AIML Contest 2016

Juyang Weng, General Chair
Juan Castro-Garcia, Program Chair
Zejia Zheng, Technician, Vision and Motor
Juan Castro-Garcia, Technician, Natural Language and Motor
Xiang Wu, Technician, Audition and Motor

Organizer:
Brain-Mind Institute, USA
Sponsors:
GENISMAMA LLC, USA and
INNS, USA

Many AI Competitions

IBM Watson

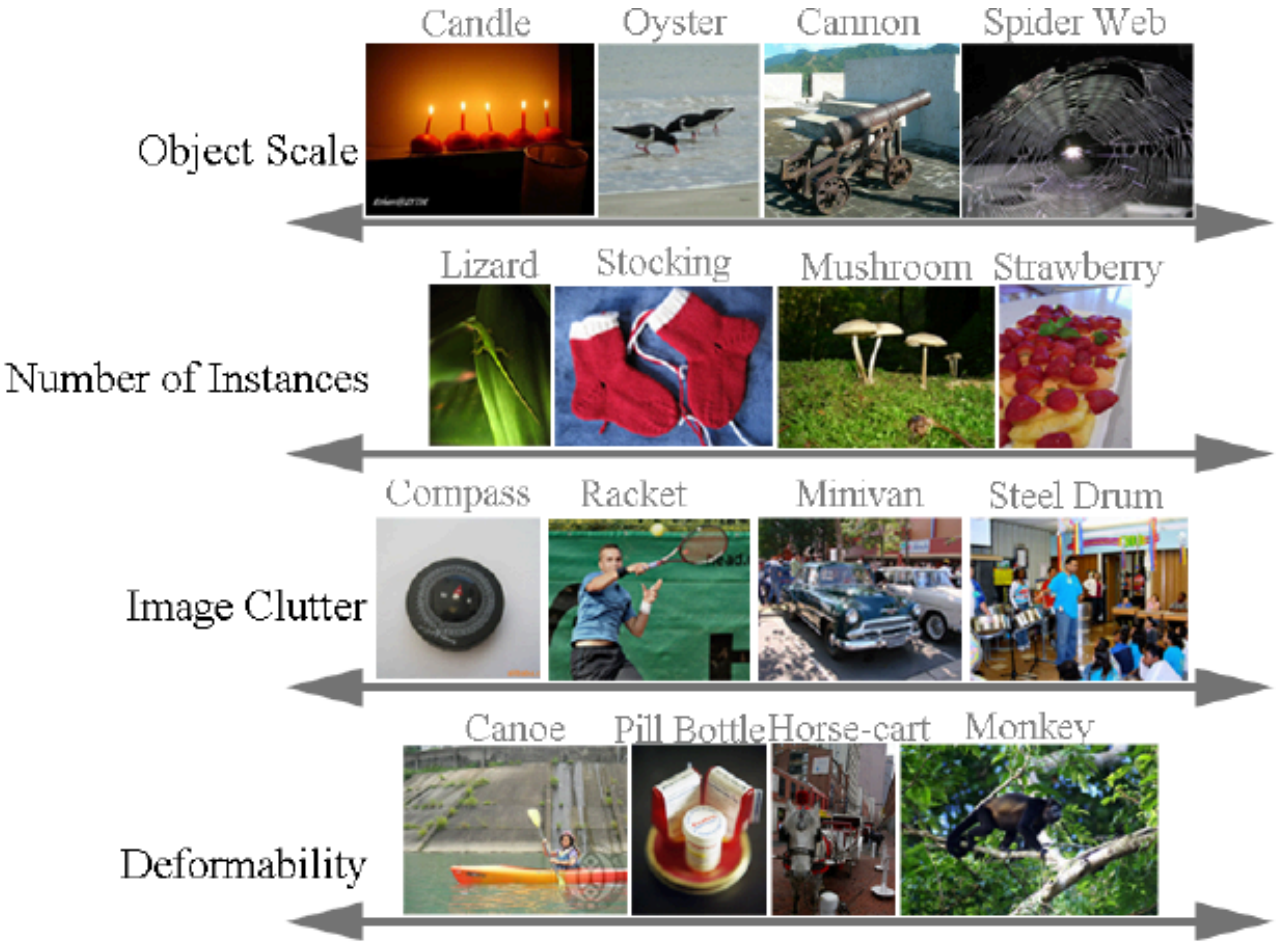
Google

kaggle

IMAGENET

What are the Purposes of a Competition?

