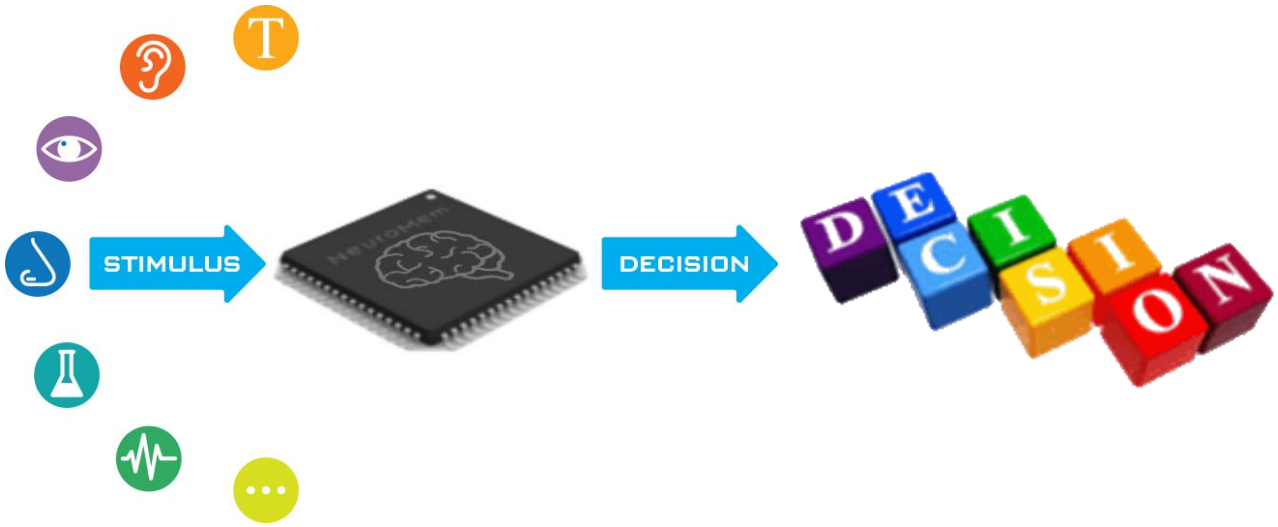
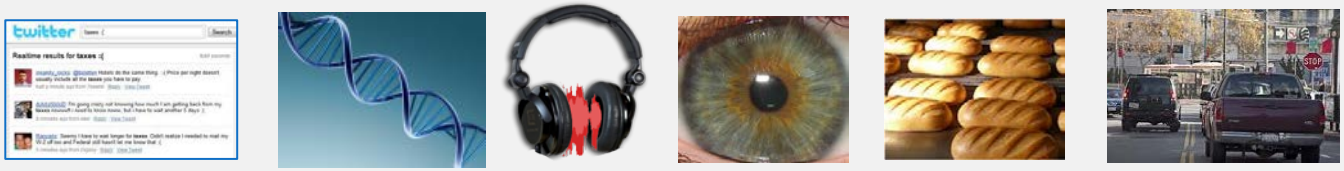
NeuroMem technology for IoT and Big Data



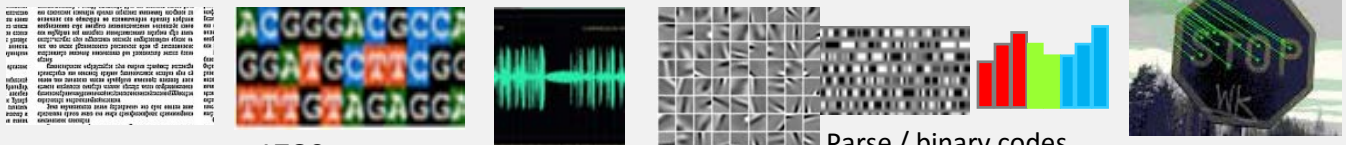
- Pattern recognition on silicon
- In-situ learning with intrinsic de-duplication
- Parallel processing performance
- Low-power, low-voltage

