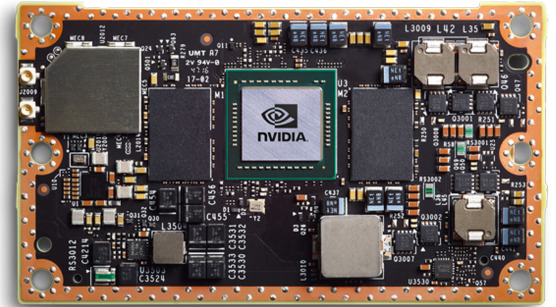


<https://developer.nvidia.com/discover/artificial-neural-network>

<https://developer.nvidia.com/embedded/jetson-tx2>

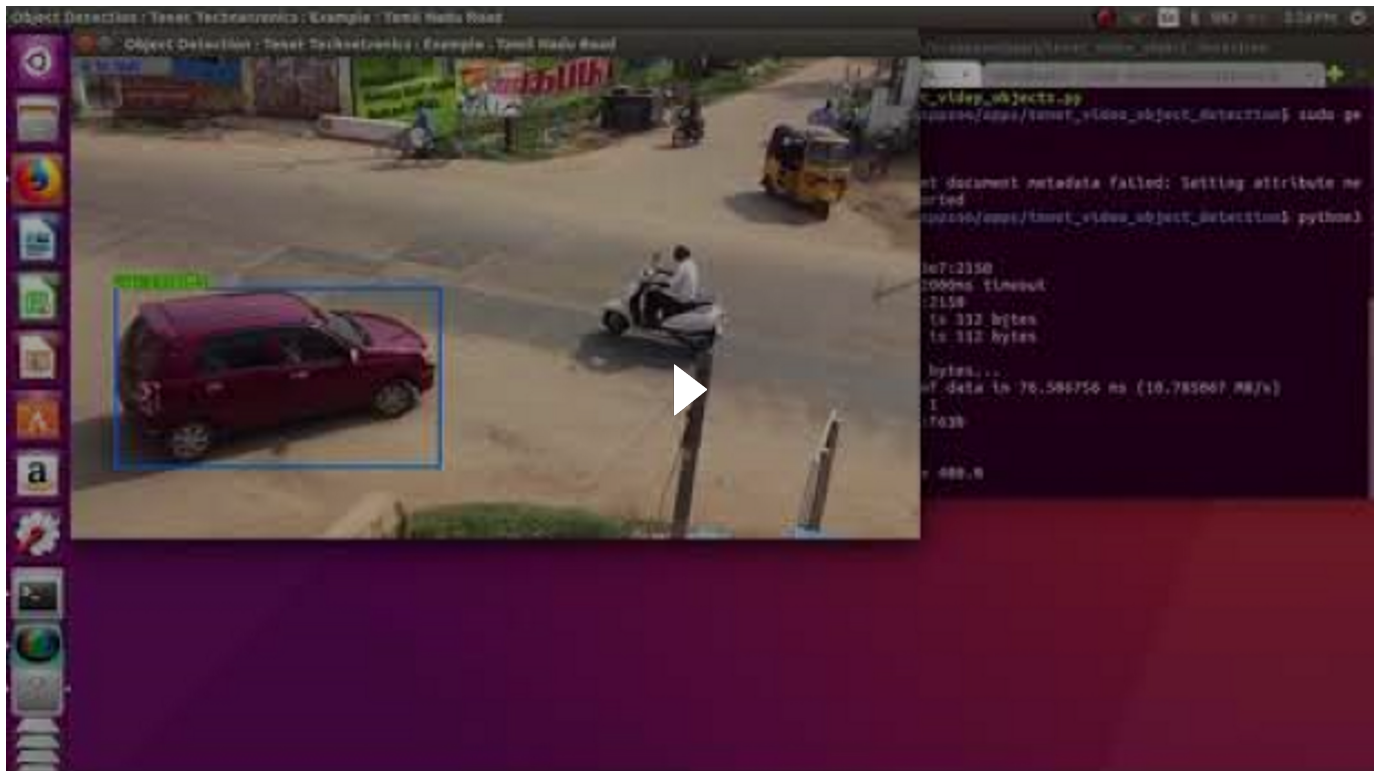
<https://developer.nvidia.com/embedded/jetson-nano-developer-kit>

