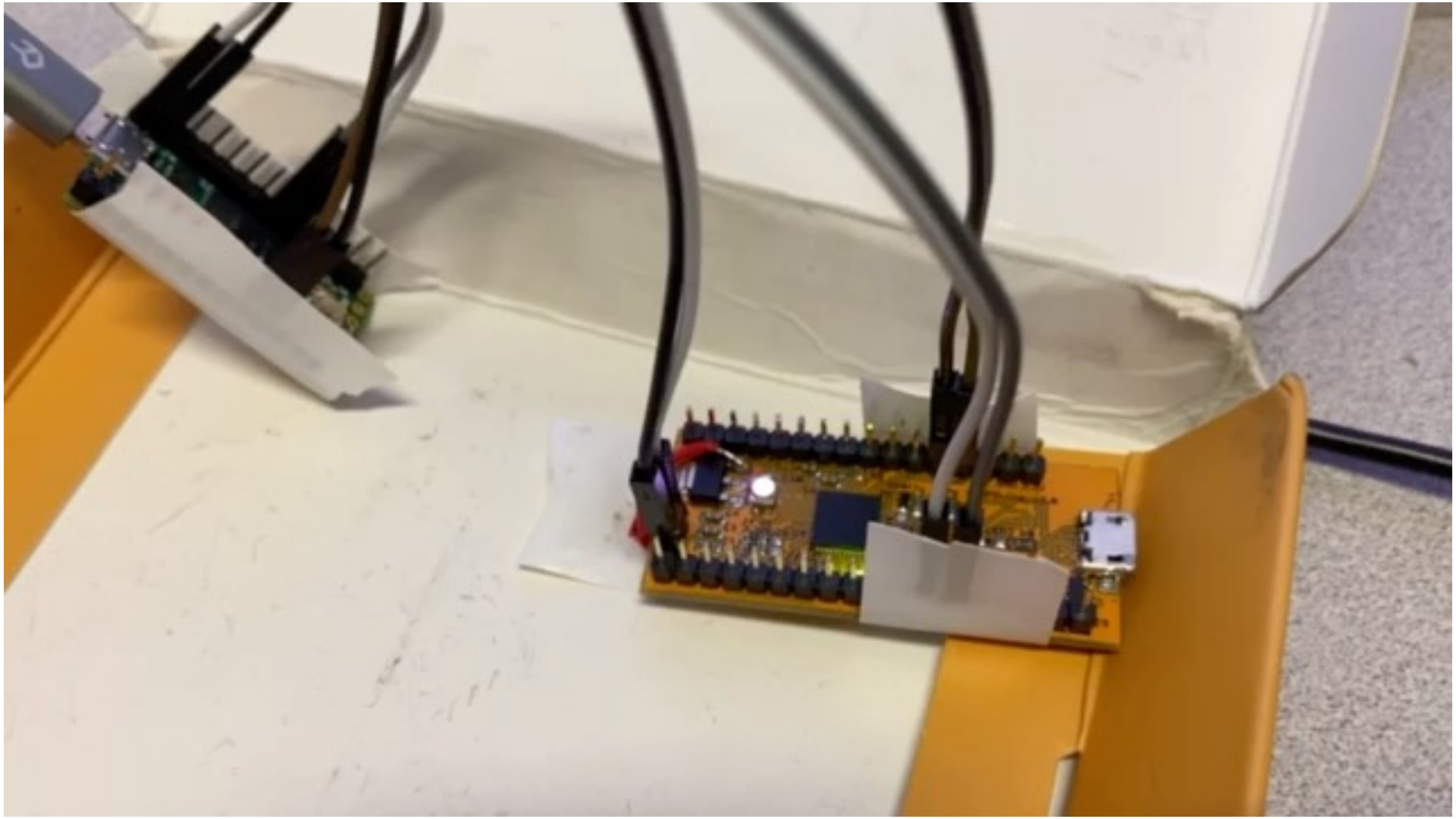
- Developed a complete workflow to train a spiking neural network (SNN) to deploy to an FPGA-based neuromorphic hardware system for internal combustion engine control.
- SNN-based approach outperforms fixed control strategies in terms of fuel efficiency in simulation while still meeting acceptable performance metrics.
- Currently deploying SNN trained on Summit to neuromorphic hardware in-the-loop with engine at National Transportation Research Center.