GENERAL VISION

Collect data source



Extract features or signatures



ASCII | ATGC sequence | Parse / binary codes | SURF/SIFT

Stimuli

NeuroMem

classifier and trainable neural network

Category

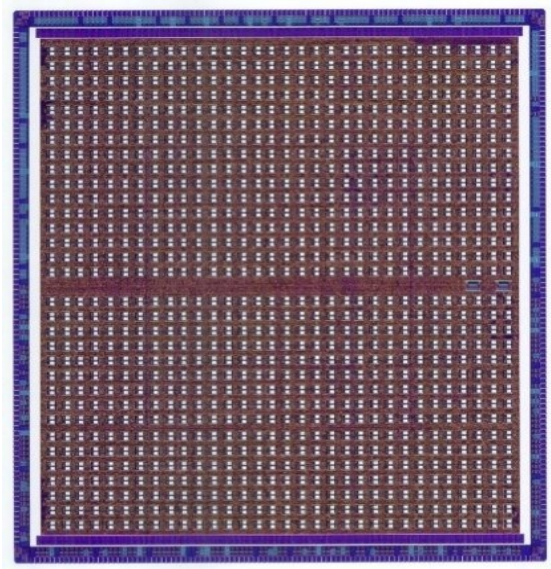
Produce insights, meta data and decisions

Combine | Prioritize | Filter

How is NeuroMem different?

- Pattern recognition chip:
  - Radial Basis Function and K-Nearest Neighbor
- Match 1 among N in 500 ns to 2.5 µsec
- Highly scalable due to natively parallel architecture

