Memristor Crossbar Based Low Power Computing

June 30, 2016

Tarek M. Taha

Electrical and Computer Engineering Department

University of Dayton

Areas of Research

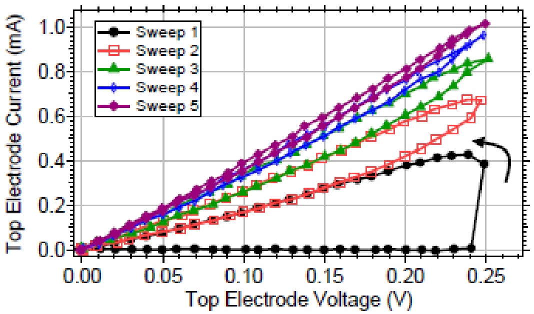
- Application acceleration:
  - Cognitive computing: Autonomous agent for UAVs and decision making
  - Cybersecurity
- Neuromorphic applications:
  - Porting algorithms to IBM TrueNorth, and our internal neuromorphic architectures
  - Examples: cognitive agent, cybersecurity, image processing
- Neuromorphic multicore architectures:
  - Digital CMOS (verified via FPGA implementation)
  - Memristor crossbar
  - Both learning and recognition
  - Specialized versions for: deep learning, cybersecurity, convolution networks, controls
- Memristor devices:
  - SPICE Modeling
  - Fabrication

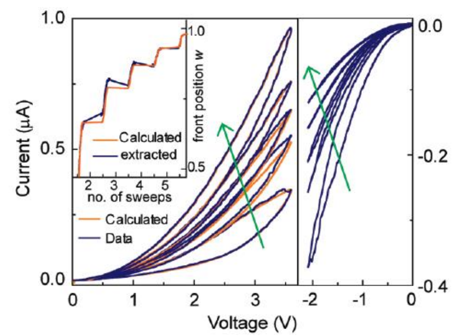
Device Modeling

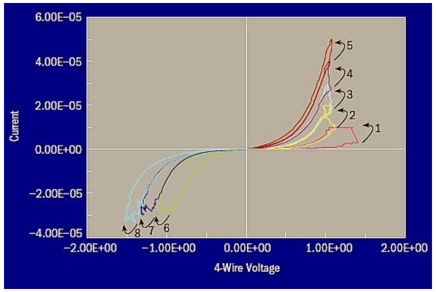
Device SPICE Model

	Boise State	Univ of Michigan	HP Labs

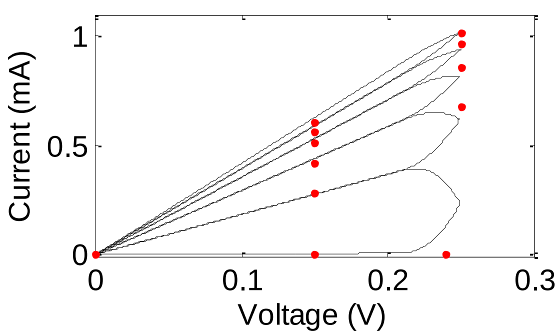
Actual Device

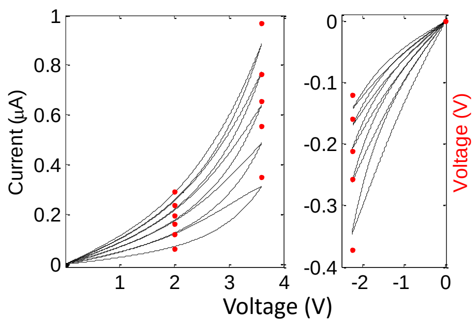


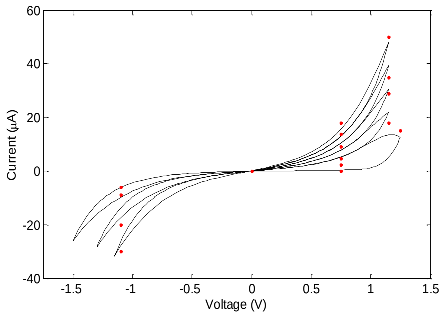




SPICE Model







C. Yakopcic, T. M. Taha, G. Subramanyam, and R. E. Pino, "Memristor SPICE Model and Crossbar Simulation Based on Devices with Nanosecond Switching Time," IEEE International Joint Conference on Neural Networks (IJCNN), August 2013. [**BEST PAPER AWARD**]

Integrated into Sandia XYCE

Wed 1/20/2016 1:02 PM
XYCE <xyce@sandia.gov>
Xyce version 6.4 has been released.

The Xyce (TM) team is pleased to announce the release of Xyce (TM) Version 6.4. This release fixes a number of bugs in Xyce (TM) 6.3 and includes improvements to existing features of Xyce (TM) 6.4. Please see the Release Notes for a complete list of new features and enhancements.

Highlights for Xyce Release 6.4 include:

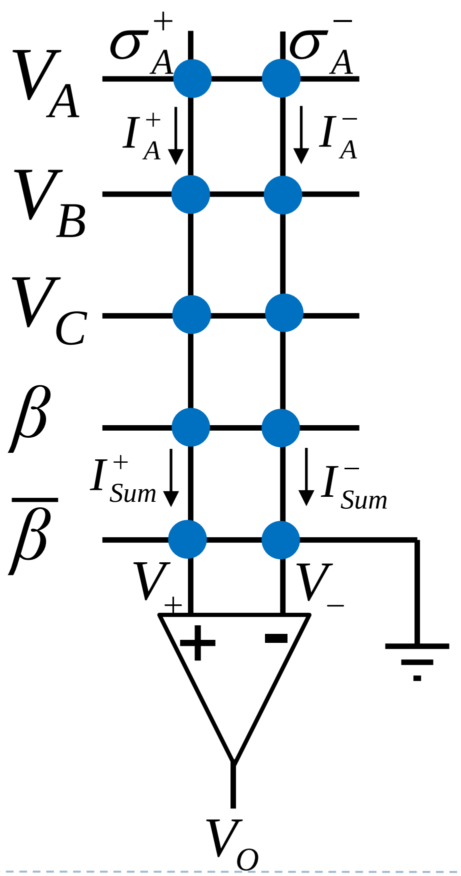
New Devices and Device Model Improvements

* VBIC version 1.3, 3- and 4-terminal variants (Q levels 11 and 12)
* MEXTRAM 504.11 with self-heating (Q level 505)
* New memristor device using the **Yakopcic model**
* Support for Reactive Power limits in the Power Grid Generator Bus model.

Circuit Design

Memristor Based Neuron

- Memristor crossbar emulates multiply-add operation in analog domain.



$$I_A^+ = V_A \sigma_A^+$$

$$I_A^- = V_A \sigma_A^-$$

$$I_{Sum}^+ = V_A \sigma_A^+ + V_B \sigma_B^+ + V_C \sigma_C^+ + \beta \sigma_\beta^+$$

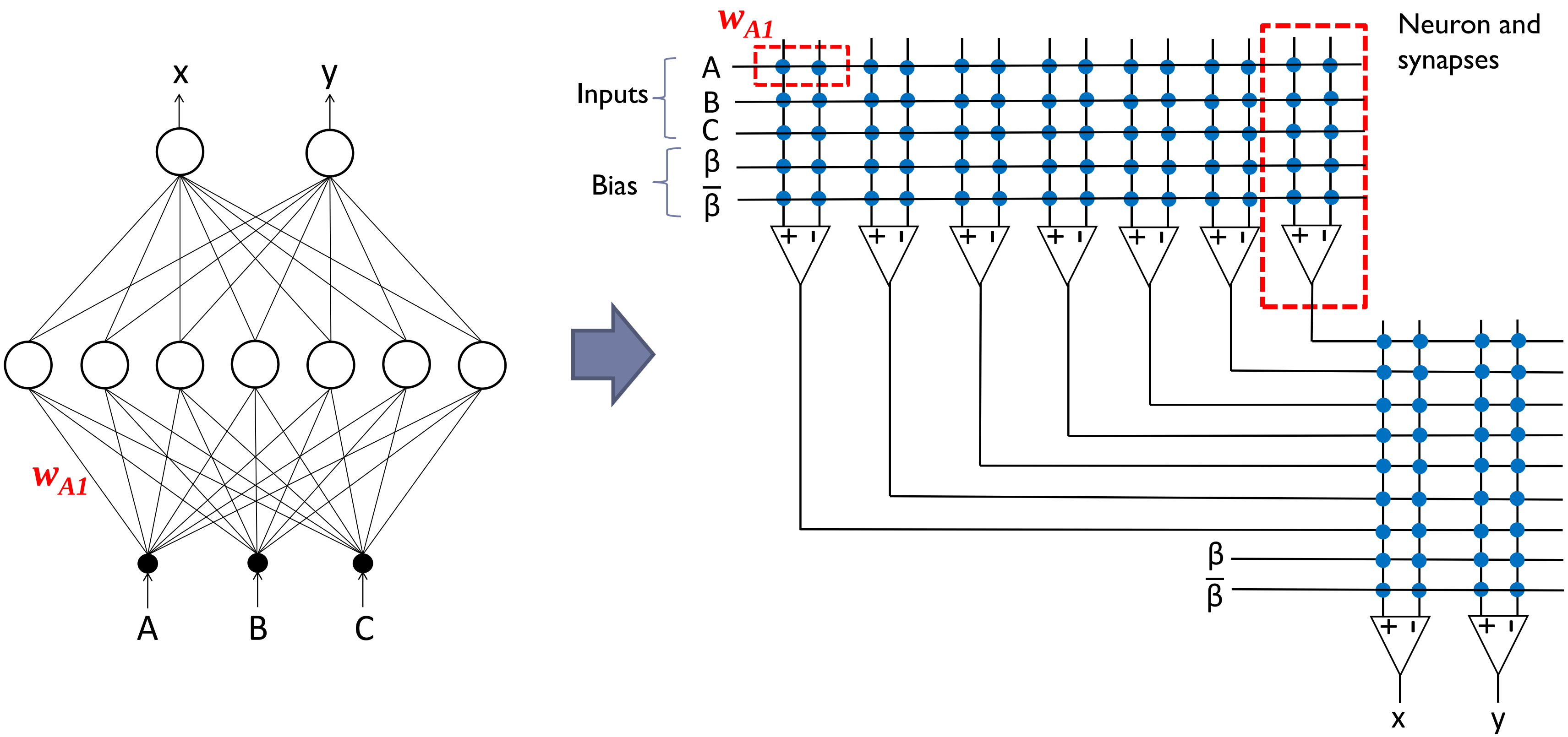
$$V_+ = I_{Sum}^+ \Omega_{\bar{\beta}}^+$$

$$V_{Diff} = V_+ - V_-$$

$$V_O = \begin{cases} 1, V_{Diff} > V_T \\ 0, V_{Diff} < V_T \end{cases}$$

Analog Memristor Classifier

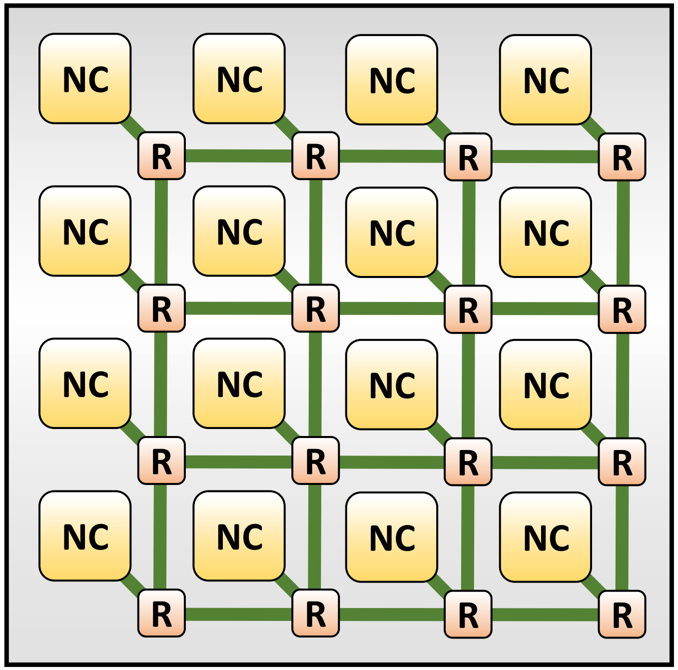
- 2 layer CLA network using analog memristor circuit
  - Based on memristor crossbars
  - Iteratively trained through MATLAB and SPICE
  - Each weight represented by two memristors (for both signs)