<https://www.nvidia.com/en-eu/geforce/graphics-cards/rtx-2080-super/>



<https://software.intel.com/en-us/neural-compute-stick>

[Vehicle detection using Movidius compute neural stick](#)



[Deep Learning with Intel](#)



<https://www.amd.com/en/technologies/deep-machine-learning>

<https://www.arm.com/products/silicon-ip-cpu/ethos/ethos-n77>

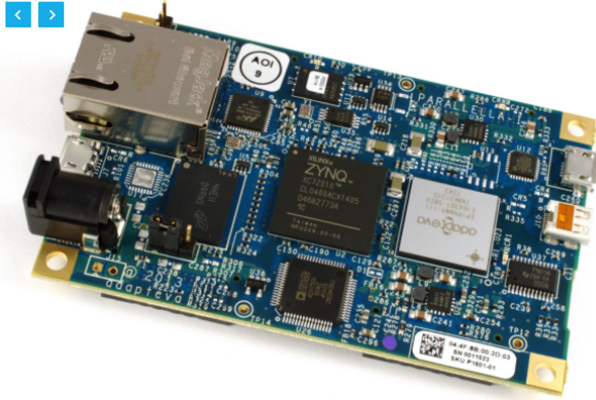
### KEY USE CASES FOR ETHOS-N57

- + Object classification
- + Object detection
- + Face detection/identification
- + Human pose detection/  
hand-gesture recognition
- + Image segmentation
- + Image beautification
- + Super resolution
- + Framerate adjustment  
(super slow-mo)
- + Speech recognition
- + Sound recognition
- + Noise cancellation
- + Speech synthesis
- + Language translation