NeuroMem CM1K

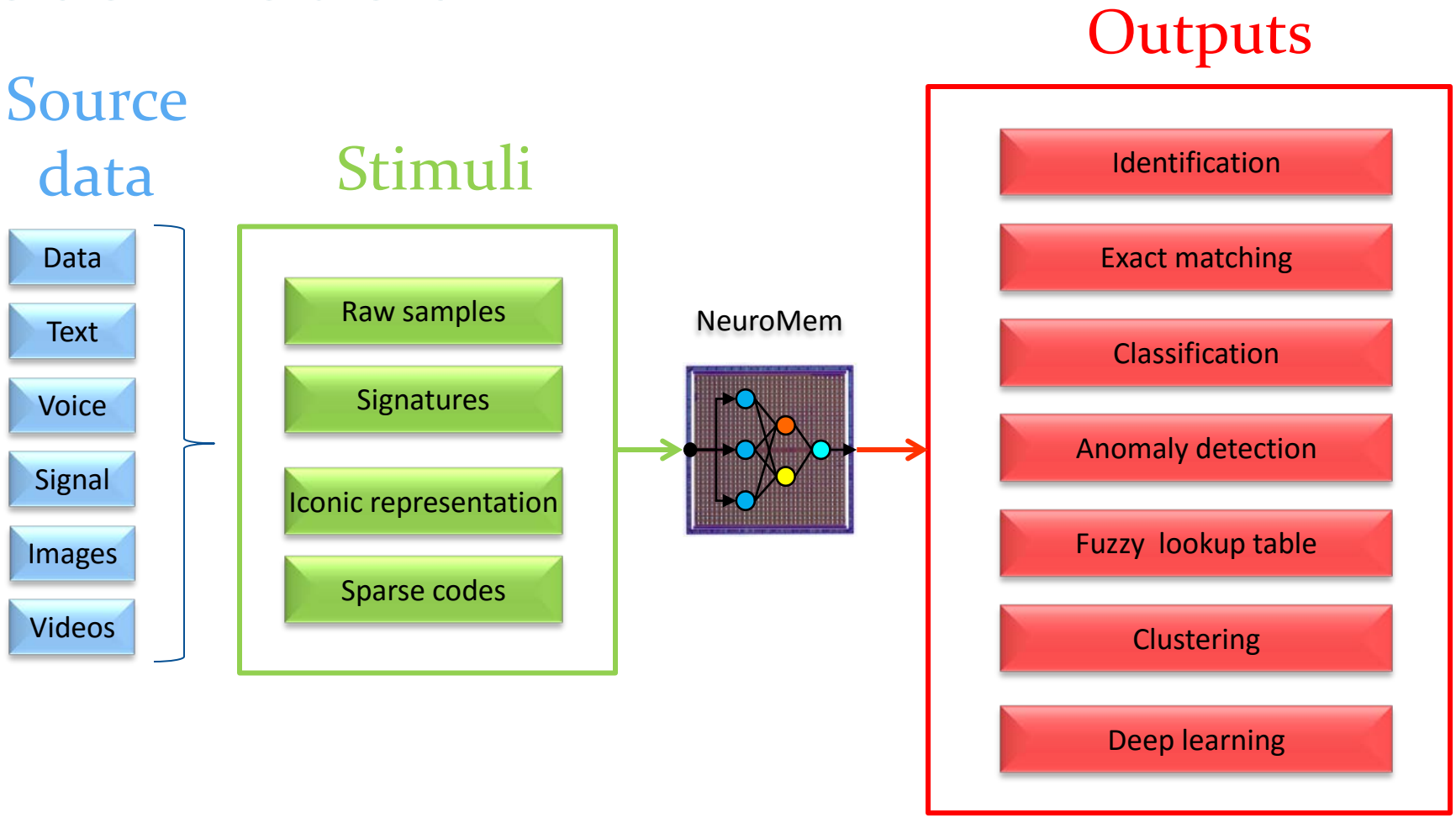


- Regular architecture, just neurons
- No fetch and decode
- Patented WTA bus (no cross bar)
- Low power (<0.5 watts)
- Self trainable
- Orthogonal inter-chip connectivity
- Commercially available (IC, Source and FPGA IP)

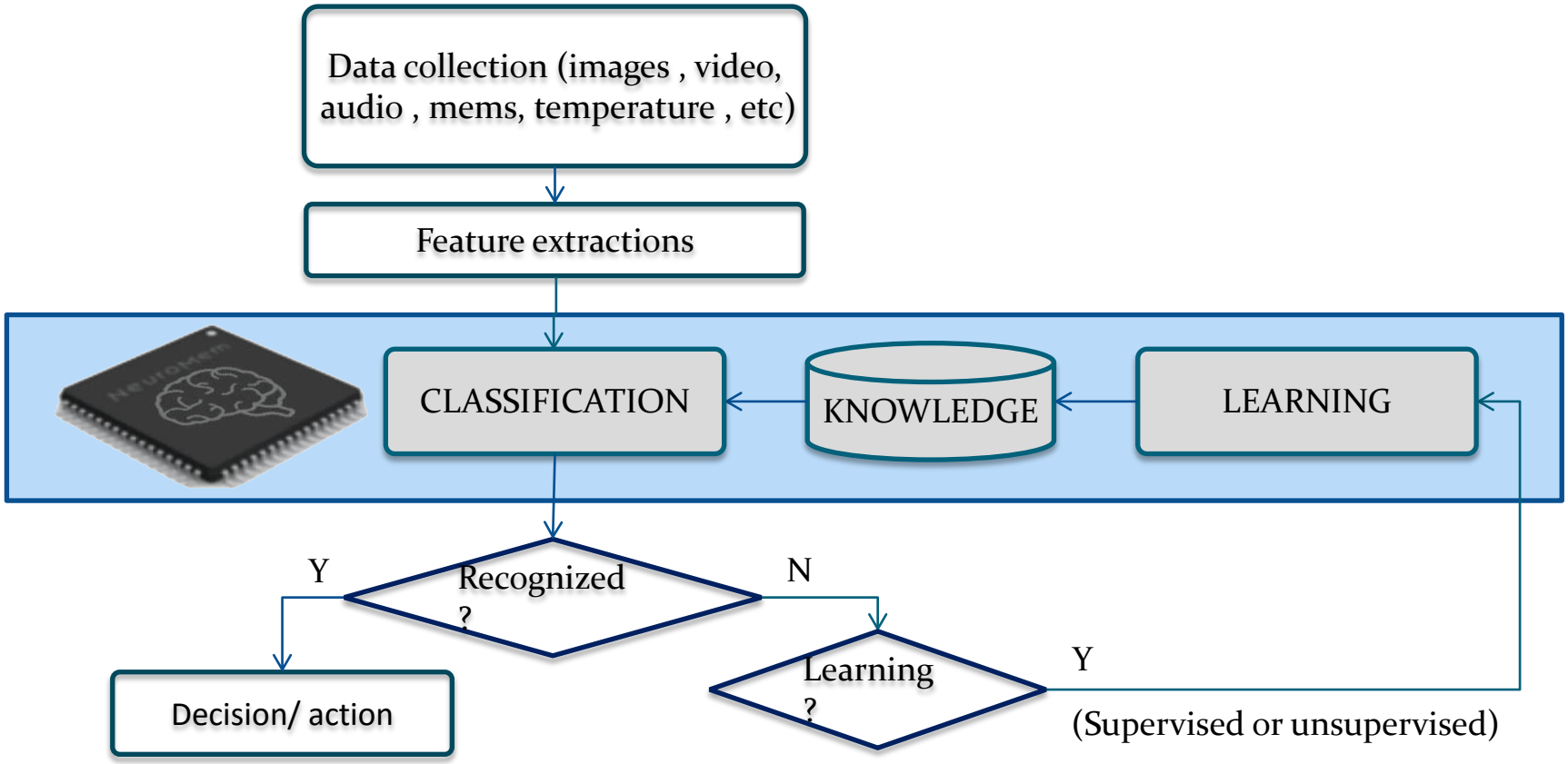
1024 identical neuromorphic memories, all interconnected (intra & inter hip)

