

Computational neuroscience offers hints for more general machine learning

David Rawlinson and Gideon Kowadlo

Motivation



Project AGI agi.io

1. Which parts of Machine Learning are moving towards AGI?

2. Are Machine Learning approaches biologically plausible?

3. If not, then can neuroscience help us towards AGI?

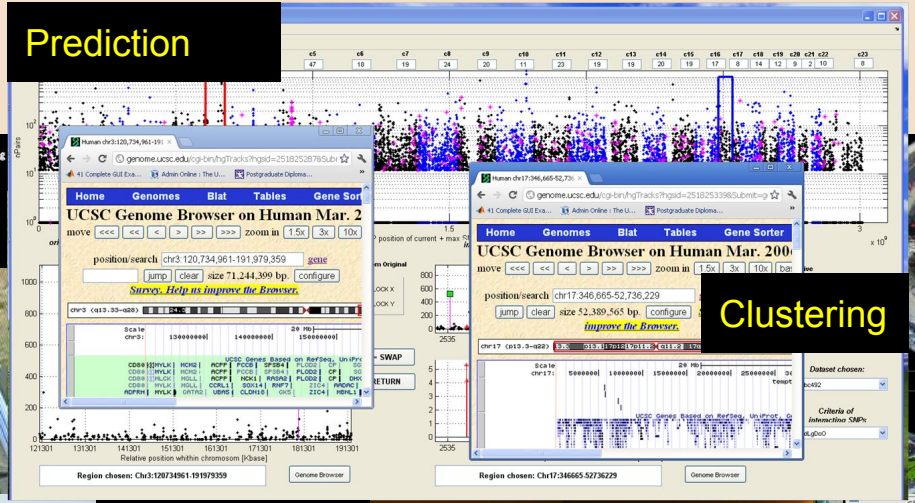
Trends

Project AGI agi.io



OCR

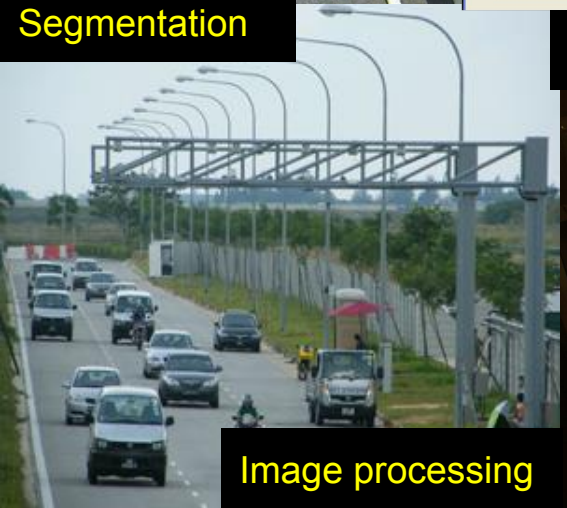
Prediction



Clustering

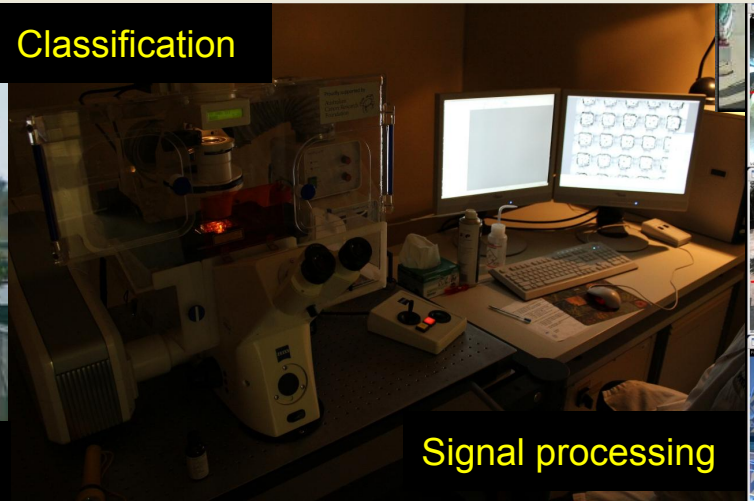


