HELSINKI UNIVERSITY OF TECHNOLOGY
Laboratory of Computational Engineering

COMPUTATIONAL MODEL OF CO-OPERATING COVERT ATTENTION AND LEARNING

ANTTI YLI-KREKOLA AND HARRI VALPOLA
Computational Neuroscience Research Group
www.lce.hut.fi/research/eas/compneuro/

Introduction

Attention is known to guide learning in the real neocortex. Selective attention should also make the learning task easier on purely theoretical grounds. Still, in state of the art computational models, these two mechanisms have not been integrated together. We developed a system-level model of the neocortex, where learning and attention support each other and improve each other's performance [1]. Later on, we intend to integrate the model into a larger cognitive architecture.

Neuroscience

Learning feature representations in the neocortex

- Hierarchical organization in both motor and sensory domains.
- The highest levels represent different forms of invariances, abstract and long time-scale targets.
- Adaptation continues throughout lifetime in all levels of the hierarchies [2].
- Learning happens mostly on attended targets [3].

Biased-competition model of attention

- Adjacent neurons compete with each other.
- Long distance excitatory connections bias this competition [4].

Covert attention emerges without any specific module controlling it. There is neurophysiological evidence for the model [5] and computational models have shown that local competition can produce global attention [6].

The Model

The mechanism

- Bottom-up feature activations are biased with lateral associations, after which competition selects the most important features.
- Higher layers have fewer neurons, which leads to invariances
- Lateral connections from motor cortex guide the visual features to be motorically useful.

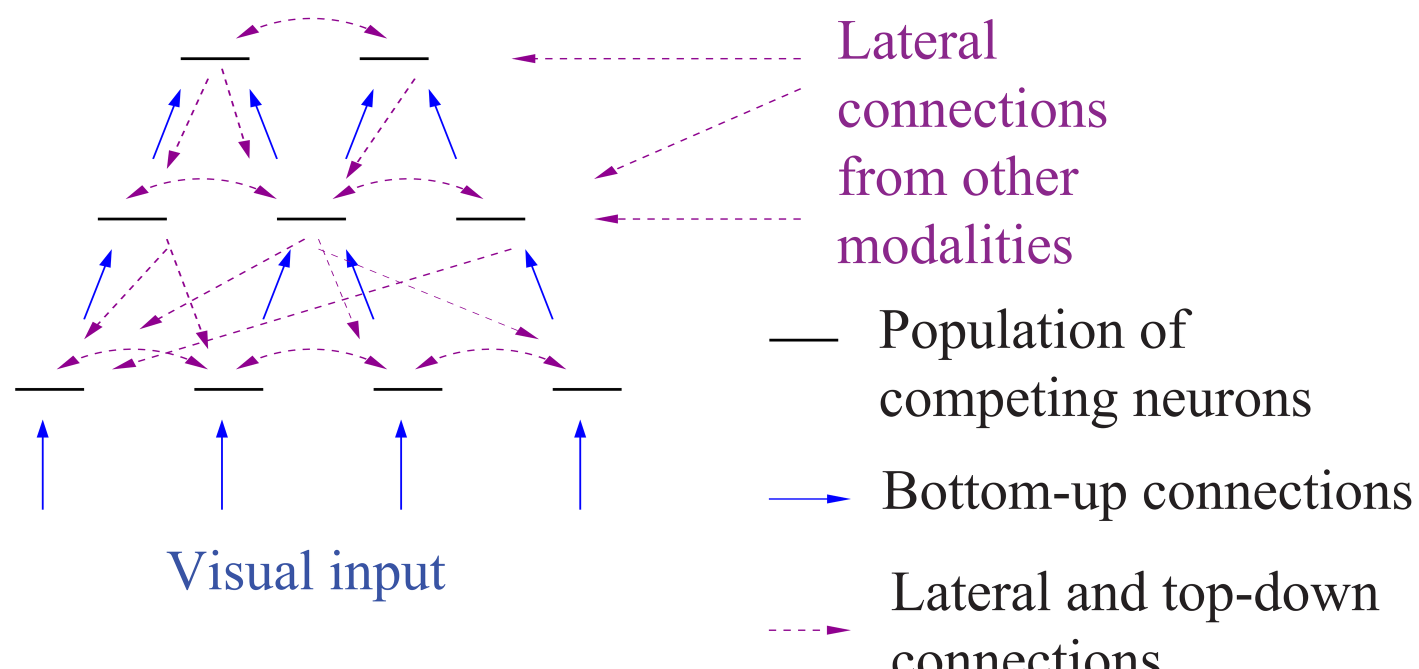

Figure 1, example architecture of the model

Consequences

Lateral associations result in corresponding features winning the local competitions and global coherent attention emerges.