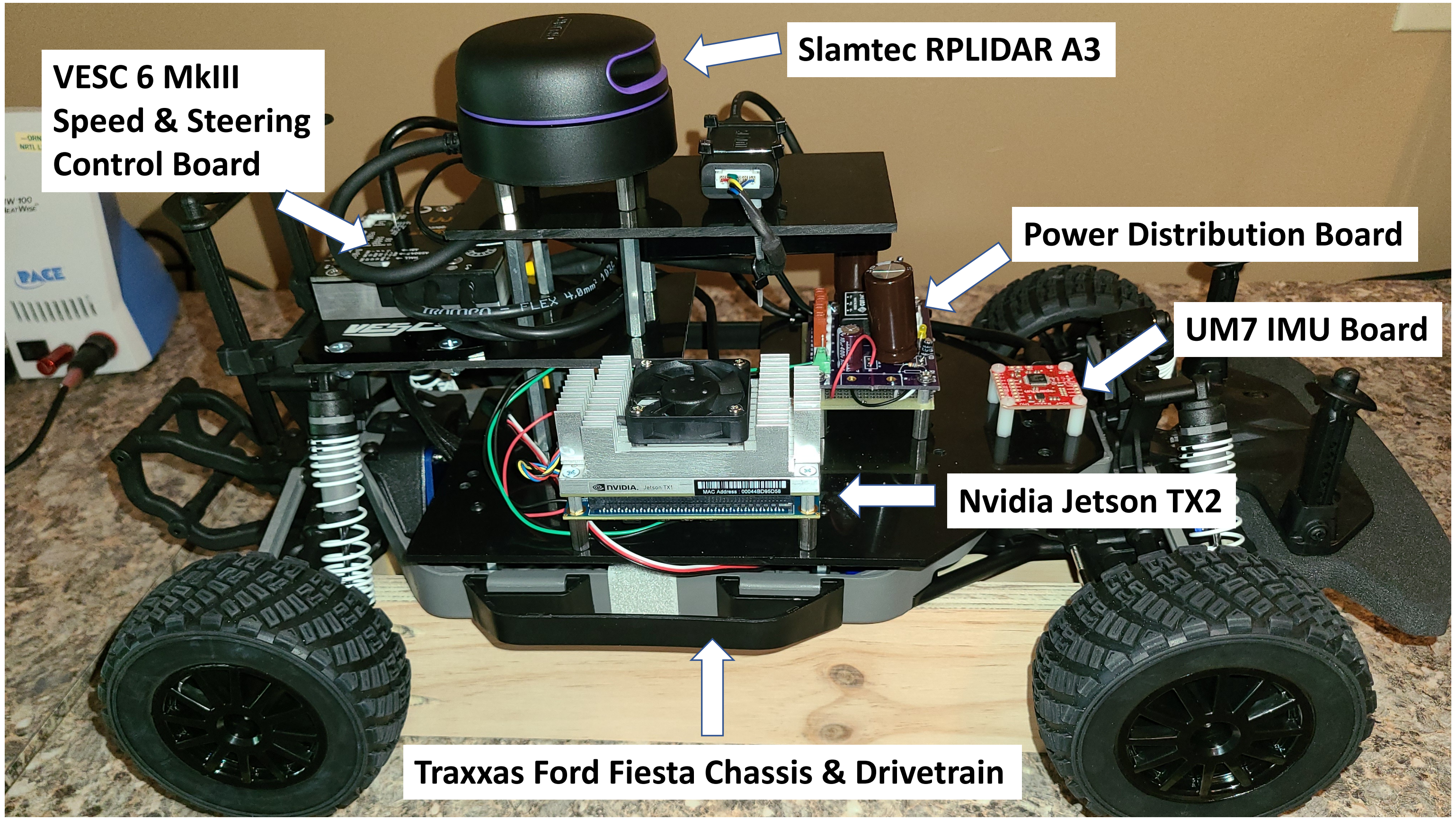

Catherine D. Schuman, Steven R. Young, J. Parker Mitchell, J. Travis Johnston, Derek Rose, Bryan P. Maldonado, Brian C. Kaul. “Low Size, Weight, and Power Neuromorphic Computing to Improve Combustion Engine Efficiency.” International Conference on Green and Sustainable Computing 2020.