GENERAL VISION

Use Models


Source data
Data
Text
Voice
Signal
Images
Videos
Stimuli
Raw samples
Signatures
Iconic representation
Sparse codes
NeuroMem
Outputs
Identification
Exact matching
Classification
Anomaly detection
Fuzzy lookup table
Clustering
Deep learning

Application Deployment



How many neurons do I need?

- Depends on the application, the variability of the data, etc. Examples:

Application	Description	Estimated Neurons /Object	Total neurons
Fish sorting	Classification of herrings (Accept, Reject, Recycle) passing on an in-line conveyor belt		≈200
Glass Inspection	Detection of anomalies of texture in patterned solar glass passing on a conveyor		≈800
Inkjet OCR	Reading of date code or serial numbers printed on a packaging	1-3 /digit	
Cooperative face recognition	Identifying a person facing front, positioned at a known distance of the camera, willing to remove her glasses if needed to be recognized	5/person	
Semantic analysis	Counting the occurrences of words from a dictionary in live tweets, posts and other text streams.	1/word	
Motion classification using MEMS	Intel Curie module is intended for motion classification and features 128 neurons		<128 ?

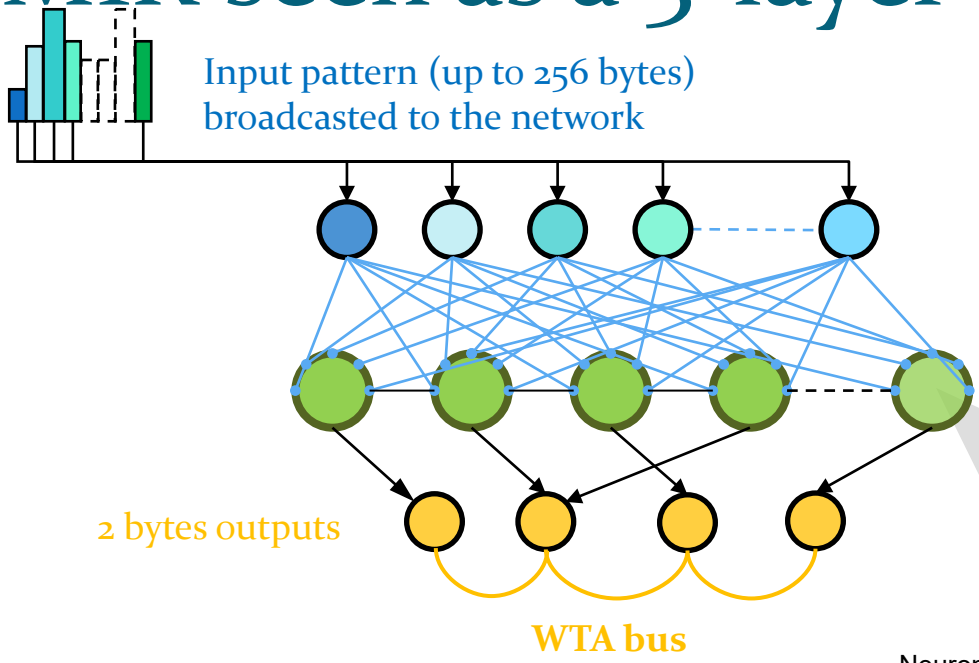
The pillars of Neuromorphic …

- **Broadcast Mode :** Queries/stimuli are broadcasted to all neurons simultaneously
- **Deterministic search time:** Firing time does not increase with the scaling of the network
- **Winner takes all:** Autonomous inhibition of the weak responders
- **Uncertain response:** Autonomous sorting of the responders in order of "lesser confidence"
- **Unknown response:** Awareness of the Unknown enabling the dynamic addition of new knowledge
- **Back propagation of error:** Autonomous inhibition of erroneous firing neurons
- **No fetch and decode of program instruction:** Software is definitively contrary to the biological model, else it's simulation, not neuromorphic…
- **Beyond biology:** Fast upload download enabling knowledge proliferation (some dream of it).

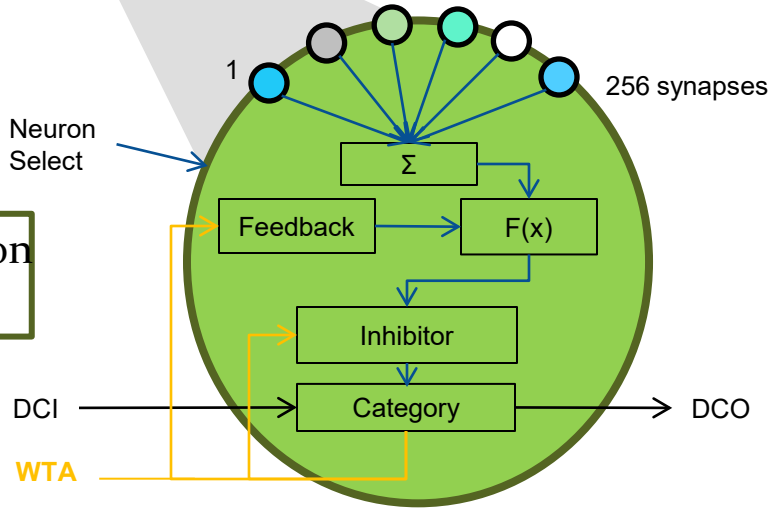
All implemented in the NeuroMem CM1K IC (continued…)

GENERAL VISION

CM1K seen as a 3-layer NN



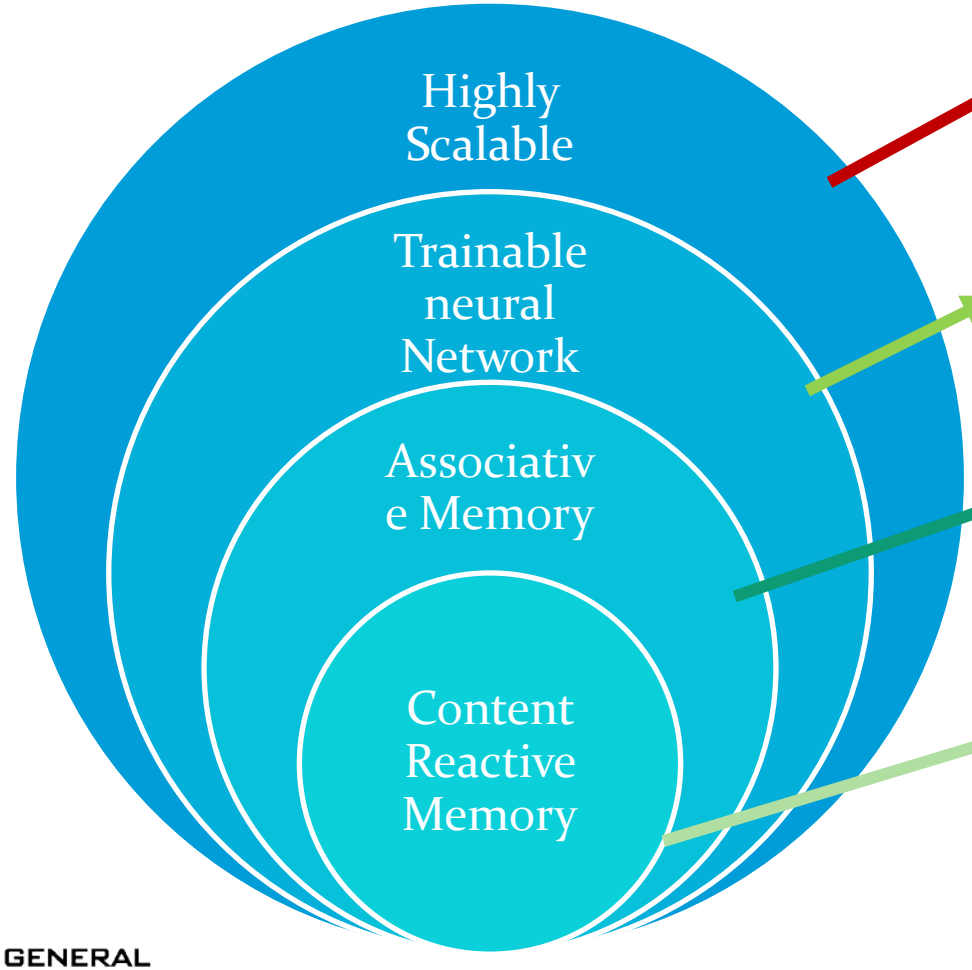
Single neuron view:
- 256 bytes RAM
- 3000 logic gates