System Design

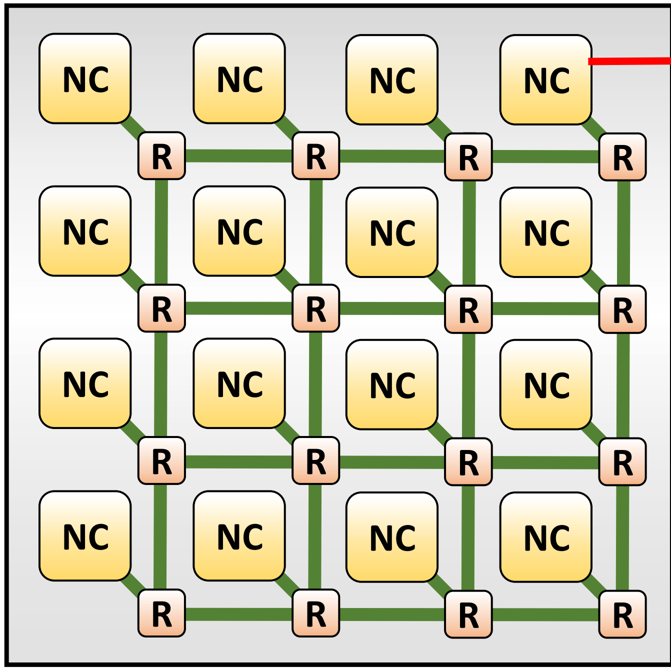
Multicore Neural Processor

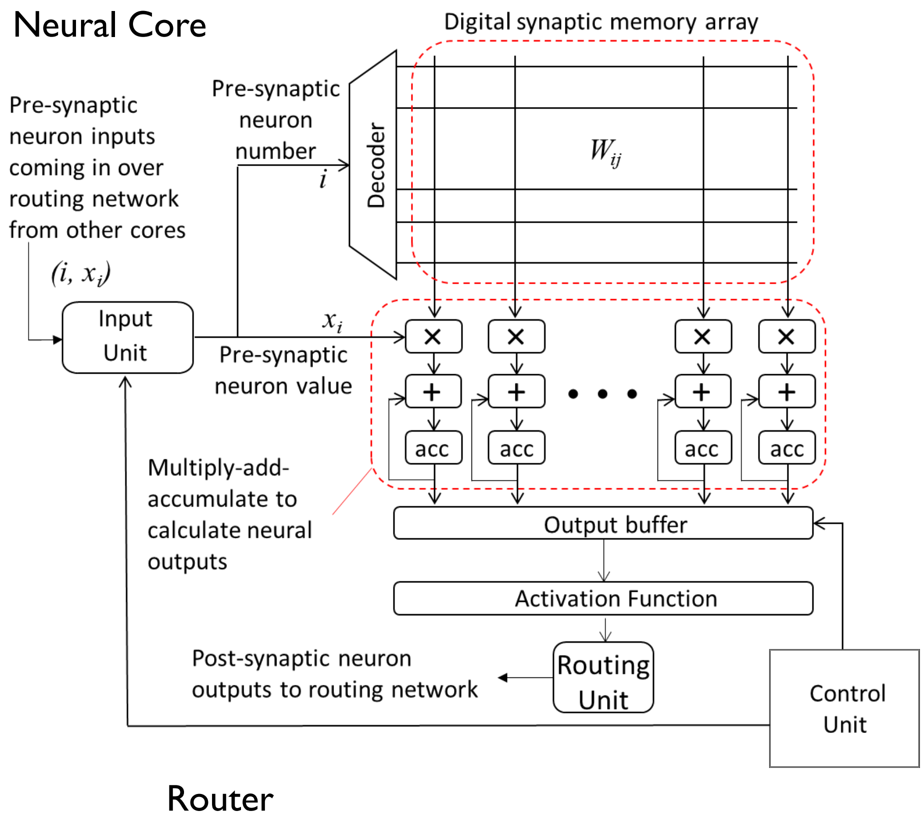
NC = Neural Core
R = Router



Multicore Neural Processor

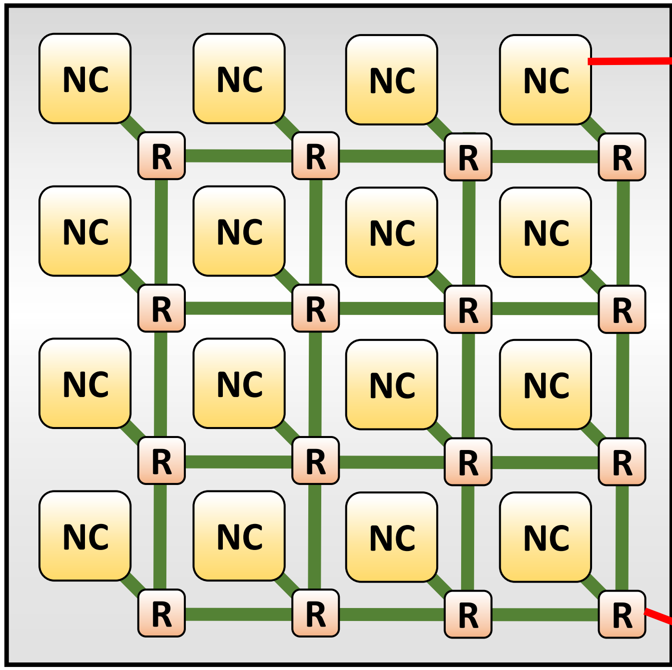
NC = Neural Core
R = Router



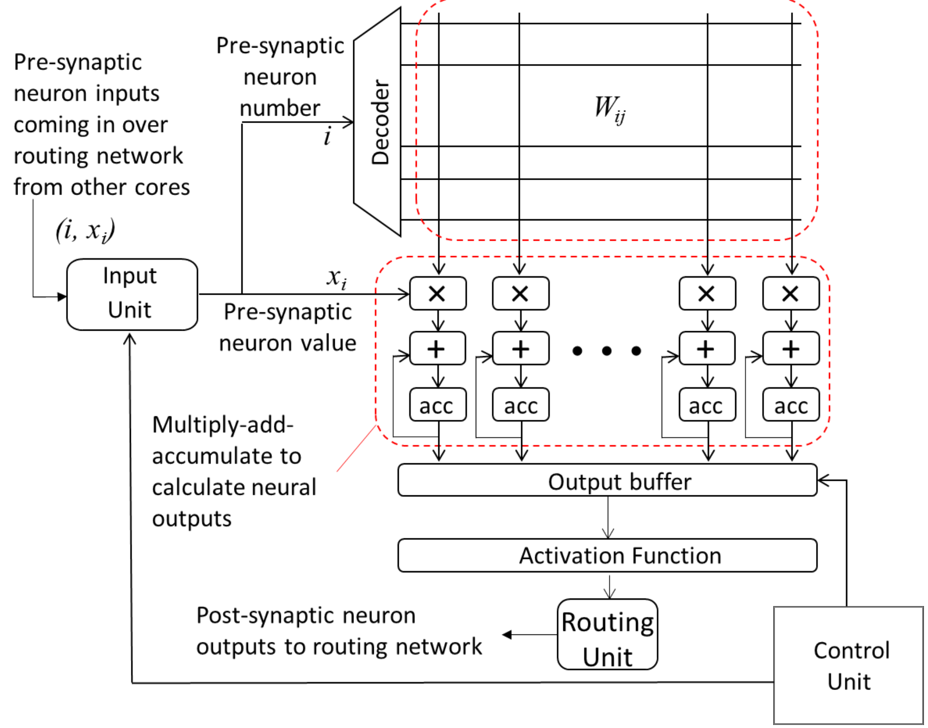


Multicore Neural Processor

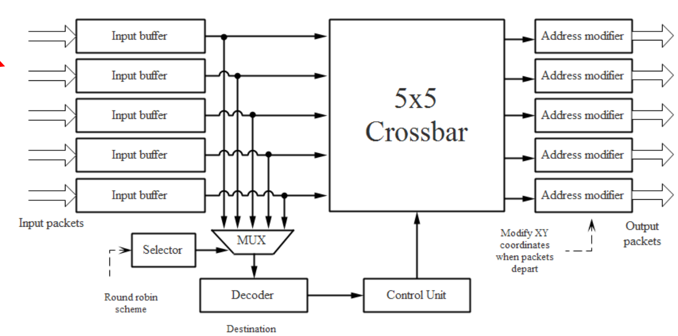
NC = Neural Core
R = Router



Neural Core

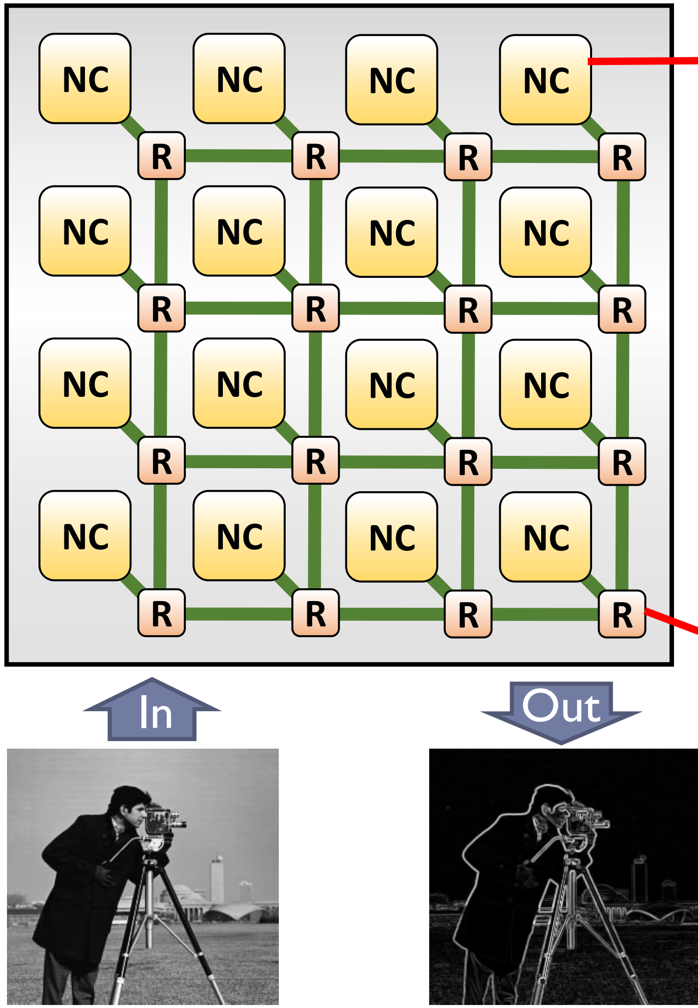


Router

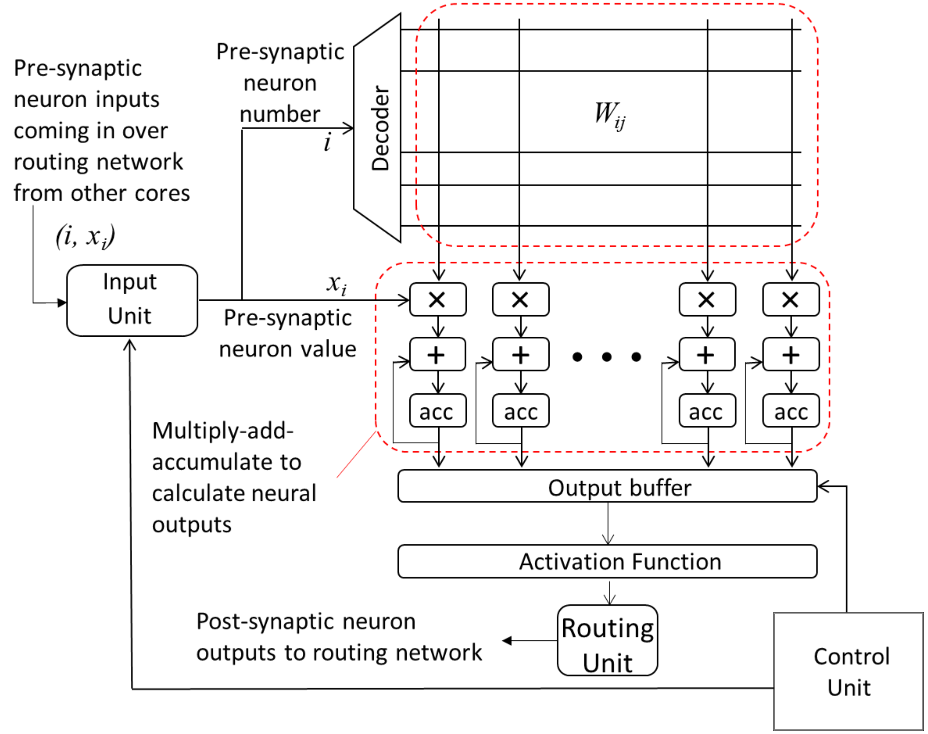


Multicore Neural Processor

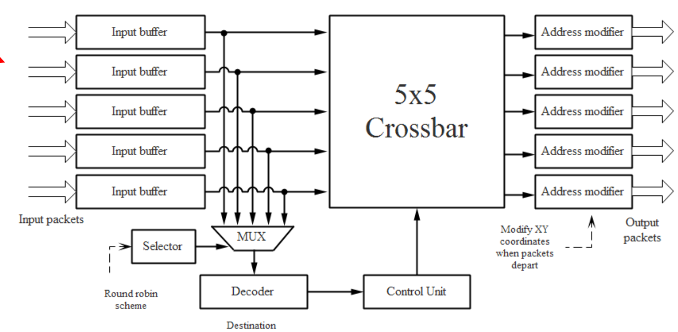
NC = Neural Core
R = Router



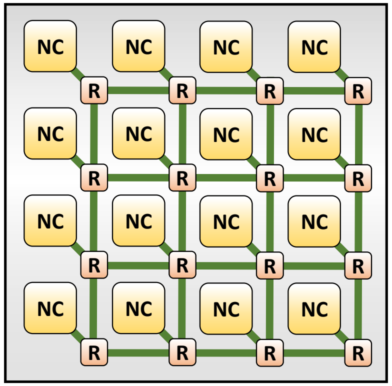
Neural Core

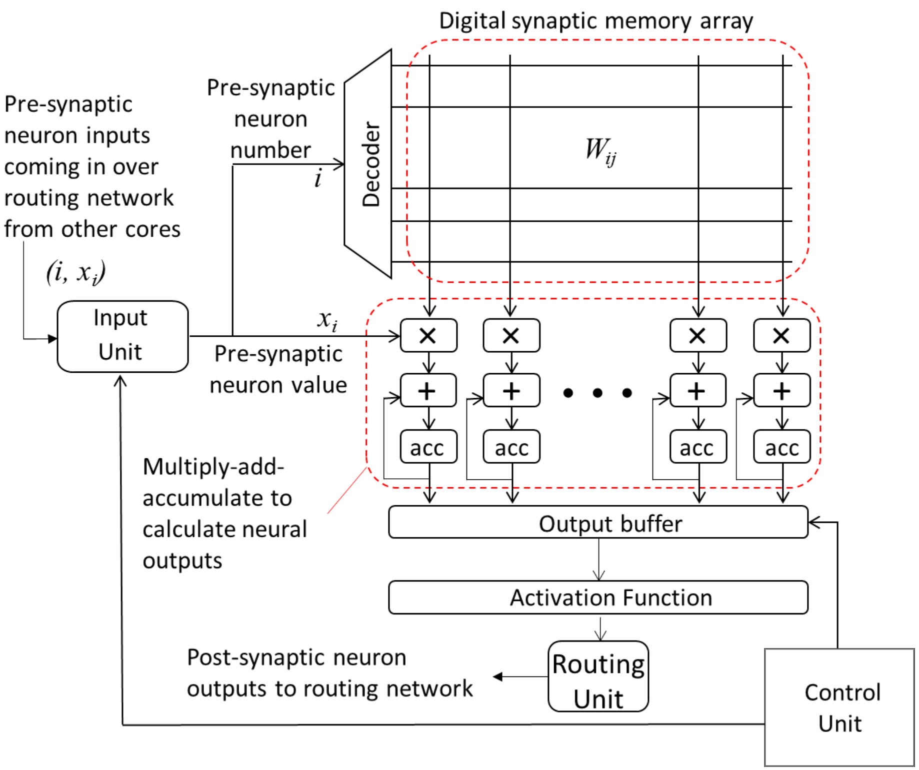


Router

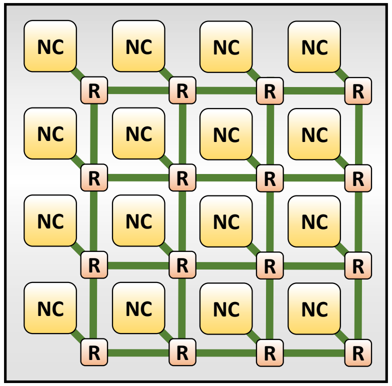


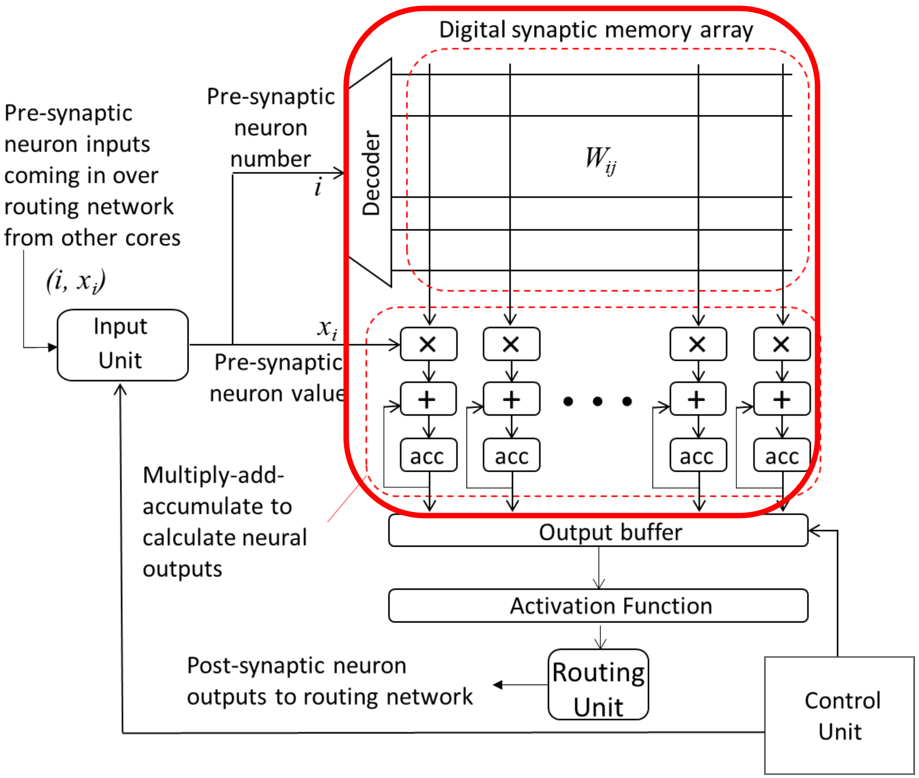
Mixed-Signal Neural Processor



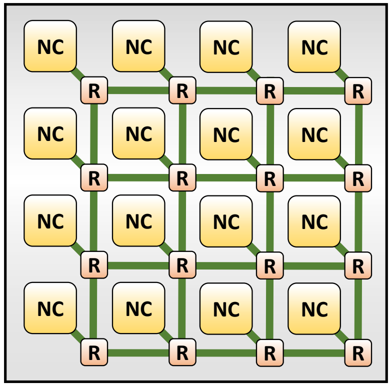


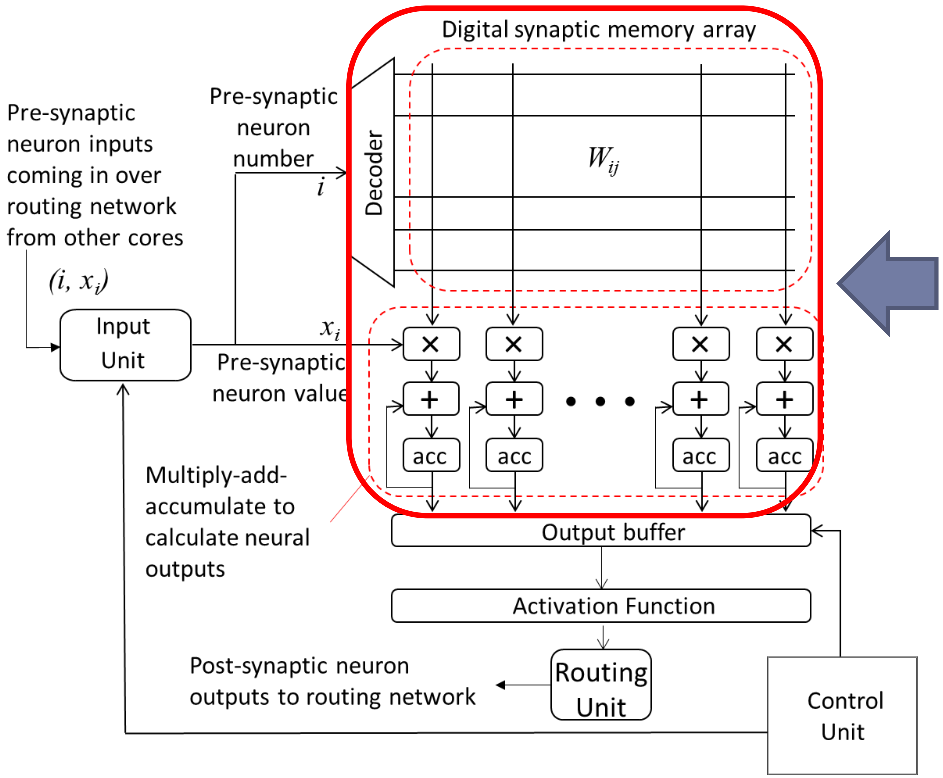
Mixed-Signal Neural Processor

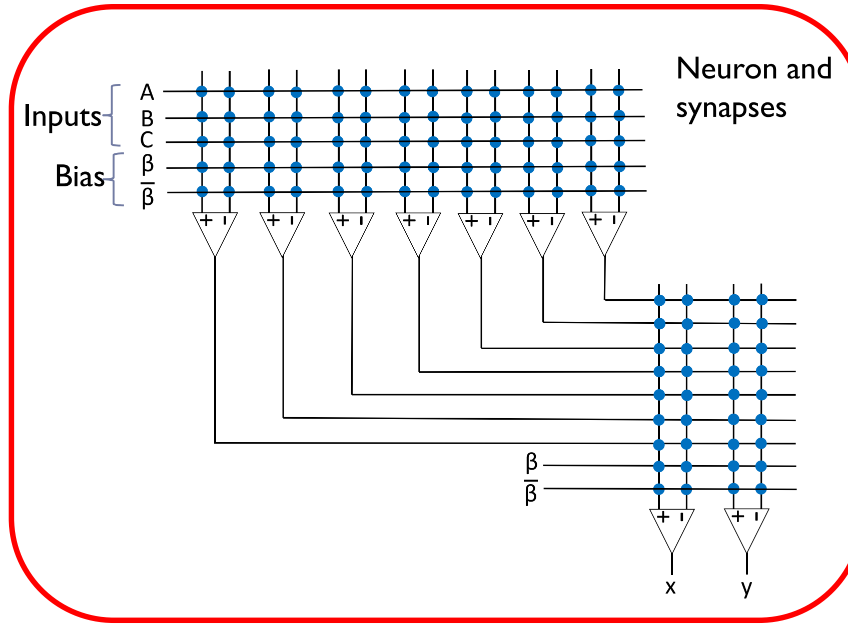