# Neuromorphic Engine Control for Fuel Efficiency



# F1Tenth

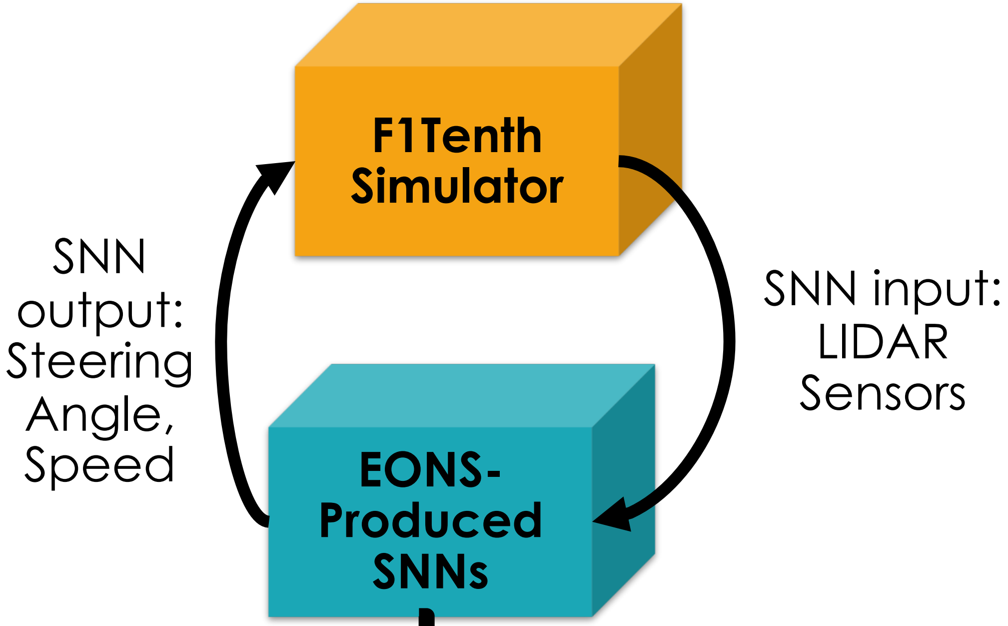
- Fully autonomous 1/10th scale racing of Formula One (https://f1tenth.org/)
- Like full scale vehicles, the need for low size, weight, and power is critical
- Relatively inexpensive real-world demonstration of what neuromorphic computing can provide