CM1K= 1024 neurons with 256 synapses per neuron
= 262,144 synaptic connections of 8-bits

GENERAL VISION

The multiple functional facets of the NeuroMem memories



Interconnect with fixed pin-count and low frequency clock

On-the-chip learning

Non linear classifier, Exact or fuzzy matching (RBF/KNN)

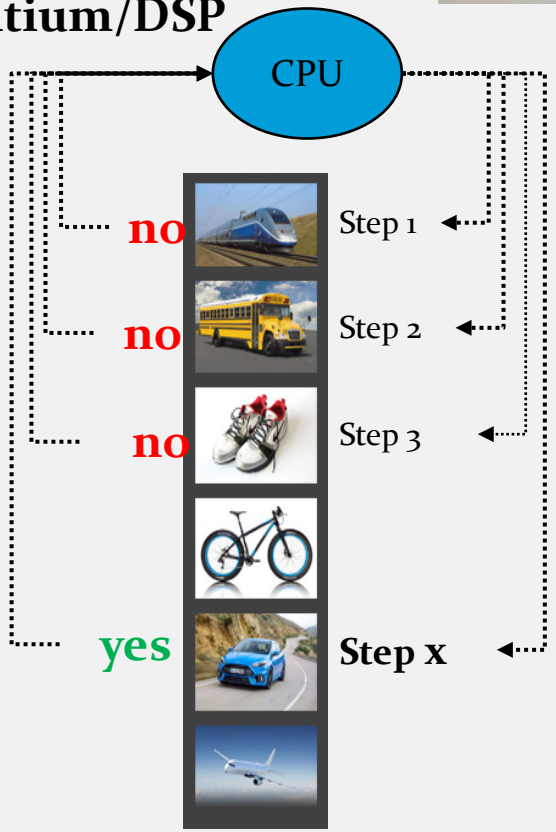
Broadcast mode and autonomous firing

GENERAL VISION

Speed Performance

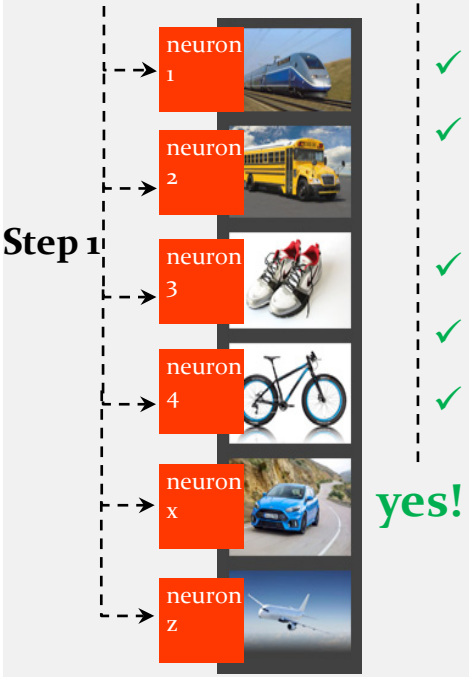
What's in this object?