ImageNet: Pattern Recognition and Falsification



F-F Li et al. IJCV 2015

ImageNet Contest is a Trap

- Static: Overwork on a hand-crafted static labels
- Plagiarism: plagiarized Cresceptron, whose key new techniques are obsolete (left-to-right scan, large-to-small scan, reduced resolution through a deep cascade, Max-pooling) by the WWN-1 to WWN-9 of DN
  Why ? *No attention. Not sensorimotor representations*
- Falsification: Manually removing data
  - Cal Tech 101 started such academic dishonesty in Comp. Vis.
  - ImageNet has spread widely the dishonesty to the community
- Read more evidence: Juyang Weng: Facebook, 科学网

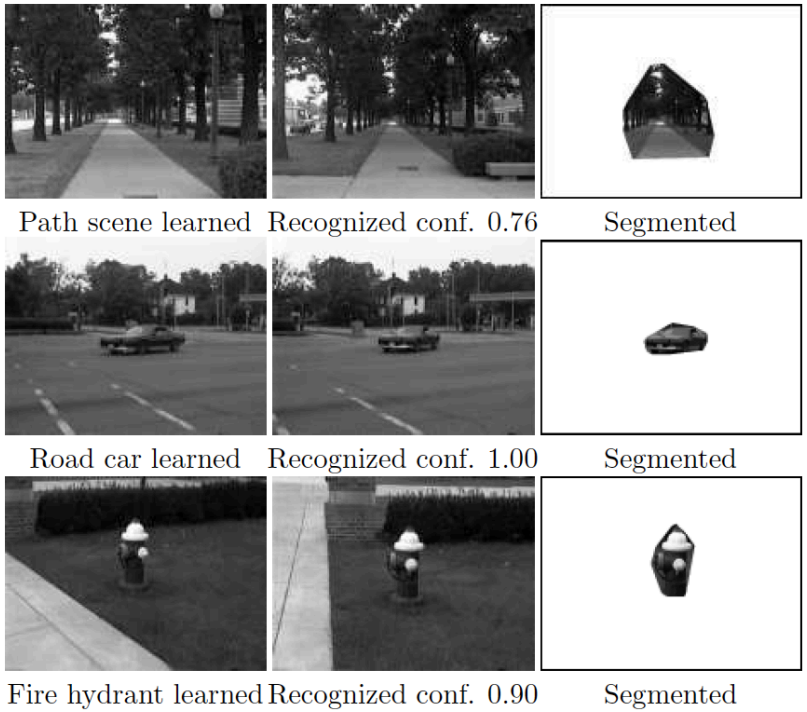
Purposes of the AIML Contest

1. For fun, like games
2. Train researchers and students
3. Add additional mechanisms to be a winner
4. Verification of methods by independent labs
5. Change of the established paradigms
6. Provide a developmental path towards a practical goal

Celebrities Plagiarized Cresceptron

- Cresceptron by Weng, Ahuja, Huang
  IJCNN 1991, ICCV 1992, IJCV 1997
- Tomaso Poggio from Nature Neuroscience 1999,
  HMAX PAMI 2007
- Li Fei-Fei from her PhD thesis 2005 at Cal Tech and her ImageNet pubs
- Yann LeCun from NIPS 2005
- Andrew Ng from ICML 2009
- Geoffrey E. Hinton from NIPS 2012
- Many in ImageNet
- More … (e.g., a connectionists@cmu manager)