Estimation & control



Segmentation

Classification



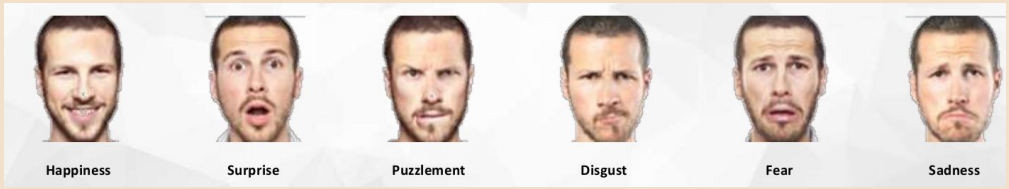


Target tracking

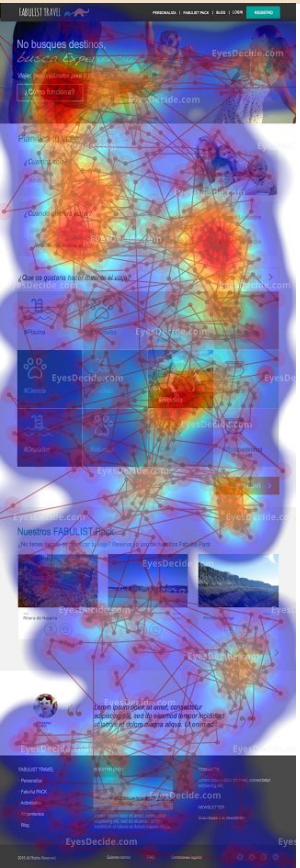
Image processing

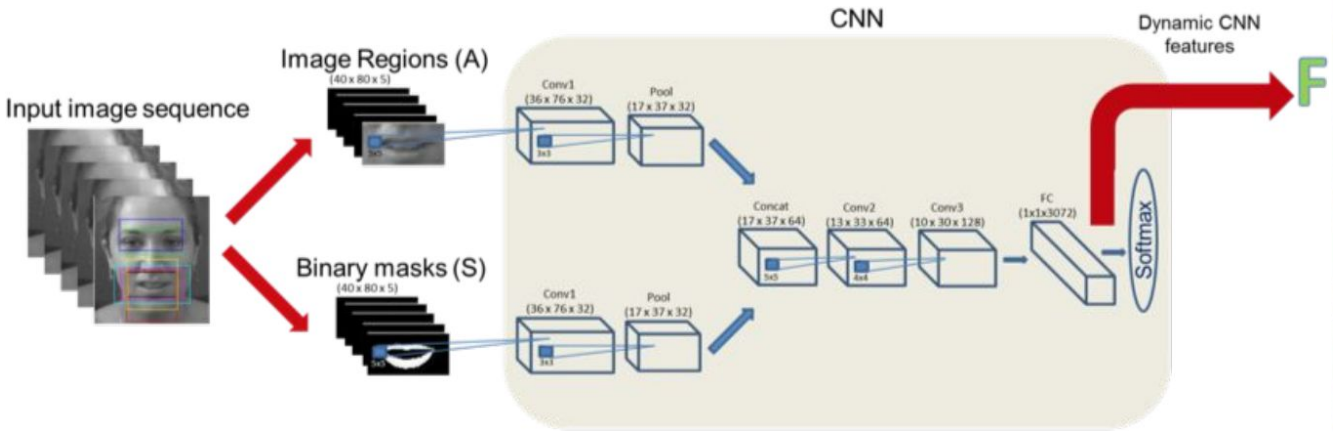
Signal processing

Now



Project AGI agi.io





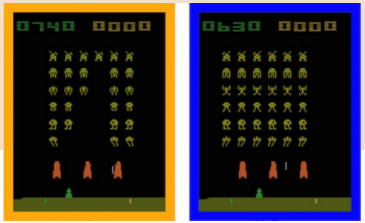
** Jaiswal, Shashank and Valstar, Michel F. (2016) Deep learning the dynamic appearance and shape of facial action units. In: Winter Conference on Applications of Computer Vision (WACV)*