Pentium/DSP



- Sequential search
- Increasing time
- High consumption (Ghz)
- Heavy software
- Complex scalability

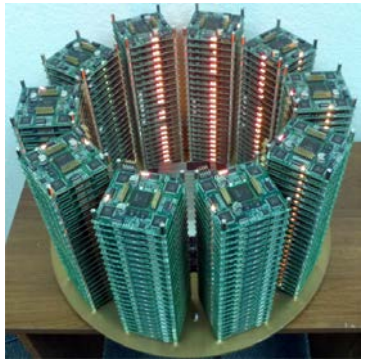
NeuroMem™

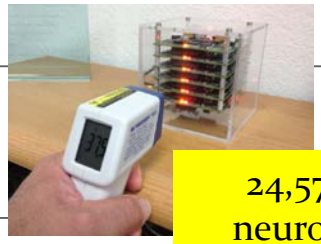


- ✓ Parallel search
- ✓ Constant search (ns)
- ✓ Low power (Mhz)
- ✓ No software
- ✓ Simple expansion

GENERAL VISION

Scalability and Power Consumption





40,960 neurons
20 Watts

1 Million neurons
500 Watts

24,576
neurons
12 Watts

4096 neurons
2Watts

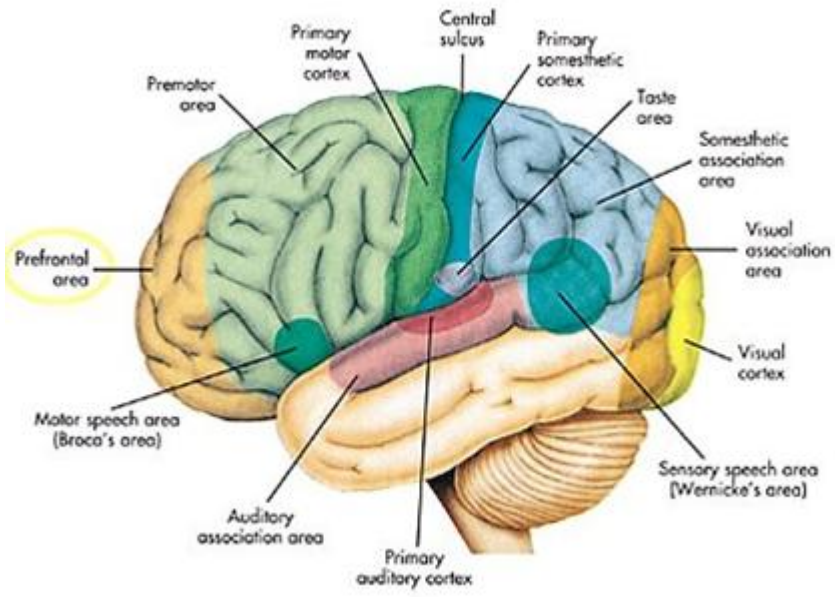
1024 neurons
0.5 Watts

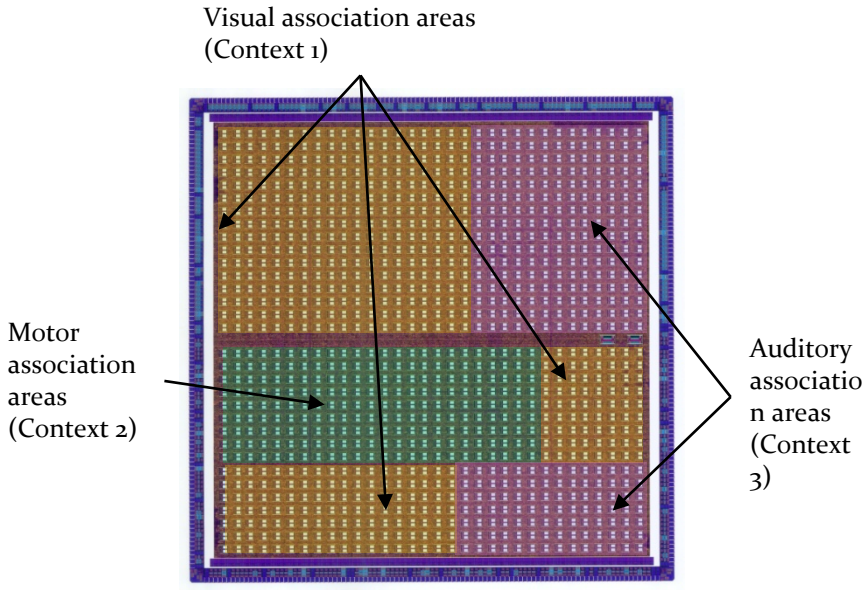
2.5 µsec