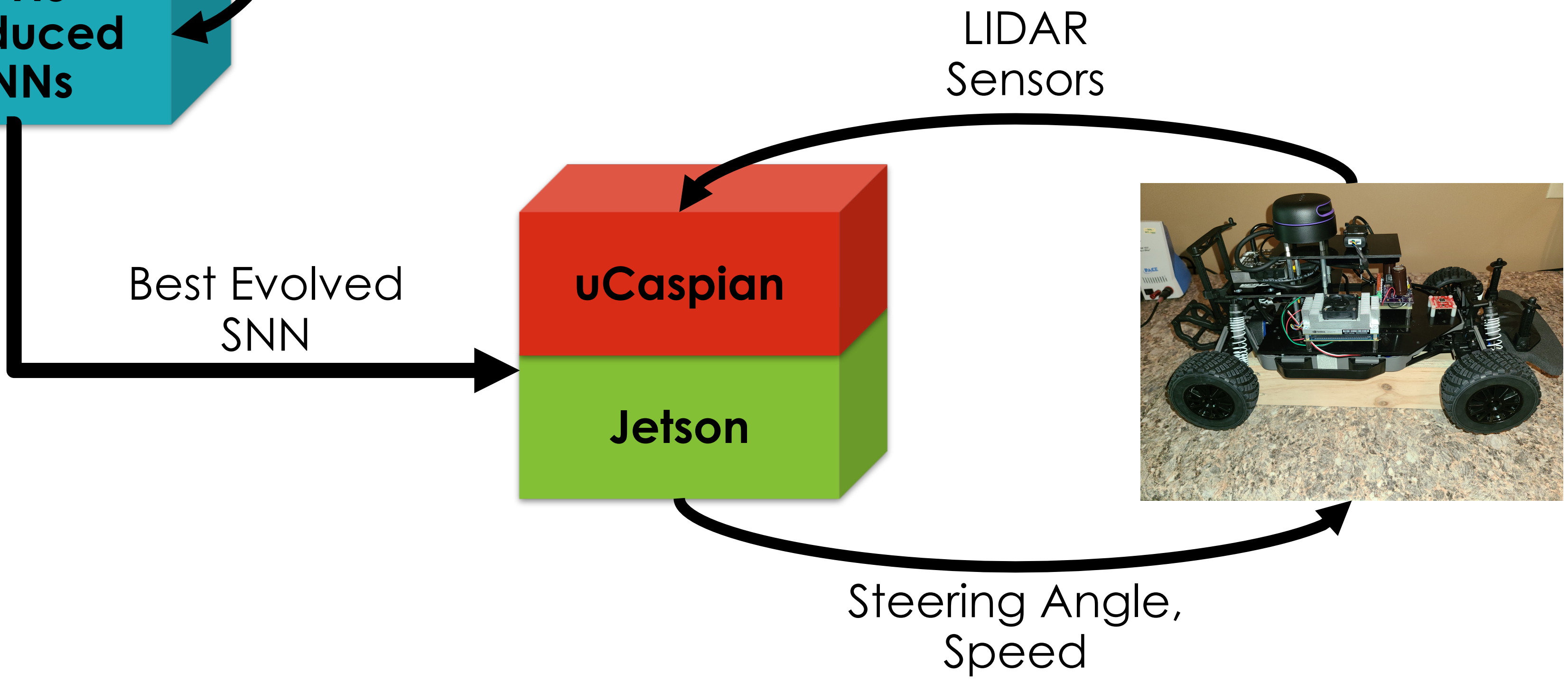


OAK RIDGE National Laboratory

# This Work




- We use EONS to train SNNs for uCaspian deployment using the F1Tenth Simulator.
- We will deploy the trained SNN to a Jetson first (CPU will simulate uCaspian) and then to the actual uCaspian to control the physical car.