EyesDecide

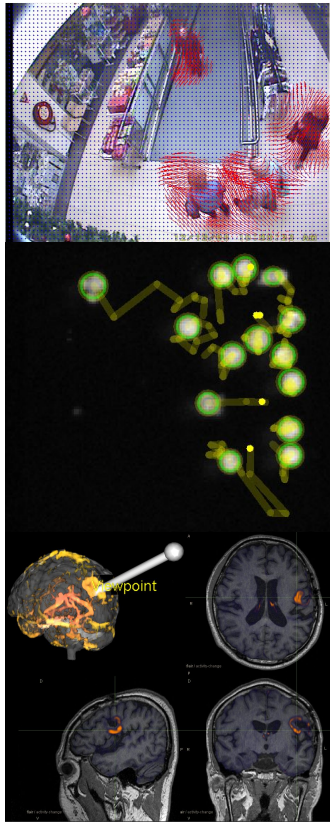
crowdemotion

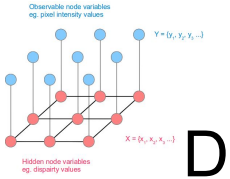


The trends are good



Project AGI agi.io





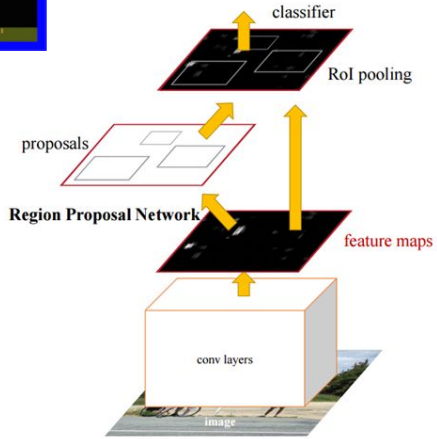
Designed → Learned

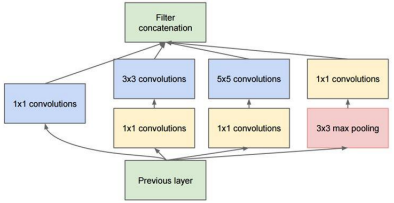
Heterogeneous → Homogeneous

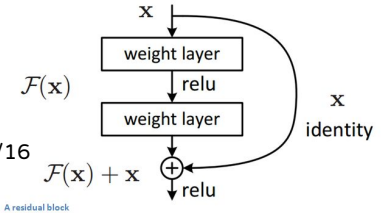
Layers → Modules, circuits

Shallow → Deep

Classification → Agents







"Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" Ren et al 2015/16
"Going Deeper with Convolutions" Szegedy et al CVPR 2015
"Deep Residual Learning for Image Recognition" He et al arXiv 2015

Dilemmas



Project AGI agi.io

- Backpropagation/SGD is common approach, but…
  - Vanishing/shattered gradients; invariances; generalization; adversarial examples
- Not the only way:
  - Deep error backpropagation not shown to be biologically plausible
- Unsupervised pre-training improves generalization via regularization
  - Detects statistically most significant features, but
  - Fails to capture rare but diagnostic features that are important for decision making

Comparison



Project AGI agi.io

ANN	Wet brain
Slow, supervised learning	Slower and faster learning! Lifelong, online, unsupervised and reinforcement learning
Needs rich, specific feedback	Rare & abstract feedback
Stationary problems (mostly)	Non-stationary problems

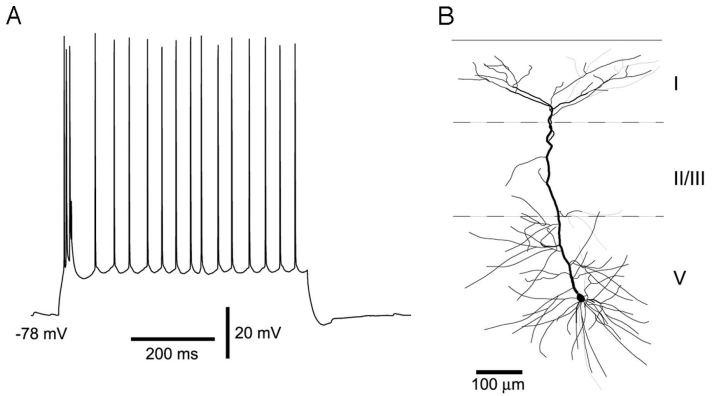
Hints



Encoding

- Spiking, recurrent neural networks
- Rate Coding (orthodoxy, inadequate)
- Temporal Coding (timing is important)
- Population Coding
- **Sparse Coding***