If attention succeeds to select one behaviourally important target:

- **Learning leads to behaviourally important representations**
- **Learning becomes easier because of decluttering the inputs**

Experiments

The data and the network

The network received 20x20 pixel images as inputs. Each image is an overlap of two different instantiations from six object classes (Fig. 2).

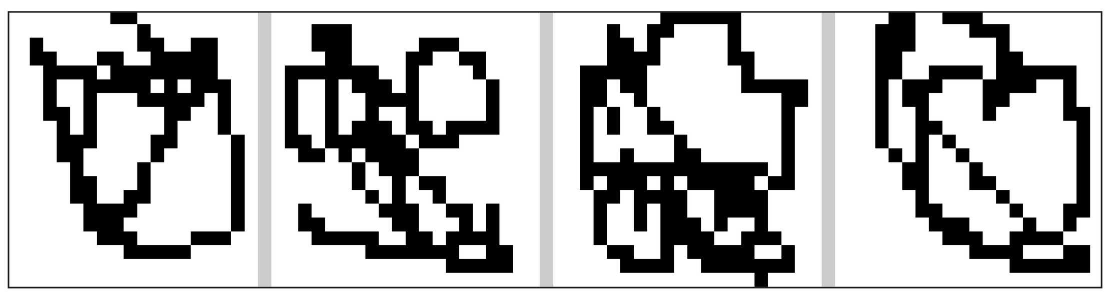

Figure 2, example input images

The network had four layers with 1120, 112, 280 and 24 neurons, respectively.

Invariant representations

The network learned individual objects from the cluttered images. The highest (fourth) layer developed representations that are invariant to object transformations. All images in the top row of Fig. 3 activated the same population code in layer four.

The process can also be reverted, so that the highest layer generates expectations of the inputs through top-down connections. Depending on the states of lower level neurons, different transformations of the object will be generated (Fig. 3).

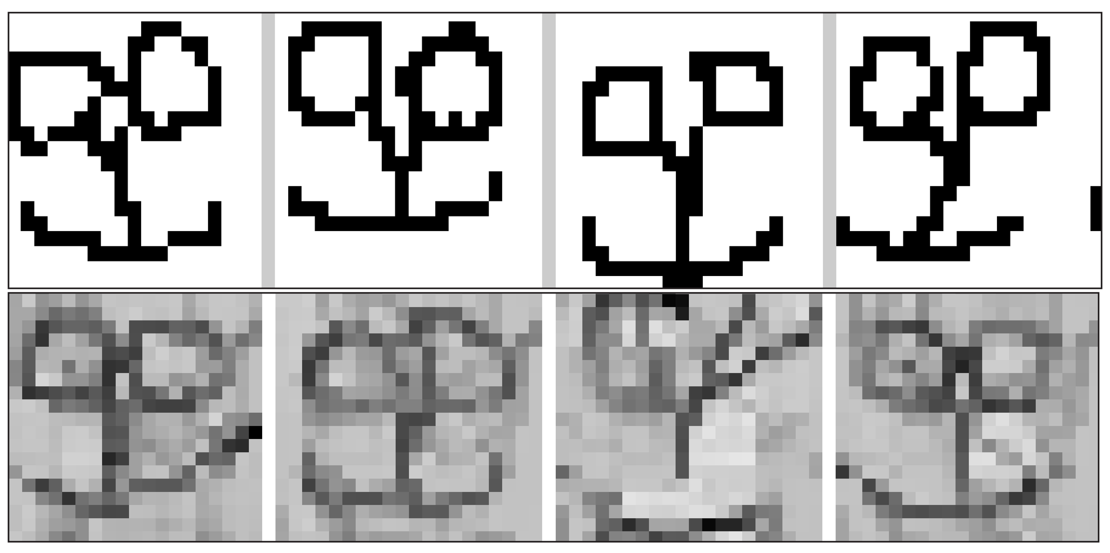

Figure 3, top row: images, which all activated the same representation in the highest layer, bottom row: predictions generated by the highest layer