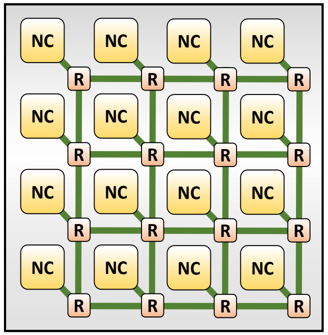
Mixed-Signal Neural Processor

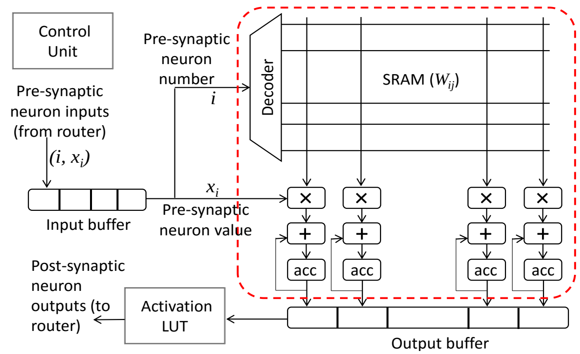


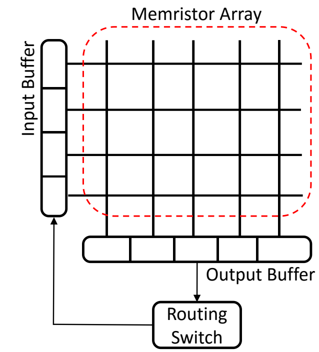




System Comparison







Deep Network

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	46	127.42	41,400.0	1
NN SRAM	31	1.91	16.3	2,540
NN Memristor	31	0.08	1.2	34,968

Edge Detection

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	43	119.11	38,700.00	1
NN SRAM	19	1.17	9.75	3,874
NN Memristor	146	0.38	5.6	6,941

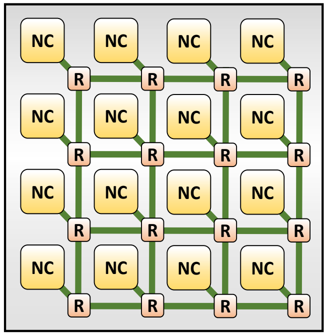
Motion Estimation

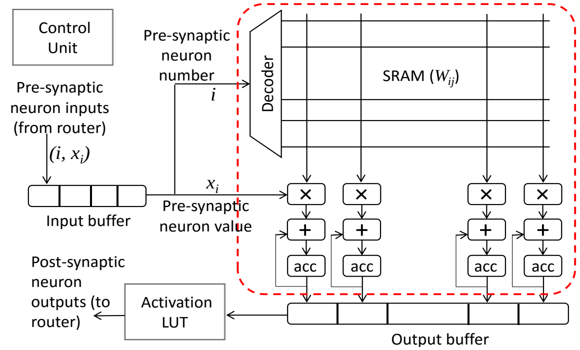
	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	2	5.54	1800.00	1
NN SRAM	2	0.12	1.05	1712
NN Memristor	4	0.01	0.15	11,783

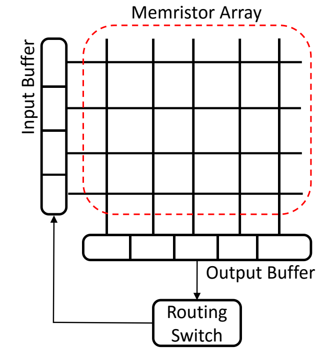
Optical Character Recognition

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	42	116.34	37,800.0	1
NN SRAM	31	1.91	16.3	2,319
NN Memristor	31	0.08	1.2	31,928

System Comparison