Does it matter? Well, it seems to help…



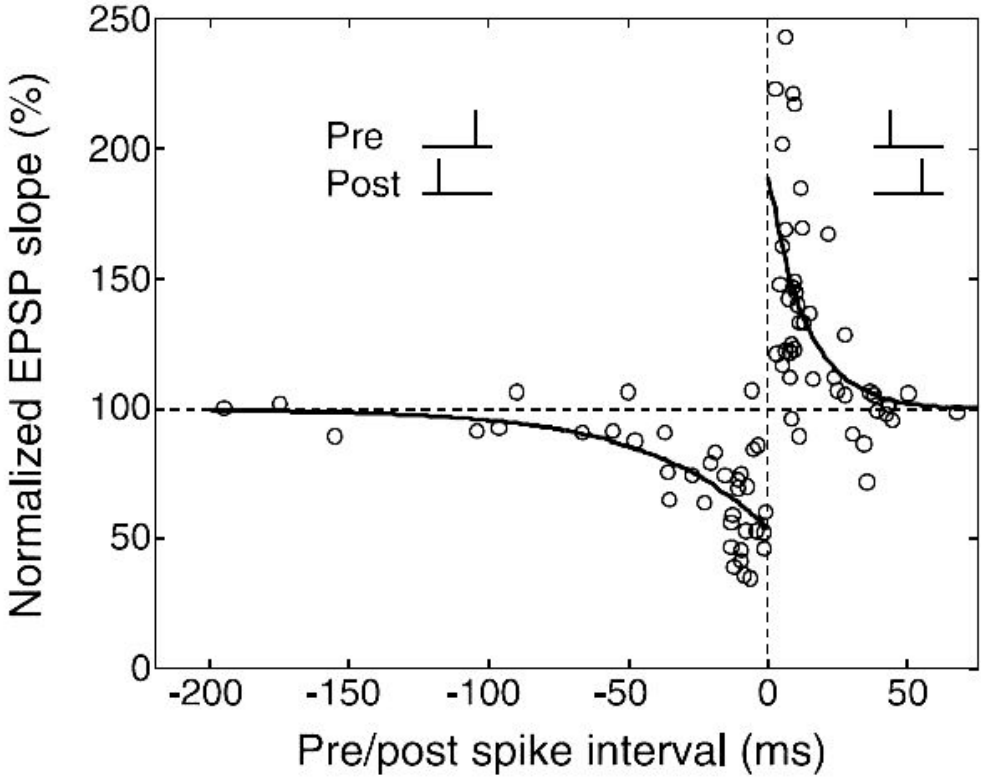
* “Building high-level features using large scale unsupervised learning” Le et al ICML 2012

Learning rules
(real world doesn't have labels)

Spike Timing Dependent Plasticity:

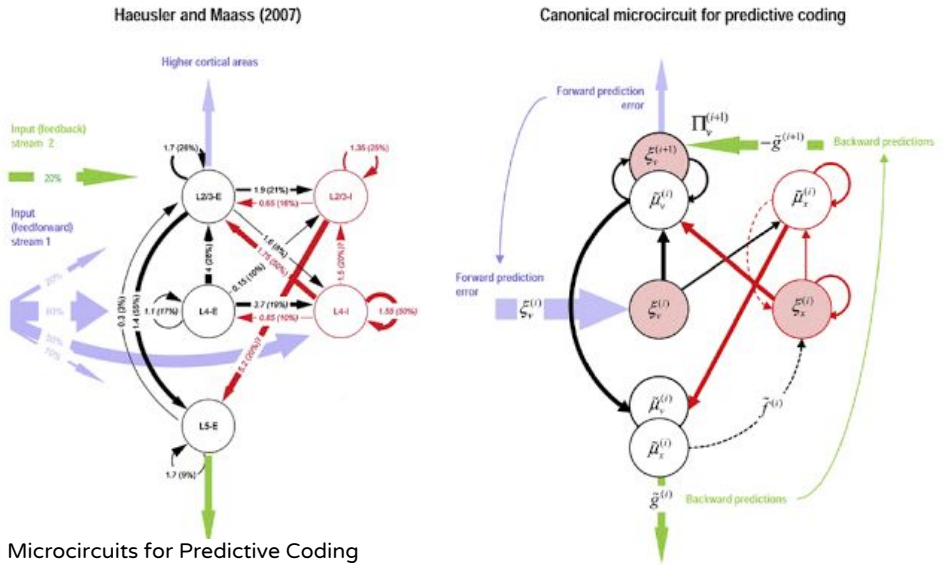
- An unsupervised local learning rule
- Repeatedly confirmed by direct observation in vitro