Recognition latency

GENERAL VISION

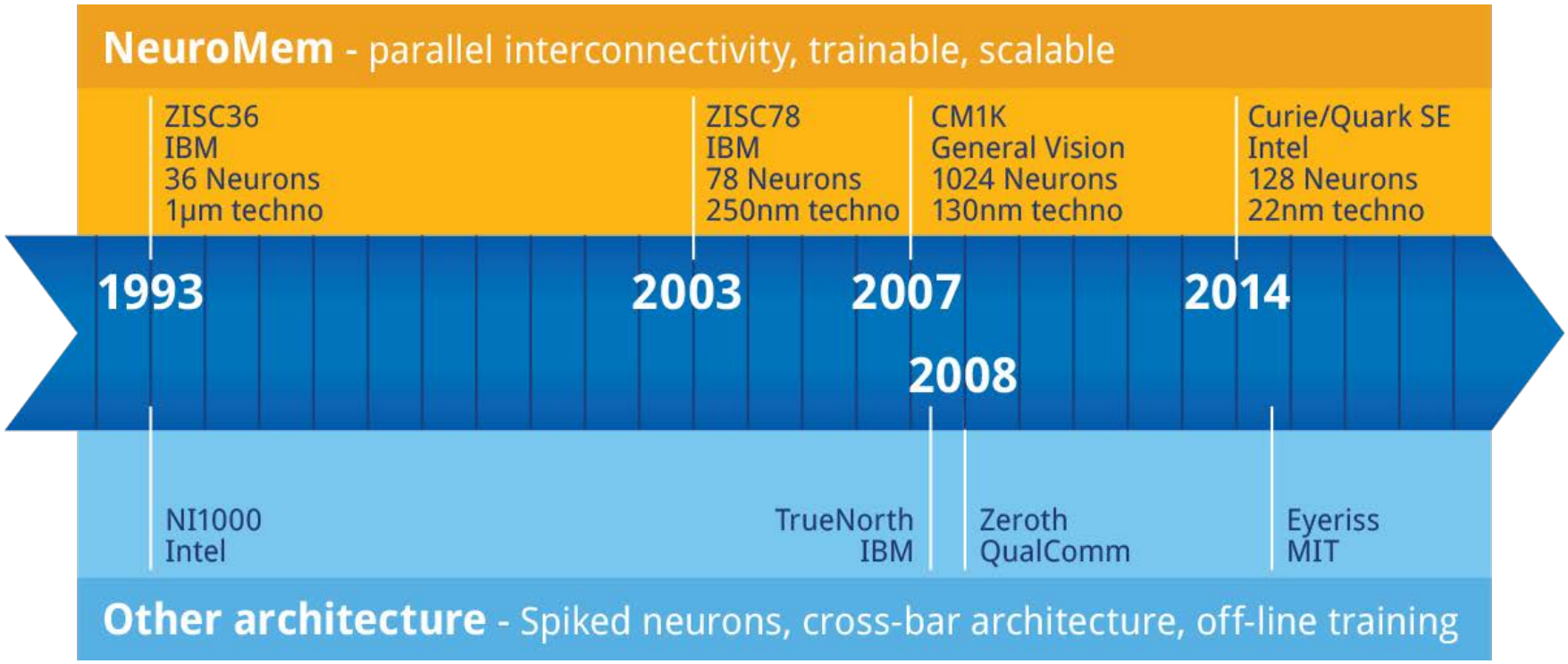
Context segmentation

Neurons can be assigned to different contexts for Sensor fusion, Feature fusion and robust decision





The Genesis


NeuroMem - parallel interconnectivity, trainable, scalable
ZISC36
IBM
36 Neurons
1µm techno
ZISC78
IBM
78 Neurons
250nm techno
CM1K
General Vision
1024 Neurons
130nm techno
Curie/Quark SE
Intel
128 Neurons
22nm techno
1993
2003
2007
2008
2014
NI1000
Intel
TrueNorth
IBM
Zeroth
QualComm
Eyeriss
MIT
Other architecture - Spiked neurons, cross-bar architecture, off-line training