Deep Network

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	46	127.42	41,400.0	1
NN SRAM	31	1.91	16.3	2,540
NN Memristor	31	0.08	1.2	34,968

Edge Detection

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	43	119.11	38,700.00	1
NN SRAM	19	1.17	9.75	3,874
NN Memristor	146	0.38	5.6	6,941

Motion Estimation

	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	2	5.54	1800.00	1
NN SRAM	2	0.12	1.05	1712
NN Memristor	4	0.01	0.15	11,783

Optical Character Recognition

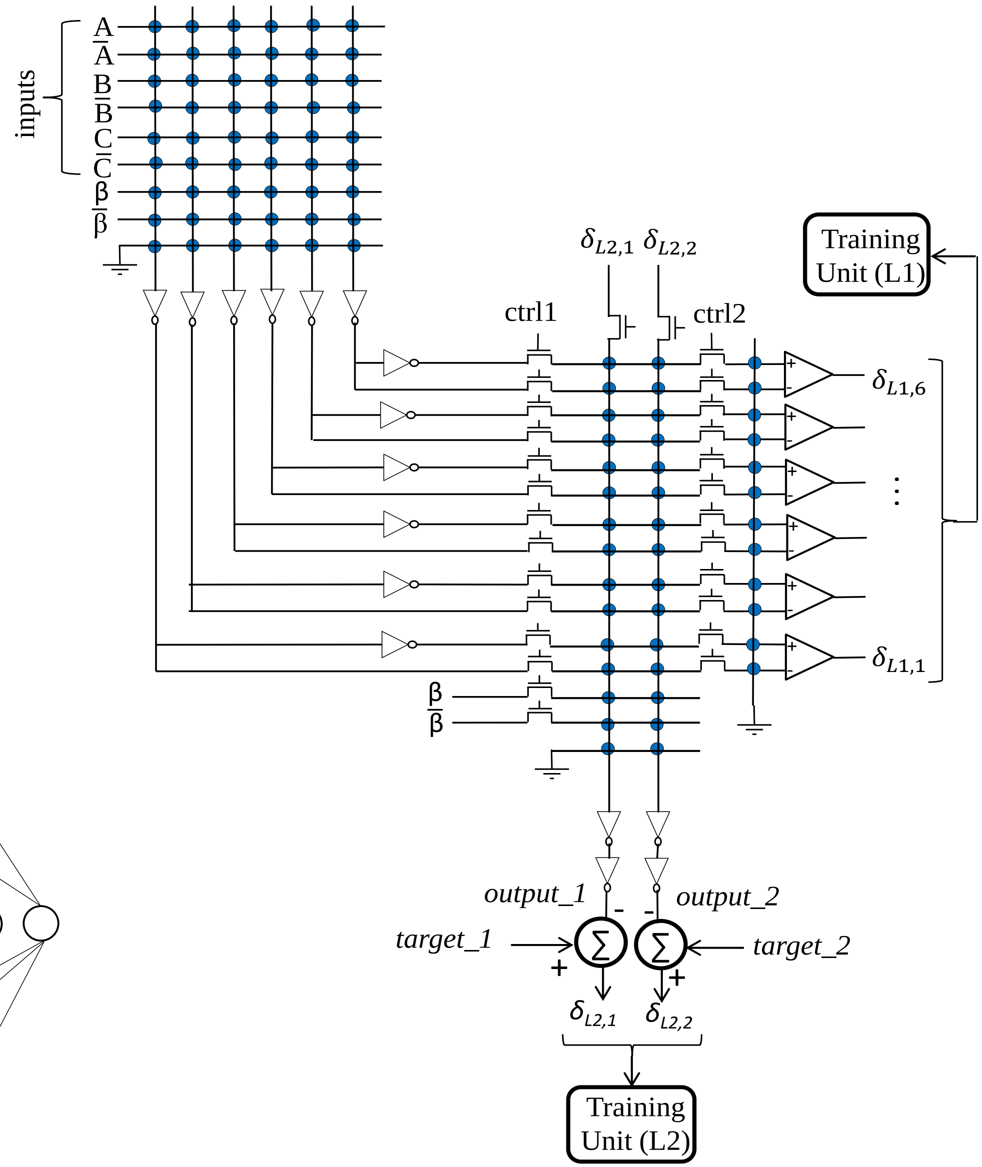
	Number of core	Area (mm^2)	Power (mW)	Power efficiency over RISC
RISC	42	116.34	37,800.0	1
NN SRAM	31	1.91	16.3	2,319
NN Memristor	31	0.08	1.2	31,928

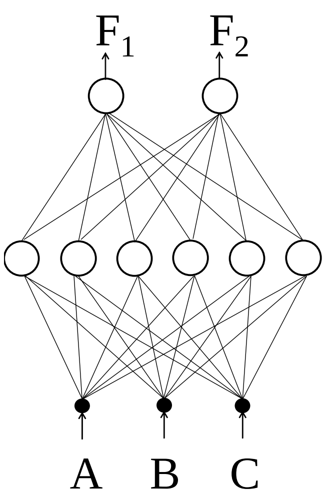
	Deep Net		Edge		Motion		OCR	
	mJ	Energy eff	mJ	Energy eff	mJ	Energy eff	mJ	Energy eff
RISC	41,400	1	38,700	1	1800	1	37,800	1
NN SRAM	6.888	6,010	8.533	4,536	0.667	2,700	6.888	5,488
NN Memristor	0.032	1,303,491	0.197	196,734	0.003	585,627	0.032	1,190,144