(with exceptions of course…)

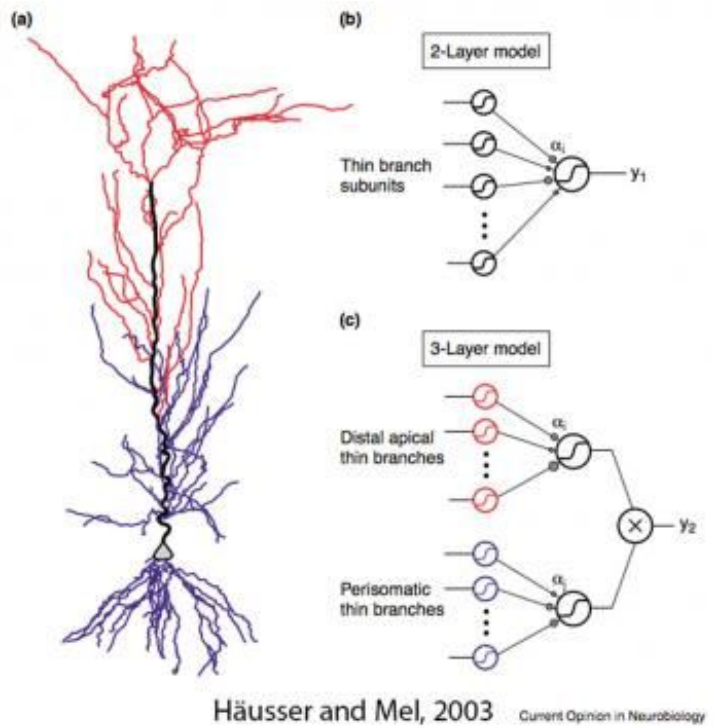


Microcircuitry

- Dendrite integration & computation
- Recurrently connected microcircuits
- Predictive coding



Canonical Microcircuits for Predictive Coding
Bastos, AM. Usrey, WM. Adams, RA. Mangun, GR. Fries, P. Friston, KJ. (2012)
Neuron 76: 695-711.

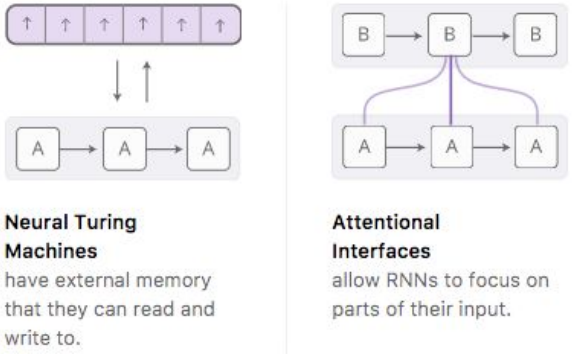


FN Error coding: "Principles of dendritic integration in CA1 pyramidal neurons" Spruston et al

Predictions

- Major strides in Reinforcement learning
  Example: reward value distribution
- Possible revolution in representation. Better solutions for non-stationary problems; local and unsupervised learning
- AGI architectures with RNNs: Attention, working memory, synaptic consolidation
- Agent decision-making: Exploration/exploitation, self-teaching
- Continuous, online, lifelong learning

"A Distributional Perspective on Reinforcement Learning"
Bellemere et al arXiv 2017



Neural Turing Machines
have external memory that they can read and write to.

Attentional Interfaces
allow RNNs to focus on parts of their input.

https://distill.pub/2016/augmented-rnns/#neural-turing-machines

The End.



Project AGI agi.io

Go back.