Cresceptron:
Learning 3D objects from 2D Images without a 3-D model

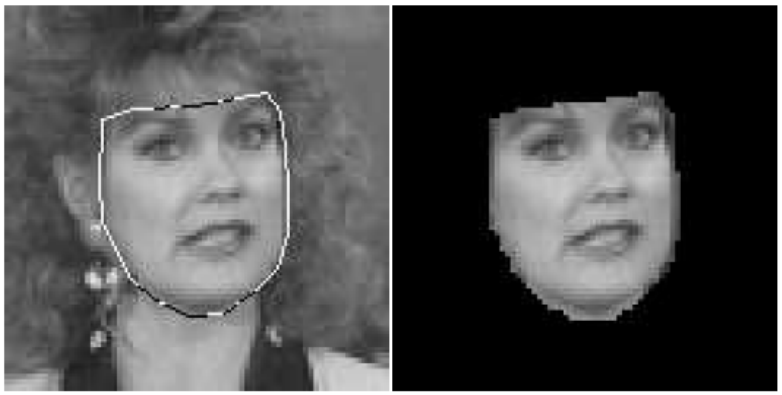


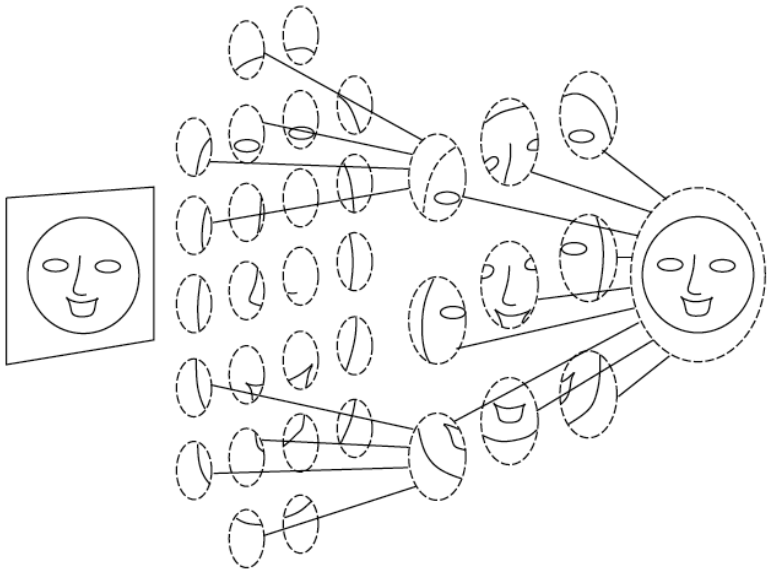
First 3D Visual Learning Without a Given Model: Cresceptron

- The 1st deep learning network that adapts its connection structure.
- The 1st visual learning program for both detecting and recognizing general objects from cluttered complex natural background.
- Also did segmentation from learning, but in a separate top-down segmentation phase.
- The 1st that proposed what is now called max-pooling

Cresceptron: Deep Learning

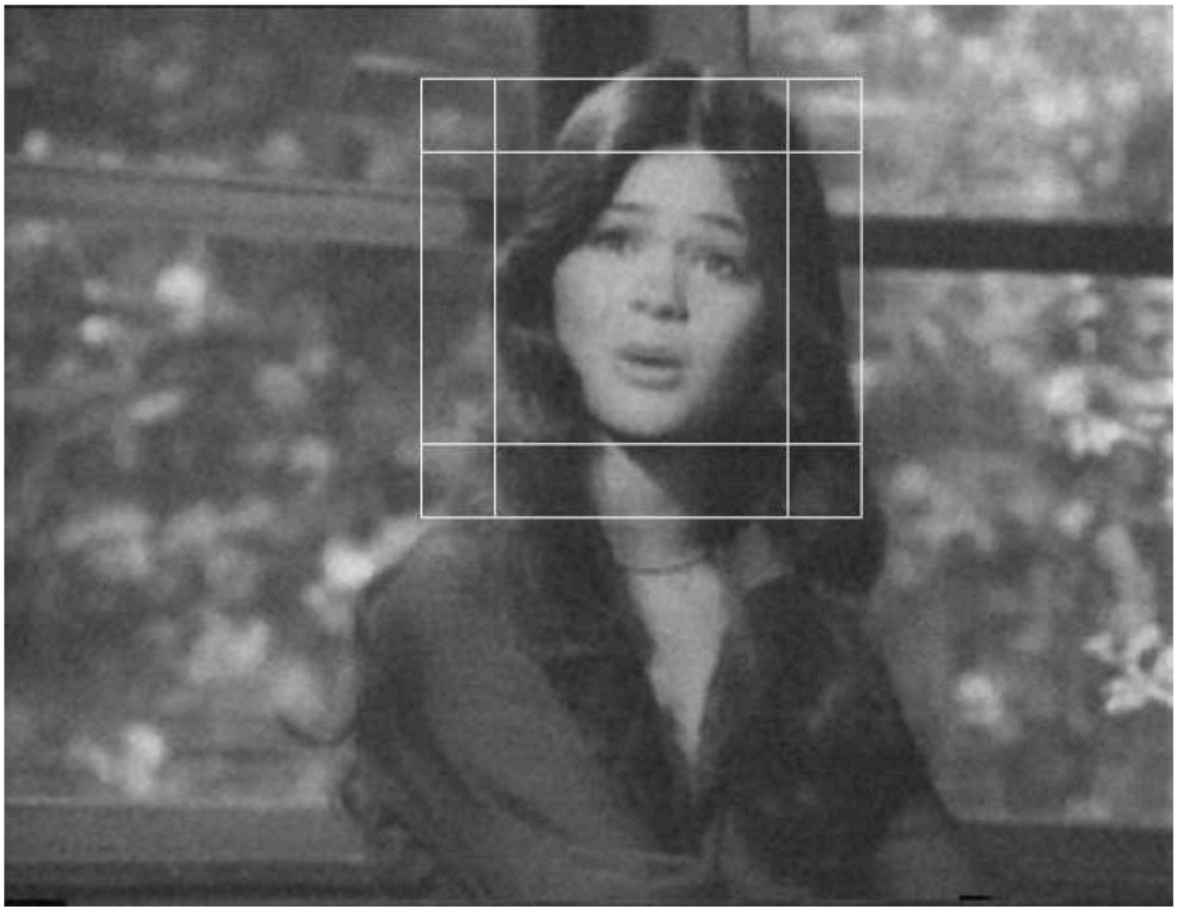
- The first Deep Learning Network for General Cluttered Scenes
- The first to proposed max-pulling