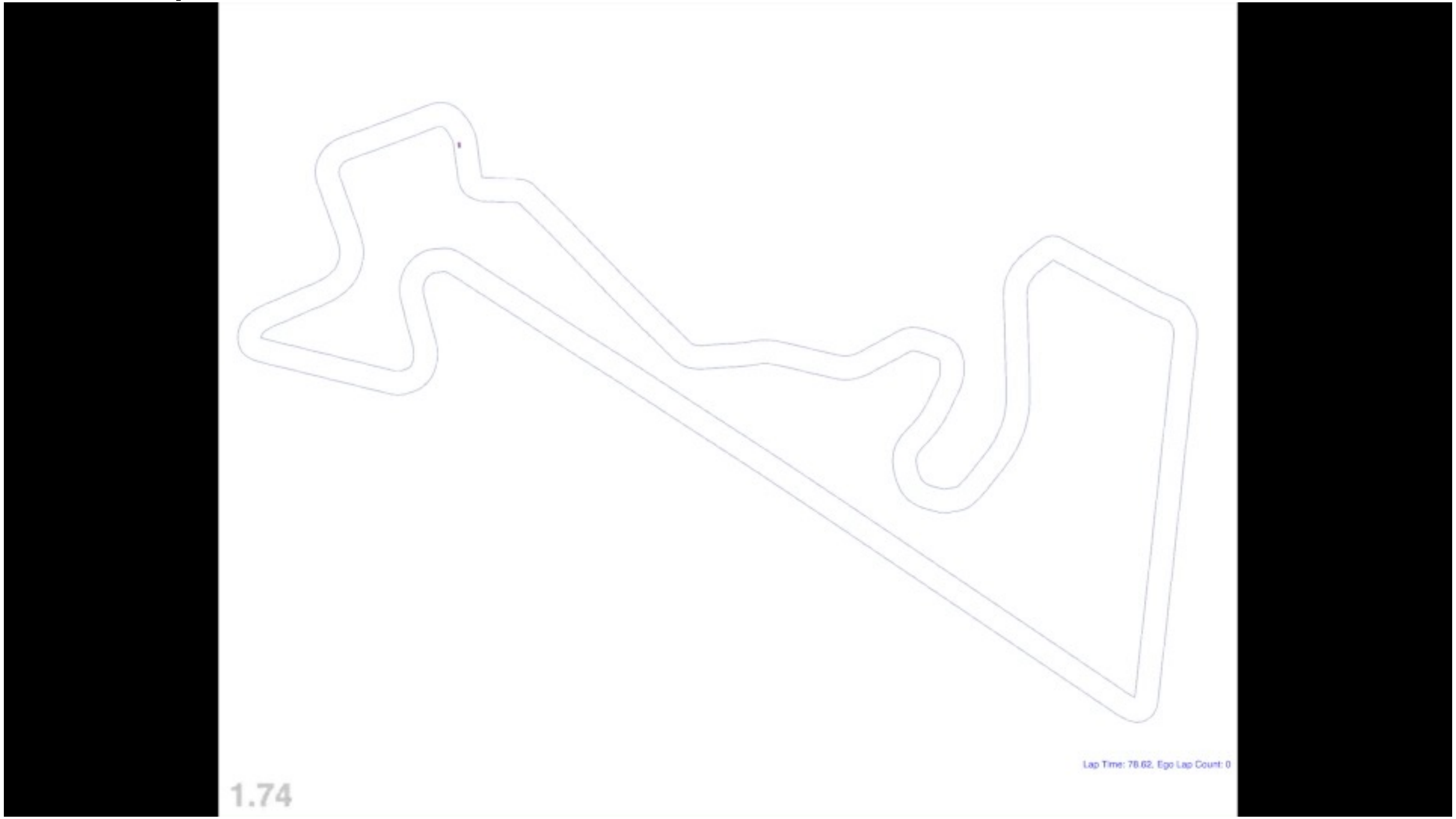
# Training Tracks



Oak Ridge National Laboratory

# Testing Tracks

- It’s (usually) able to extend what it has learned



# Physical Deployment

Robert Patton, Catherine Schuman, Shruti R. Kulkarni, John Mitchell, N. Quentin Haas, Christopher Stahl, Spencer Paulissen, Prasanna Date, Thomas Potok, Shay Snyder and Maryam Parsa, "Neuromorphic Computing for Autonomous Racing." International Conference on Neuromorphic Systems (ICONS) 2021.

# Summary

- Neuromorphic computing systems are promising for the future of AI at the edge.
- We have developed neuromorphic hardware and algorithms to evaluate neuromorphic computing solutions for edge applications.
- We have successfully demonstrated neuromorphic technologies on multiple applications in robotics and transportation.

OAK RIDGE
National Laboratory

TENN LAB
NEUROMORPHIC
ARCHITECTURES. LEARNING. APPLICATIONS.



Work supported by:

Department of Energy

Air Force Research Lab

# Thank you!

# Questions?

**Contact:**
**Email: cdschuman@gmail.com**

**Website: catherineschuman.com**

**Twitter: @cdschuman**