### Specifications

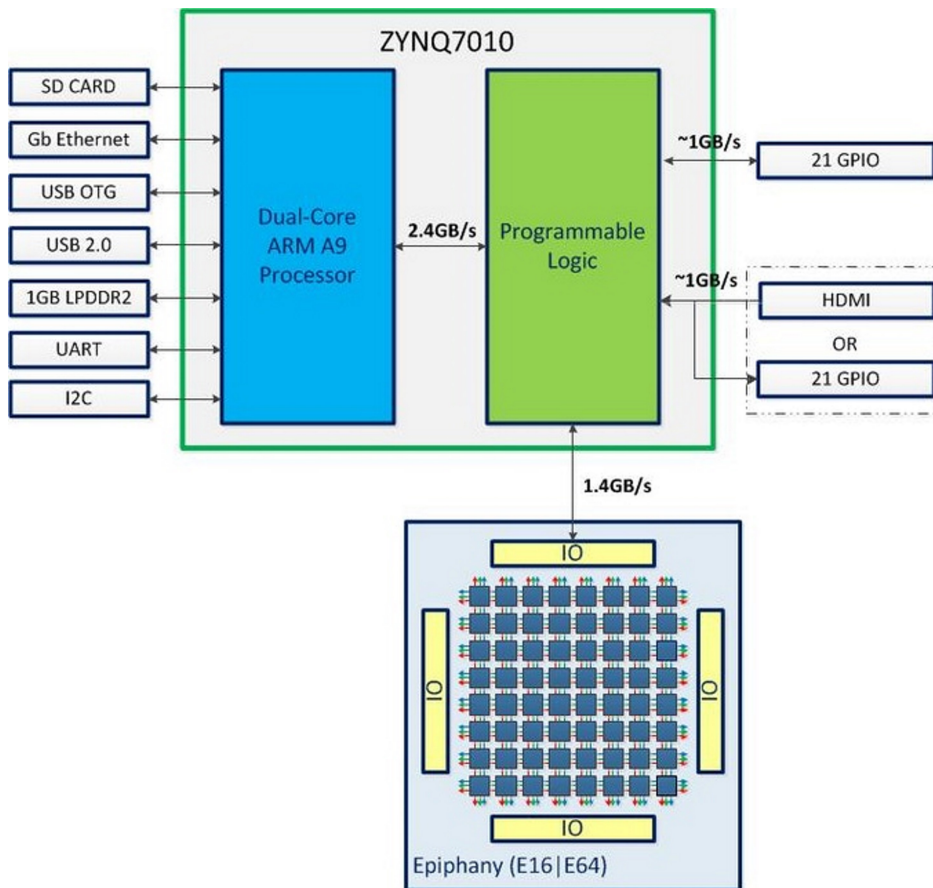
Key Features	Performance (at 1GHz)	Up to 1 TOP/s
	MACs (8x8)	512
	Data Types	Int-8 and Int-16
	Network Support	CNN and RNN
	Efficient Convolution	Winograd support
	Sparsity	Yes
	Secure Mode	TEE or SEE
	Multicore Capability	8 NPU's in a cluster 64 NPU's in a mesh
Memory System	Embedded SRAM	512KB
	Bandwidth Reduction	Extended compression technology, layer/ operator fusion
	Main Interface	1xAXI4 (128-bit), ACE-5 Lite
Development Platform	Neural Frameworks	TensorFlow, TensorFlow Lite, Caffe2, PyTorch, MXNet, ONNX
	Neural Operator API	Arm NN, AndroidNN
	Software Components	Arm NN, neural compiler, driver and support library
	Debug and Profile	Layer-by-layer visibility
	Evaluation and Early Prototyping	Arm Juno FPGA systems and cycle models

2008 - <https://www.parallella.org/>



# The Parallella Board

- 18-core credit card sized computer
- #1 in energy efficiency @ 5W
- 16-core Epiphany RISC SOC
- Zynq SOC (FPGA + ARM A9)
- Gigabit Ethernet
- 1GB SDRAM
- Micro-SD storage
- Up to 48 GPIO pins
- HDMI, USB (optional)
- Open source design files
- Runs Linux
- >10,000 boards shipped
- Starting at \$99



<https://en.wikipedia.org/wiki/Adapteva>

DARPA UPSIDE program (2012-2018) Unconventional Processing of Signals for Intelligent Data Exploitation

Objective: Exploit the physics of emerging devices, analog CMOS, and non-Boolean computational models to achieve new levels of performance and power for real-time sensor imaging systems.

**University of Michigan**

**UCSB**

3.8 mm      18 μm

**Key Takeaways**

Analog computing beats digital on VMMs

**Challenges:**

- Comparing results (lack of data) → RTML
- Transition valley of death → RTML
- High cost of design → RTML
- Manufacturing latency too long → RTML
- Manufacturability and scalability → RTML

- Mixed signal processing (50TOPS/W)
- Sparse image reconstruction in memristors
- Numerous publications (Nature, ...)

- First memristor based multilayer perceptron
- Flash based 55nm analog computing (>10TOPS/W)
- Numerous publications (Nature, ...)