Tarek Taha

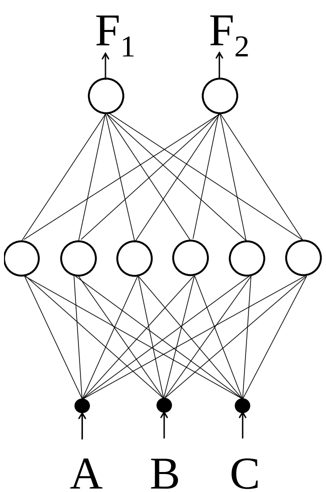
On-chip Learning

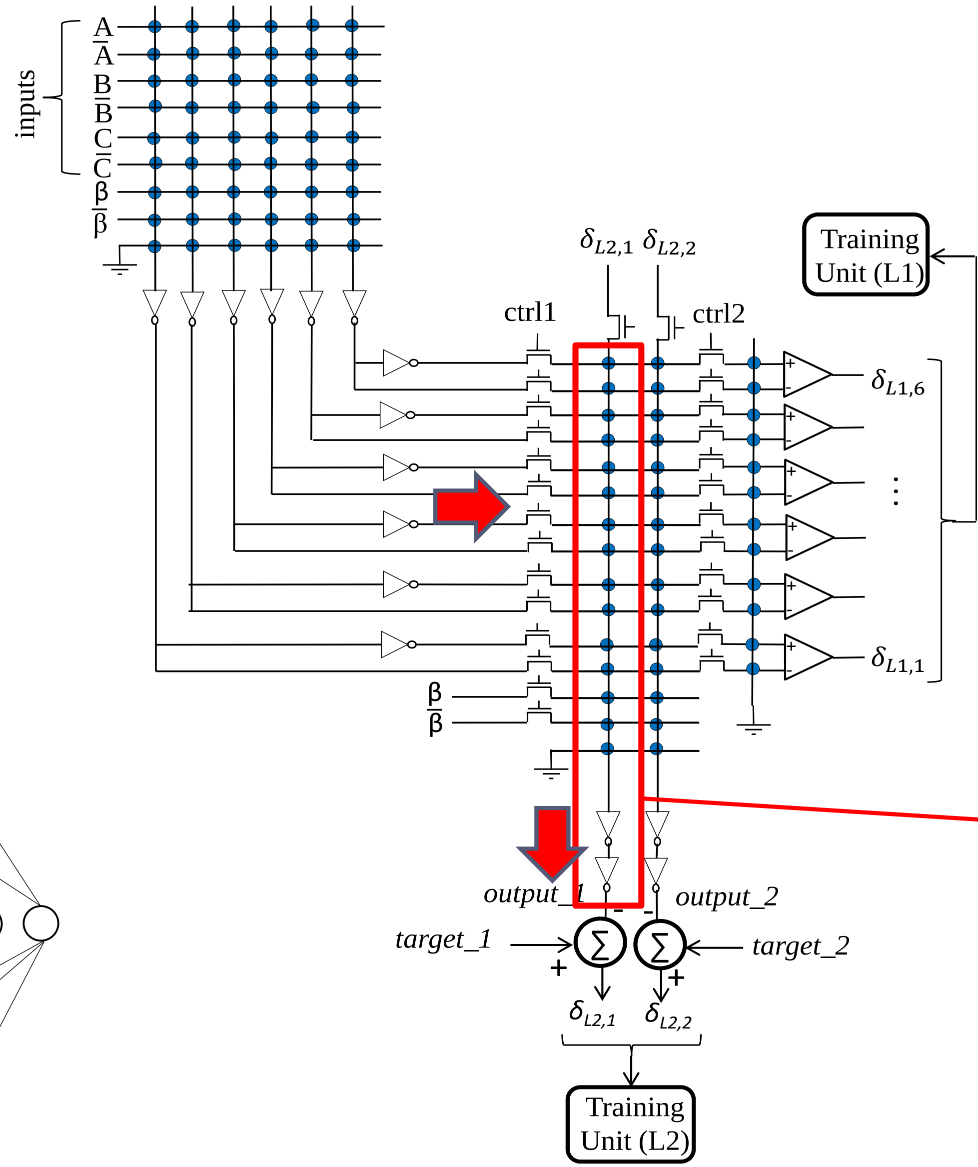
Backpropagation Circuit

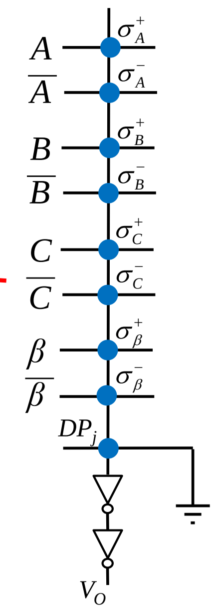




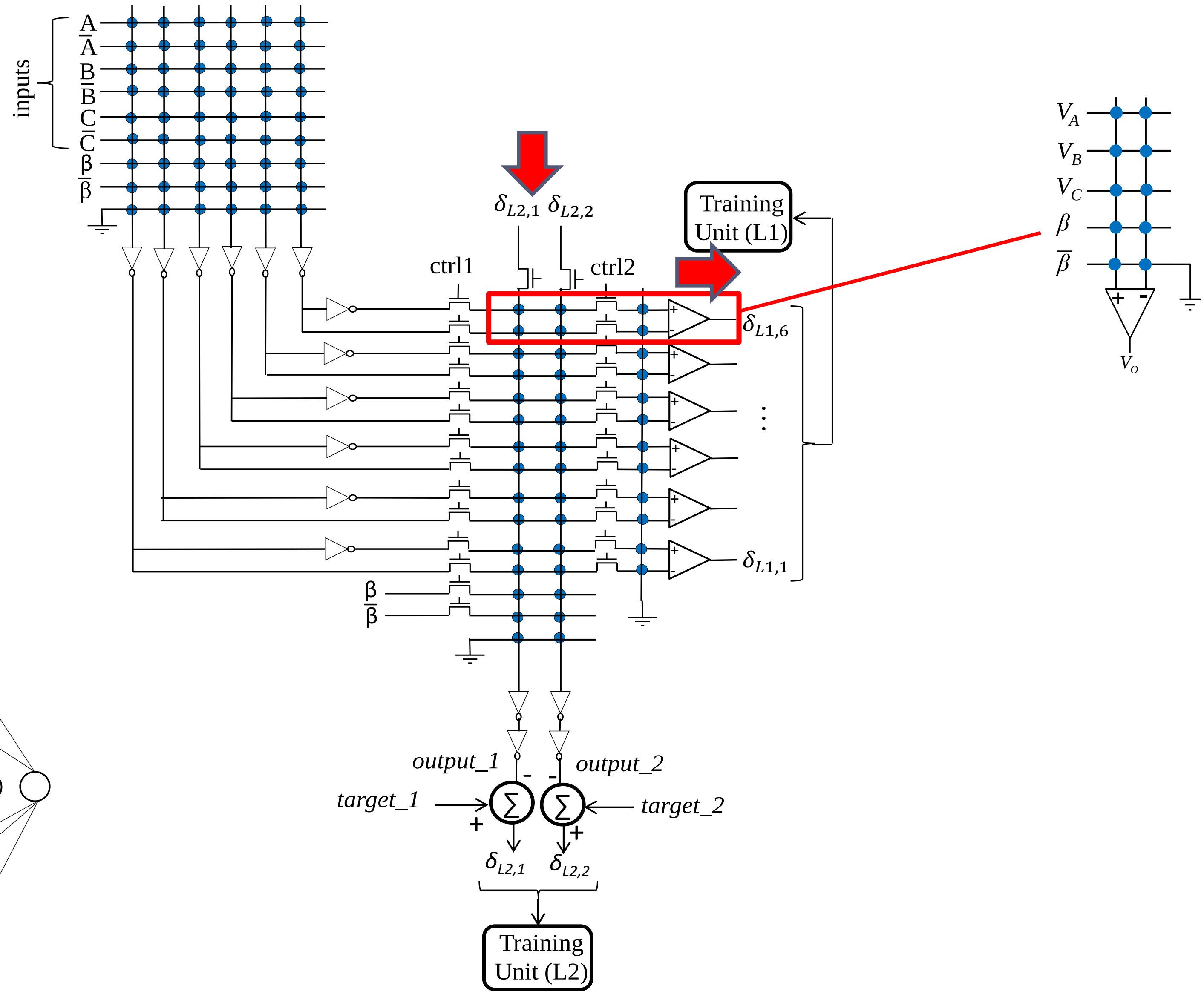
Backpropagation Circuit

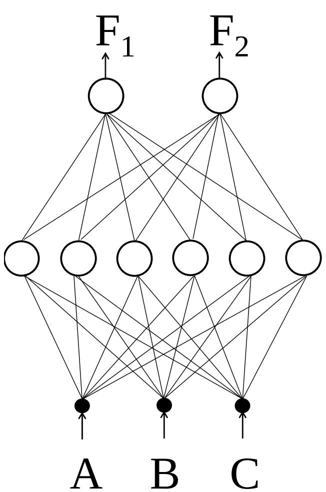




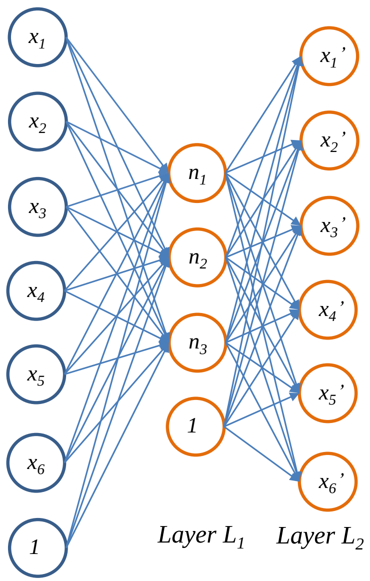


Backpropagation Circuit





Unsupervised Clustering



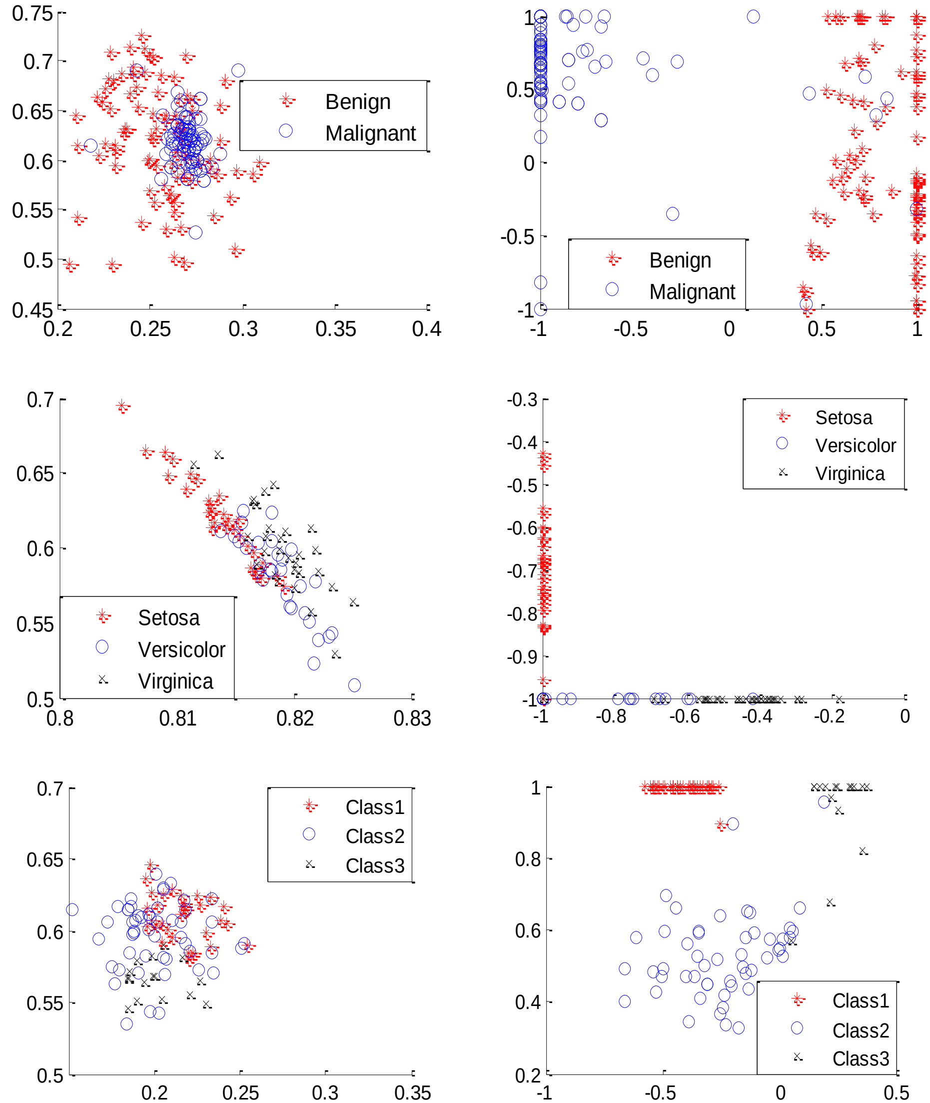
Before Training

After Training

Wisconsin

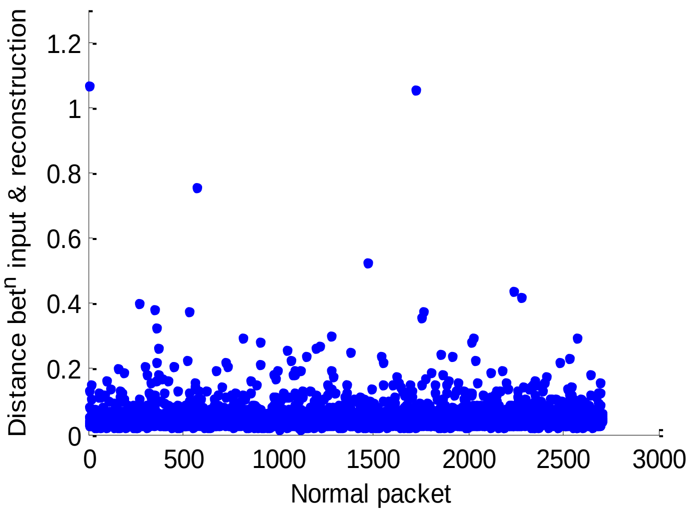
Iris

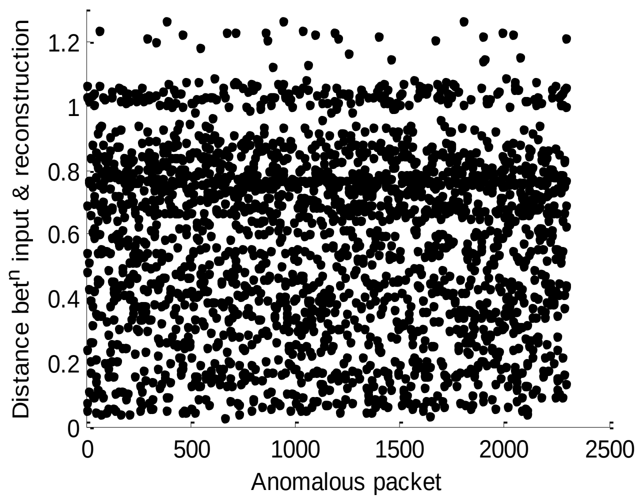