Jumping attention

When the input was constant (Fig. 4), but the neurons have a habituation property, attention will switch its focus constantly (Fig. 5 and Fig. 6)

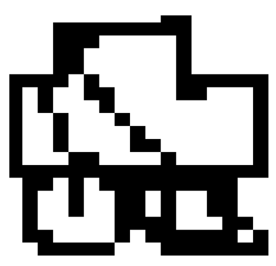

Figure 4

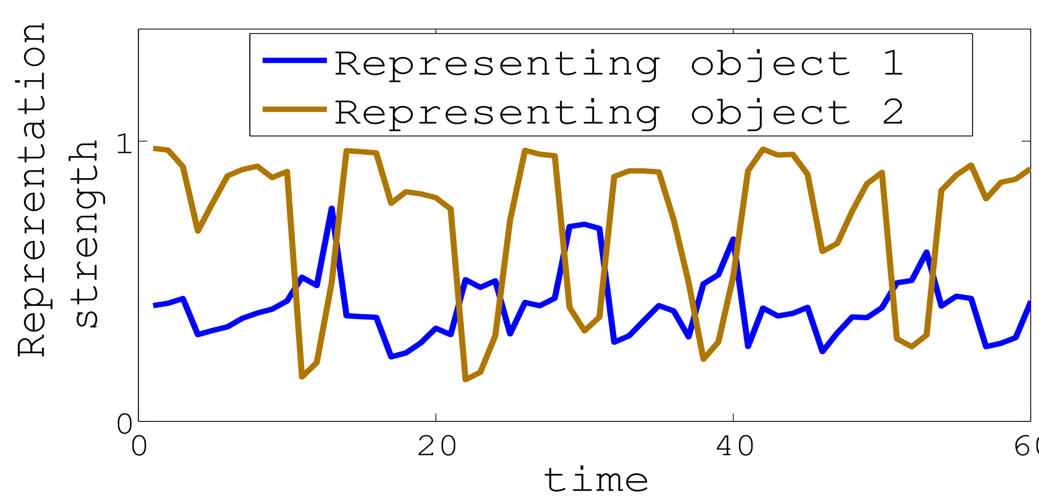

Figure 5, Activations on the highest layer on different time steps are compared to individual object representations.

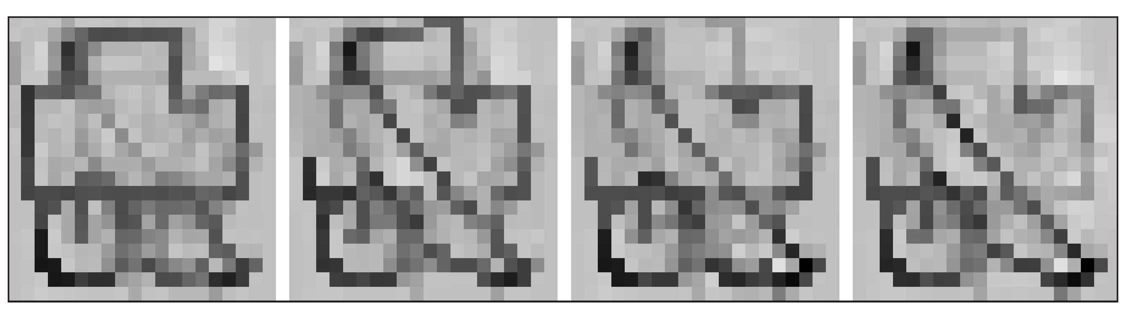

Figure 6, What the first layer (V1) sees. Images are generated with first layer's top-down weights on different time instants.

References

[1] Yli-Krekola, A. (2007). *Master's Thesis.*
[2] Gilbert, C. D. (1994). *PNAS*, 91:1195-1197.
[3] Ahissar, M. and Hochstein, S. (1993). *PNAS* , 90:5718-5722.
[4] Desimone, R. and Duncan, J. (1995). *Ann. Rev. Neurosci.*, 18:193-222.
[5] Reynolds, J. H., Chelazzi, L., and Desimone, R. (1999). *J. Neurosci.*, 19:1736-1753.
[6] Deco, G. and Rolls, E. T. (2004). *Vis. Res.*, 44:621-642.