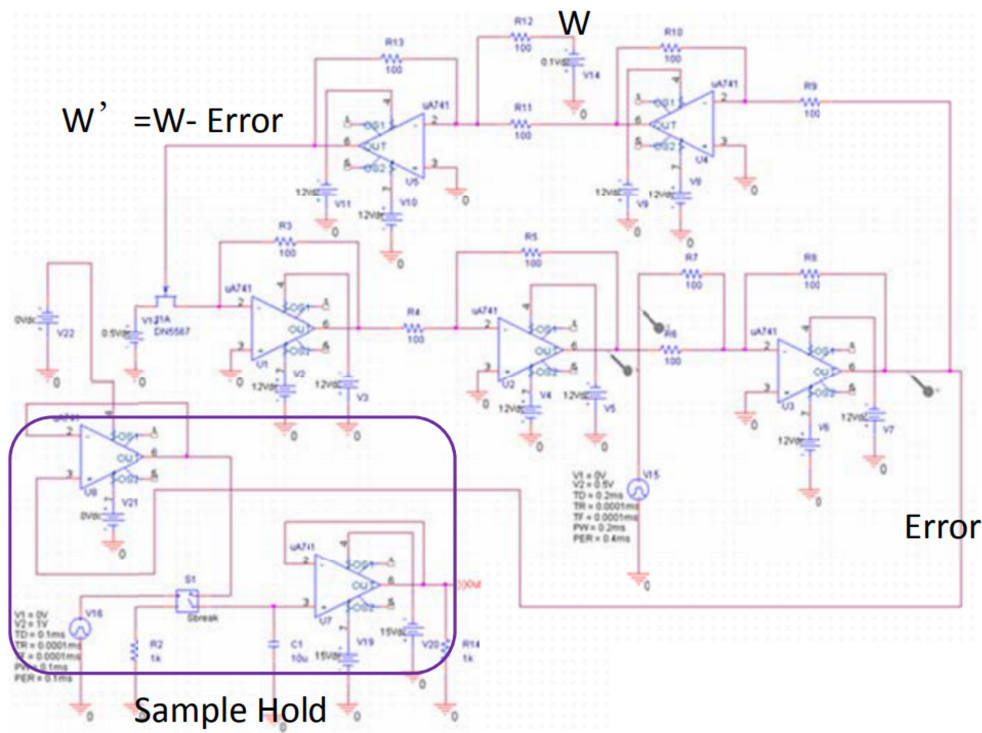
## Analog Neural Circuit and Hardware Design of Deep Learning Model

Masashi Kawaguchi<sup>a</sup>, Naohiro Ishii<sup>b</sup>, Masayoshi Umeno<sup>c</sup>

<sup>a</sup>National Institute of Technology, Suzuka College, Shiroko Suzuka Mie 510-0294, Japan

<sup>b</sup>Aichi Institute of Technology, Yachigusa Yagusa-cho Toyota 470-0392, Japan

<sup>c</sup>Chubu University, 1200 Matsumoto-cho Kasugai Aichi 487-8501, Japan



<http://www.research.ibm.com/articles/brain-chip.shtml>

## SIZE

TrueNorth is the first single, self-contained chip to achieve:

- One million individually programmable neurons--sixteen times more than the current largest neuromorphic chip
- 256 million individually programmable synapses on chip which is a new paradigm
- 5.4B transistors. By device count, largest IBM chip ever fabricated, second largest (CMOS) chip in the world
- 4,096 parallel and distributed cores, interconnected in an on-chip mesh network
- Over 400 million bits of local on-chip memory (~100 Kb per core) to store synapses and neuron parameters

From <<http://www.research.ibm.com/articles/brain-chip.shtml>>

## EFFICIENCY

TrueNorth is the first to achieve:

- 70mW total power while running a typical recurrent network at biological real-time, four orders of magnitude lower than a conventional computer running the same network
- Multi-object detection and classification application, with 240x400-pixel 3-color video input at 30 frames-per-second, the chip consumes 65mW
- 26pJ per synaptic event, which is the lowest recorded energy for any large-scale neuromorphic system, and five orders of magnitude lower than von Neumann computers
- 20mW / cm<sup>2</sup> power density which comparable to cortex but is three to four orders of magnitude lower compared to 50-100W / cm<sup>2</sup> for a CPU
- 30 times less area per neuron than than previous approaches (14.7um<sup>2</sup> compared to 400um<sup>2</sup> in analog)

From <<http://www.research.ibm.com/articles/brain-chip.shtml>>