Wine

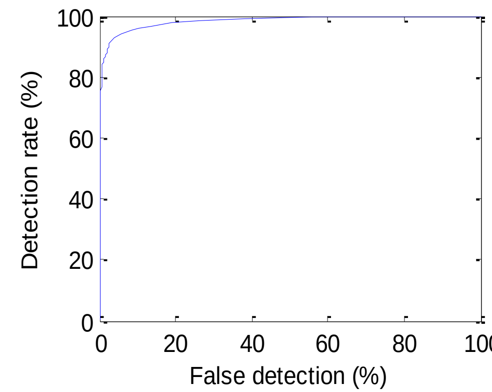


Tarek Taha

Anomaly Detection in Network Traffic

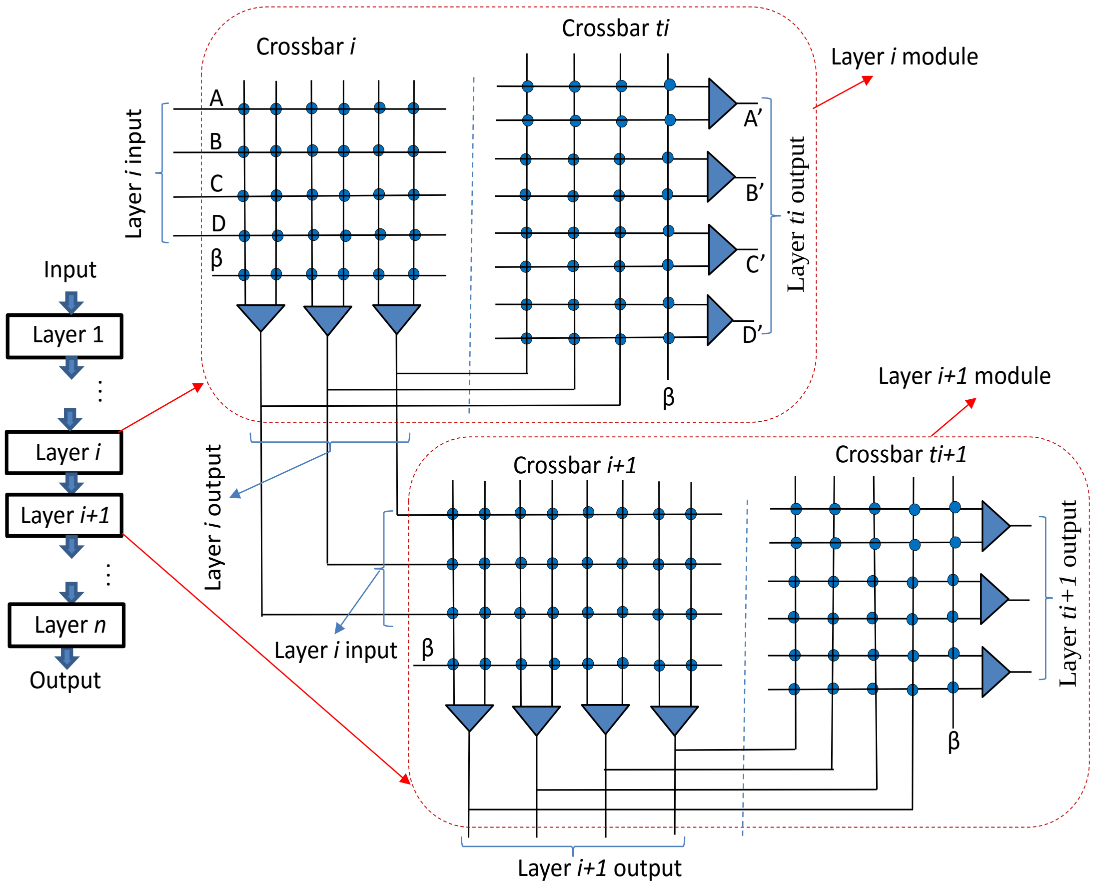




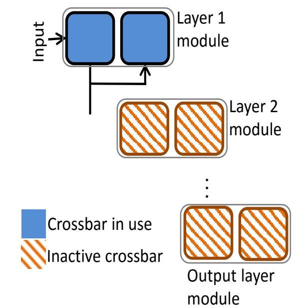


- 96.6% of the anomalous packets detected
- 4% false positive detection

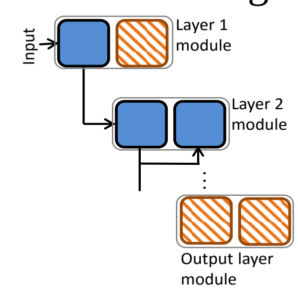
Deep Learning Architecture



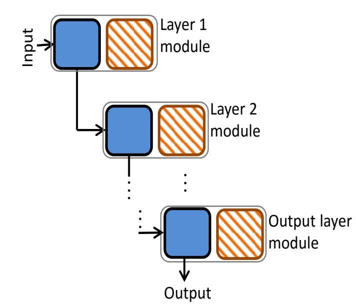
1. Pre-training layer 1



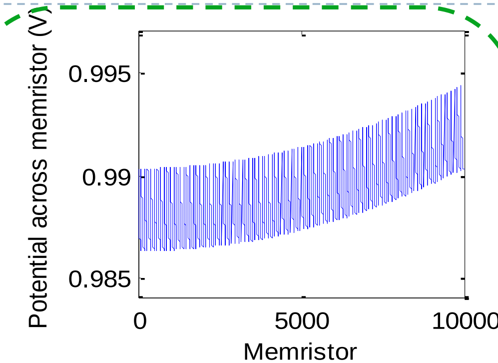
2. Pre-training layer 2



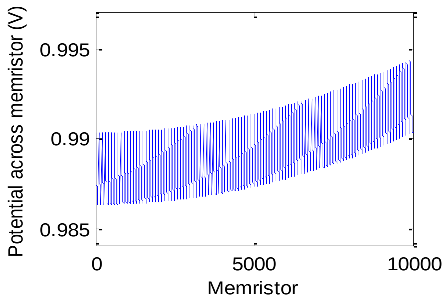
3. Supervised training of the whole network.



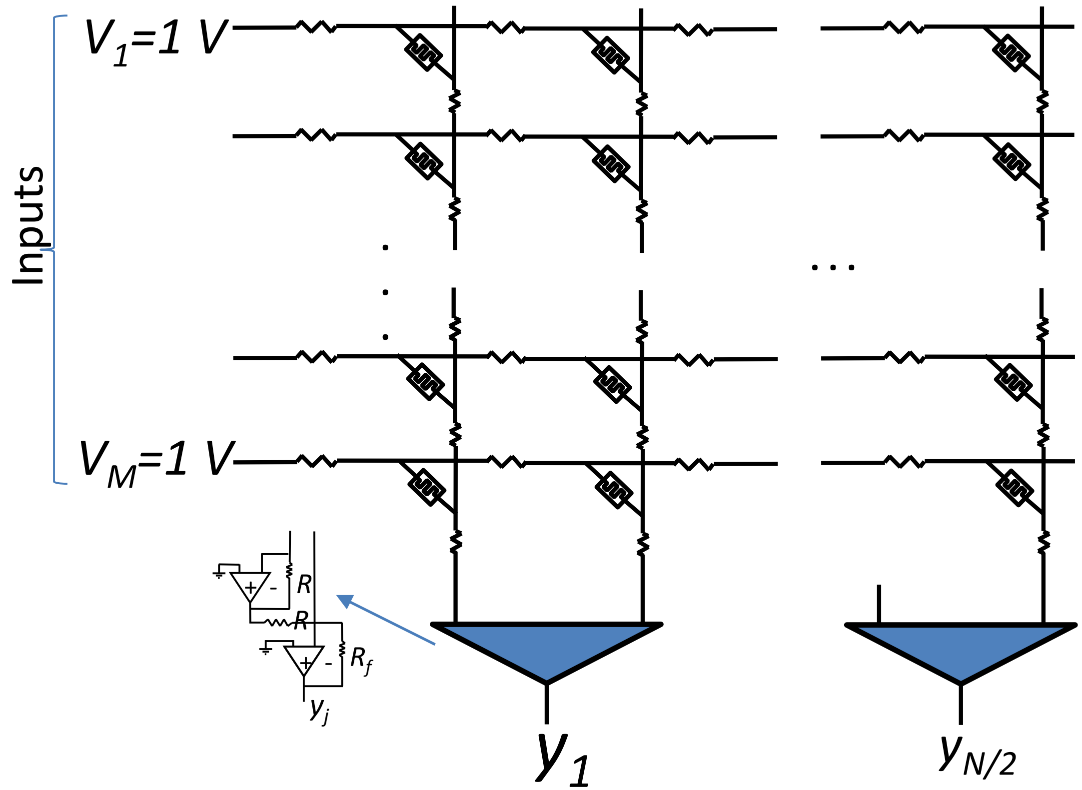
Large Crossbar Simulation

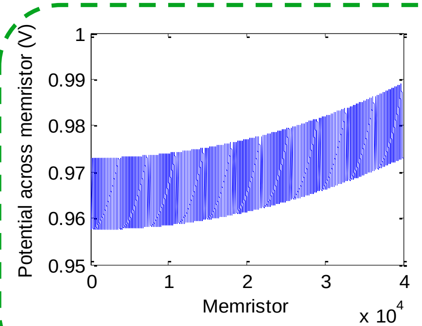


SPICE 100×100

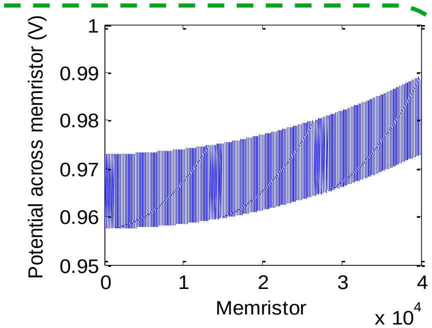


MATLAB 100×100

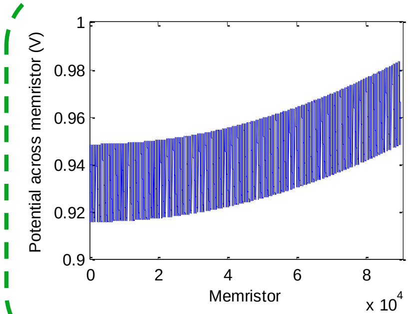




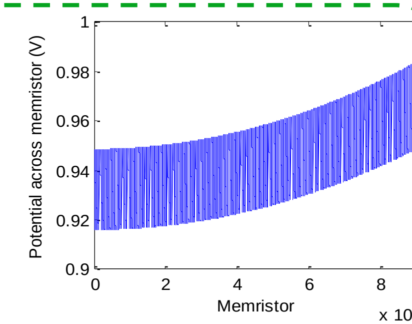
SPICE 200×200



MATLAB 200×200

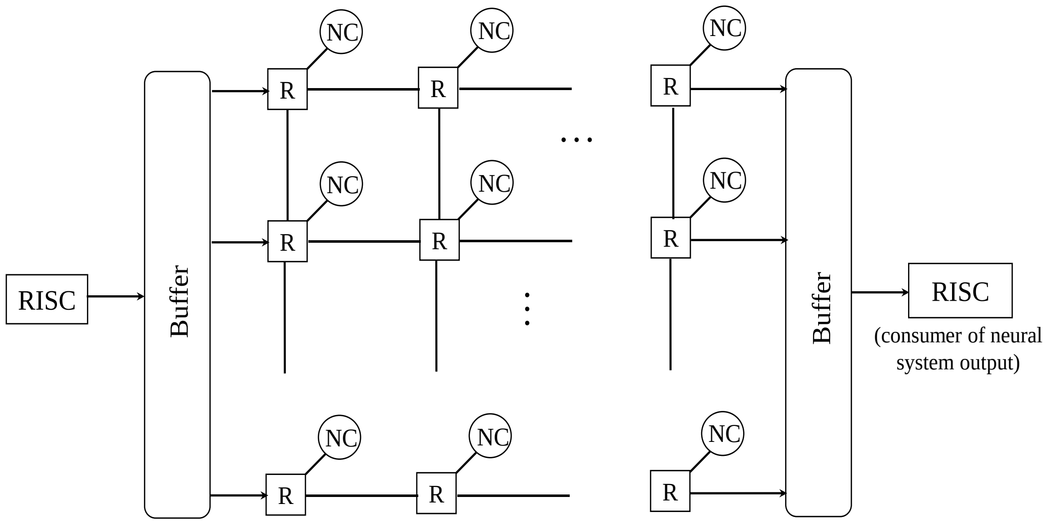


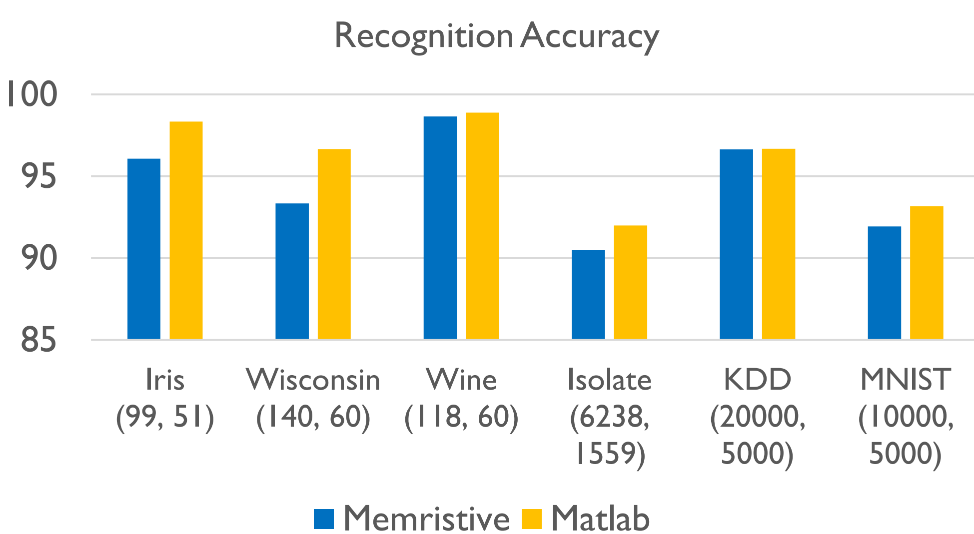
SPICE 300×300



MATLAB 300×300

Deep Learning





Performance vs. Tesla K20

Training

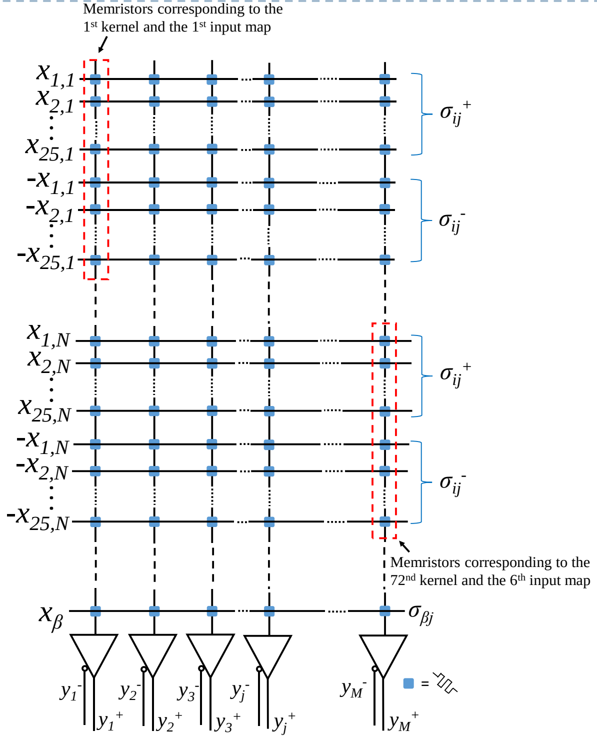
	Energy eff.	Speedup
MNIST	26,597	6.9
Isolate	12,822	4.6
KDD	239,435	3.0

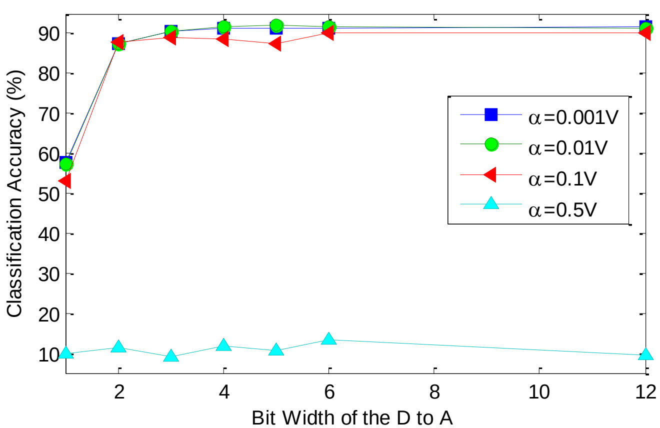
Recognition

	Energy eff.	Speedup
MNIST	314,299	41.0
Isolate	147,308	50.5
KDD	375,252	10.2

Related Projects

Convolution Neural Network





Input Image

DSP Output

Mem Output: 3 bit

Mem Output: 2 bit

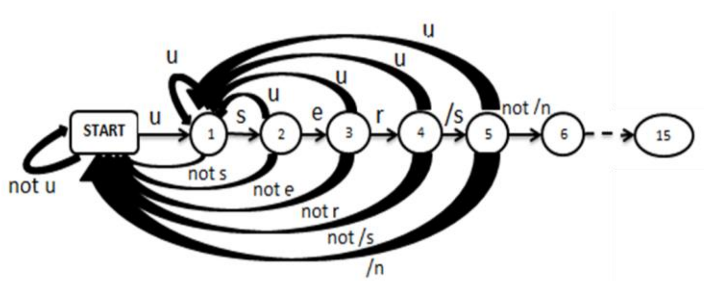
α=0.01

UNIVERSITY of DAYTON

Cybersecurity

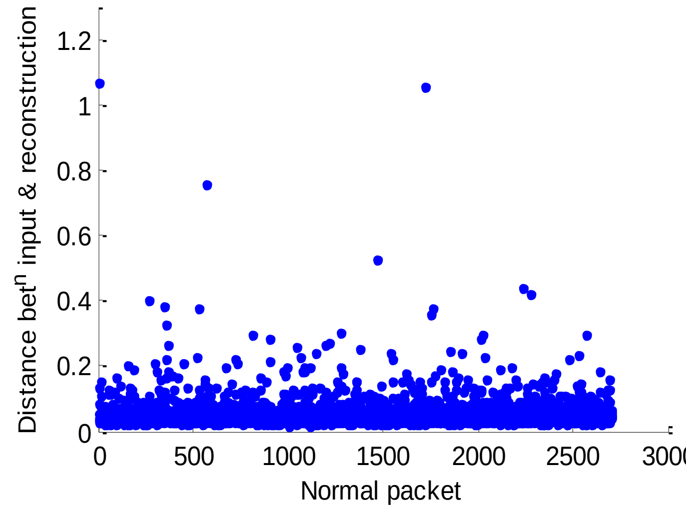
Signature Based Detection

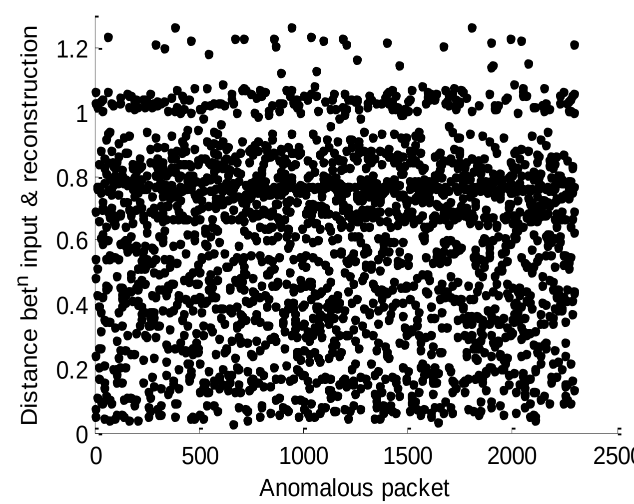
- Large collection of state machines
- 2.2 nW per rule
- 3.8 Gbps



Regex : *user\s[^\n]{10}*

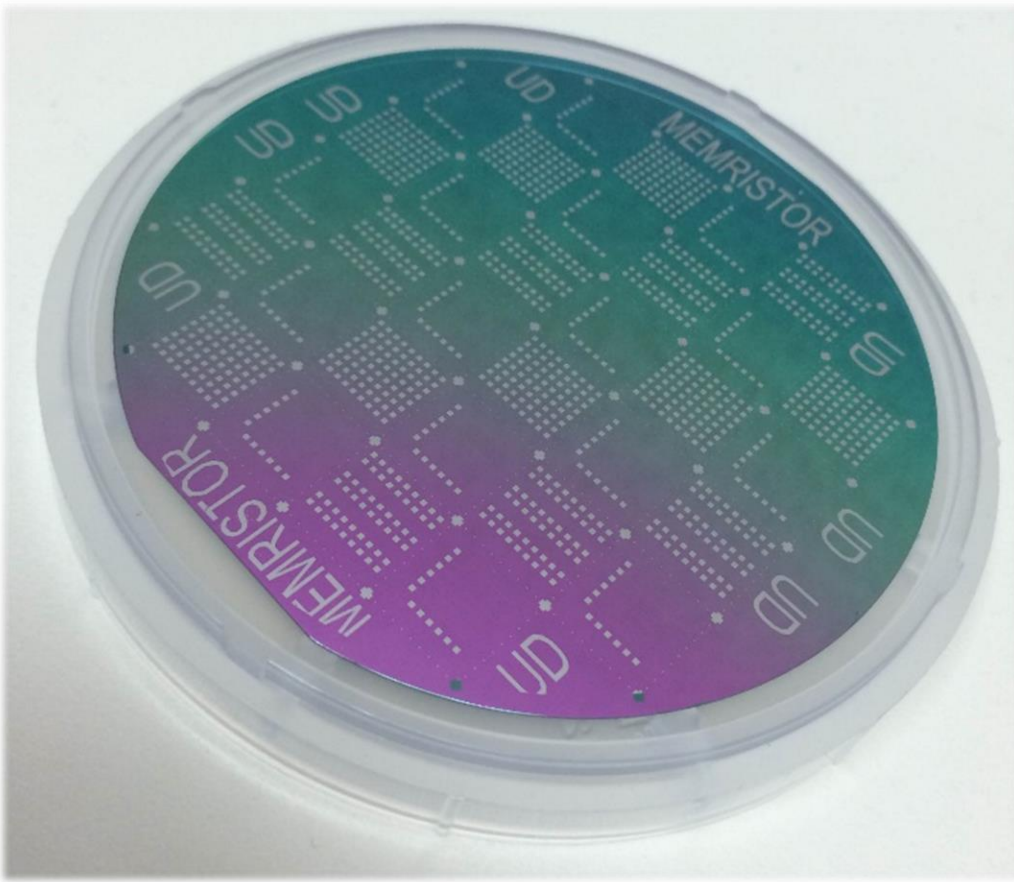
Anomaly Detection



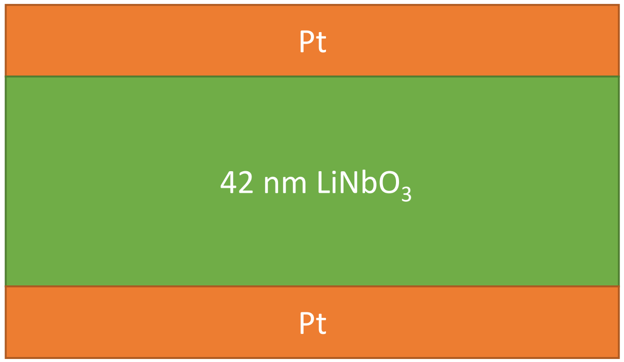


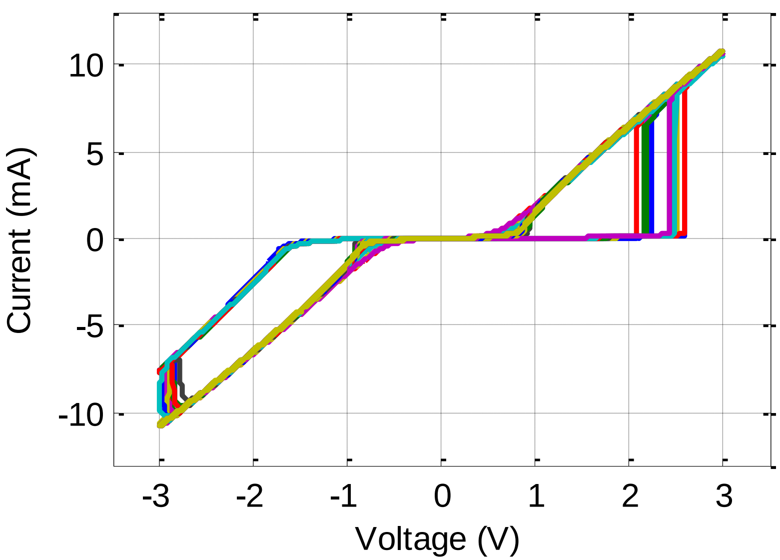
Device Fabrication

- Memristor device is based on a $LiNbO_3$ switching oxide
- IV curve shows repeatable switching

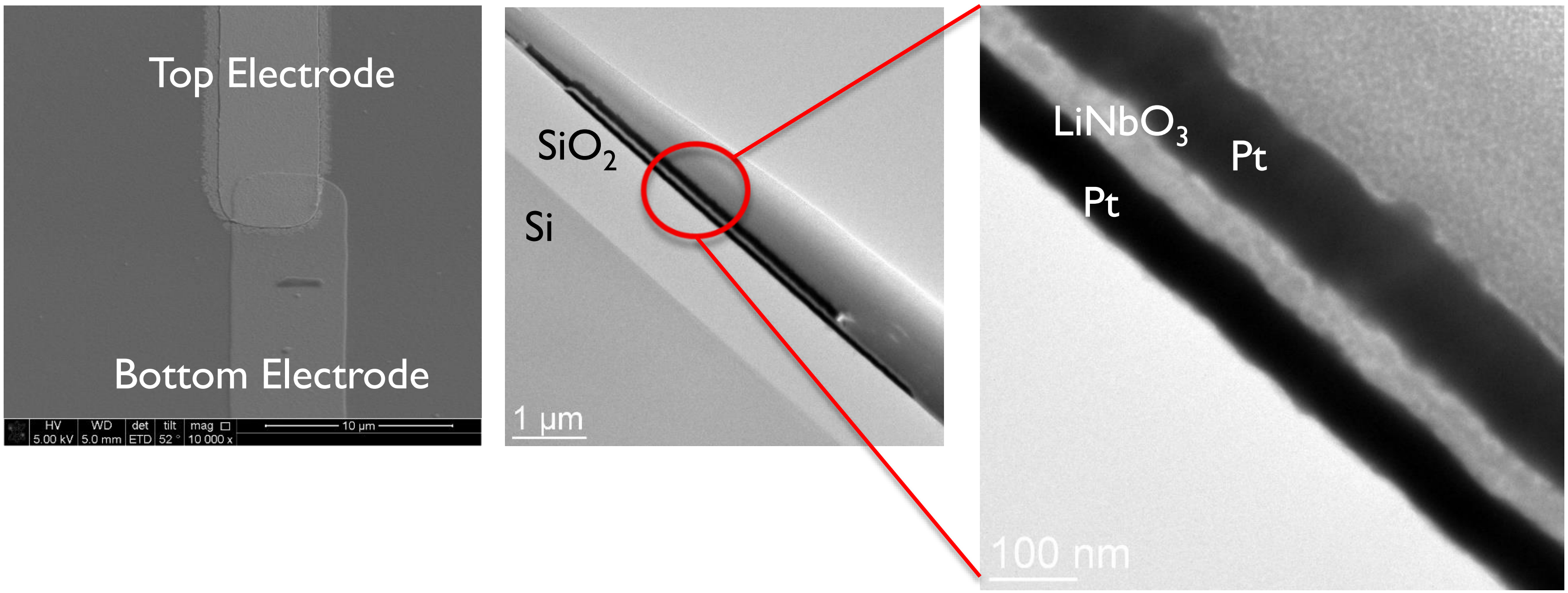


Device Structure





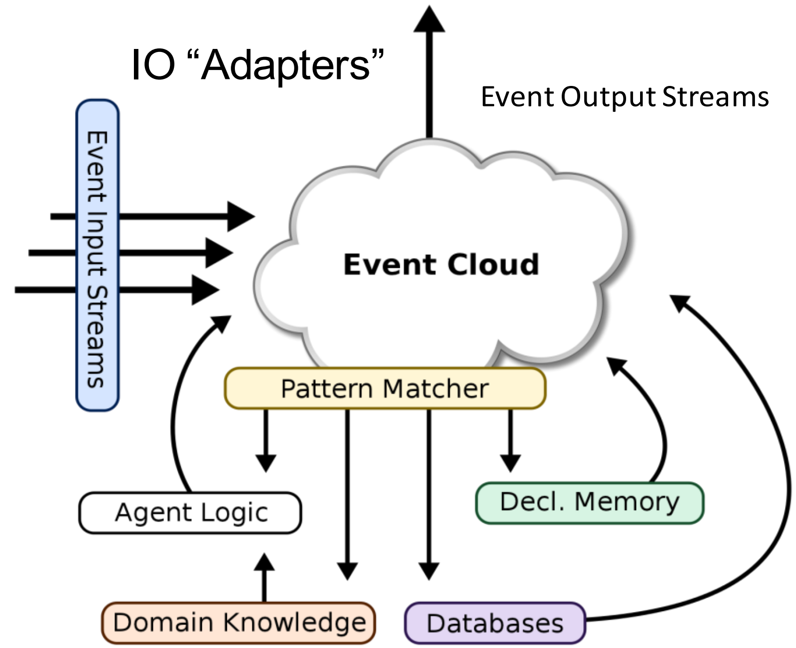
SEM/TEM results



Autonomous Agent

- Cognitively Enhanced Complex Event Processing (CECEP) Architecture:

- Consists of the following central net-centric components:
  - <u>soaDM:</u> an associative memory application that allows agents to store and retrieve declarative knowledge.
  - <u>soaCDO:</u> a knowledge representation and mining application that allows agents to store and exploit domain knowledge.
  - <u>Esper:</u> a complex event processing framework that allows agents to base actions on context assessment and procedural knowledge.



Acknowledgements

Research Engineers:
- Raqib Hasan, PhD
- Chris Yakopcic, PhD
- Wei Song, PhD
- Tanvir Atahary, PhD

Doctoral Students:
- Zahangir Alom
- Hua Chen
- Chong Chen
- Rasitha Fernando
- Ted Josue
- Will Mitchell
- Yangjie Qi
- Nayim Rahman

Sponsors:

http://homepages.udayton.edu/~ttaha1