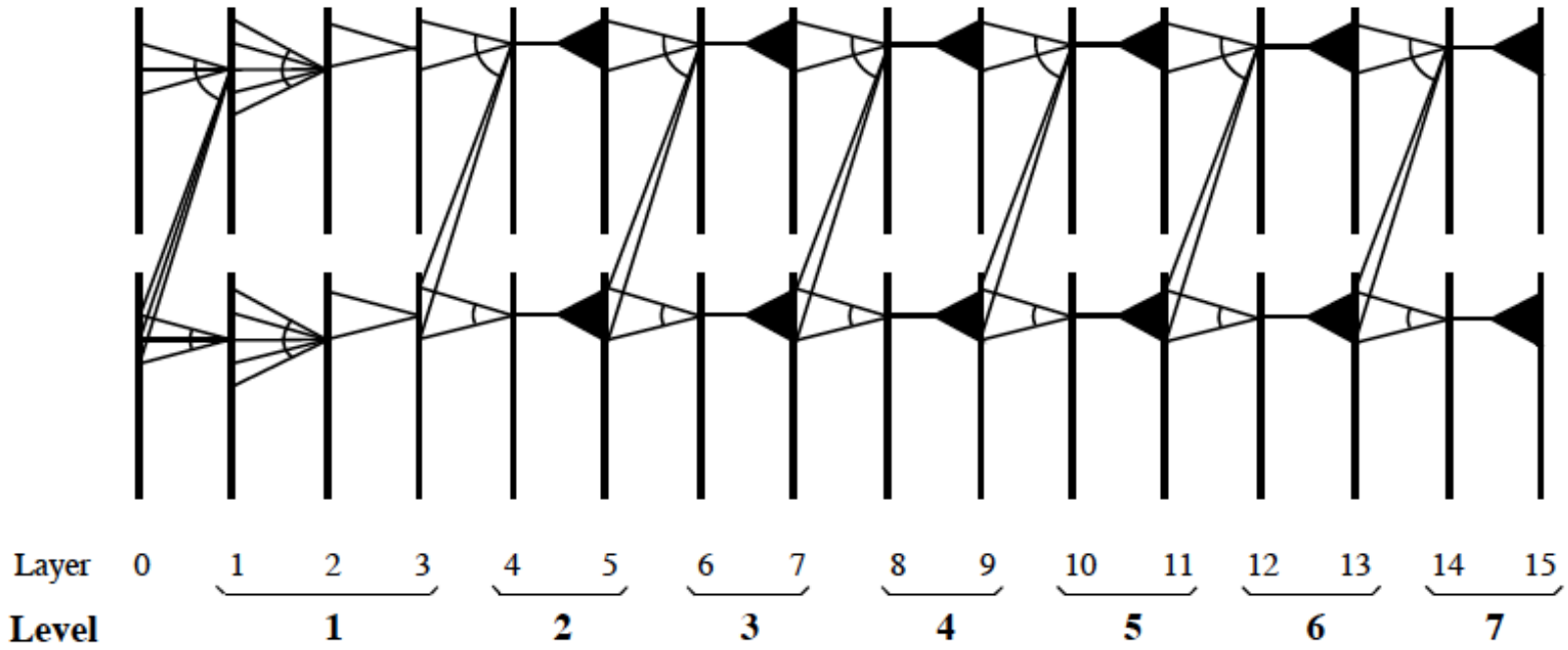


Mechanical Scan: Locations and Scales

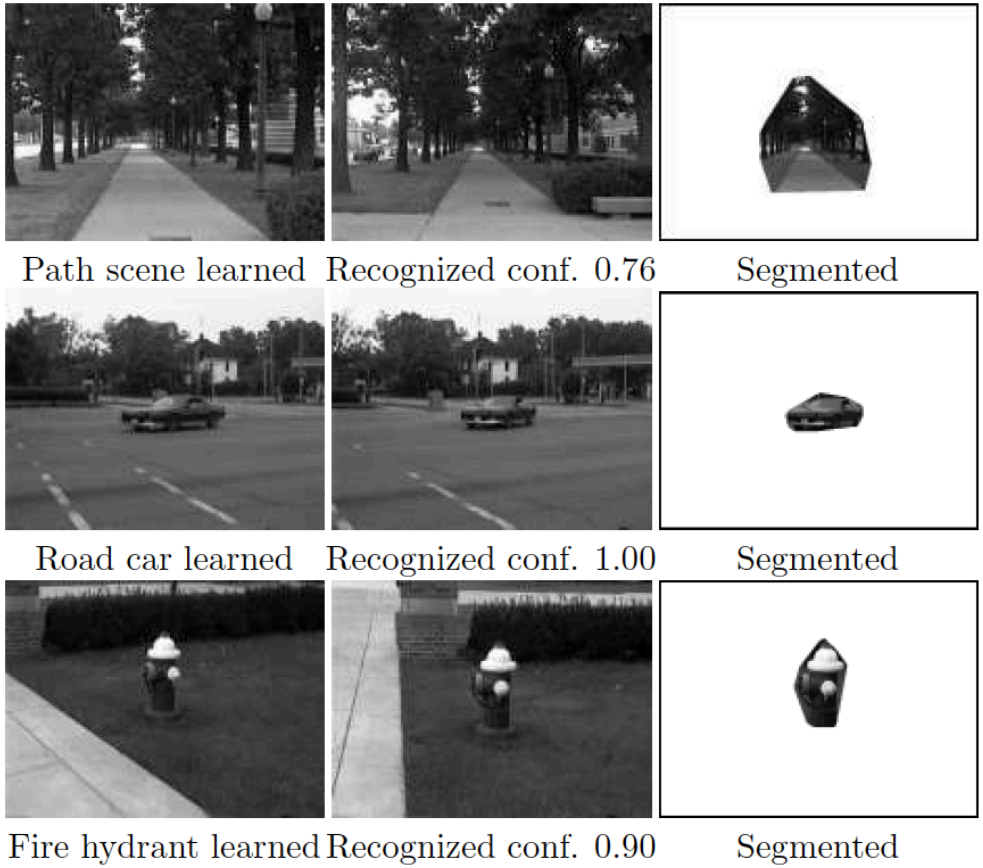
Duplicated Windows



Deep Learning: A Cascade



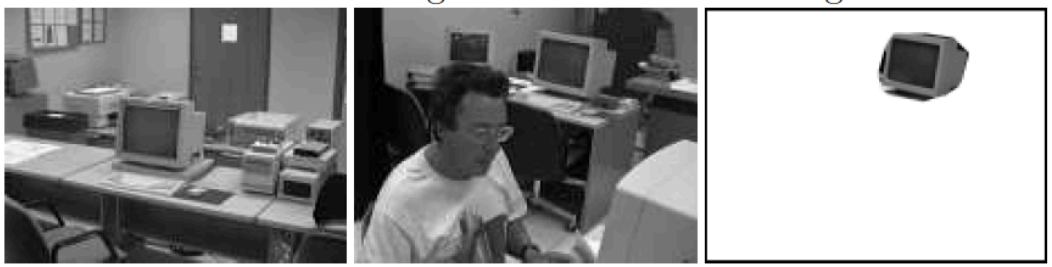
Detect, Recognize, and Segment (1)



Detect, Recognize, and Segment (2)

Telephone learned Recognized conf. 0.68 Segmented

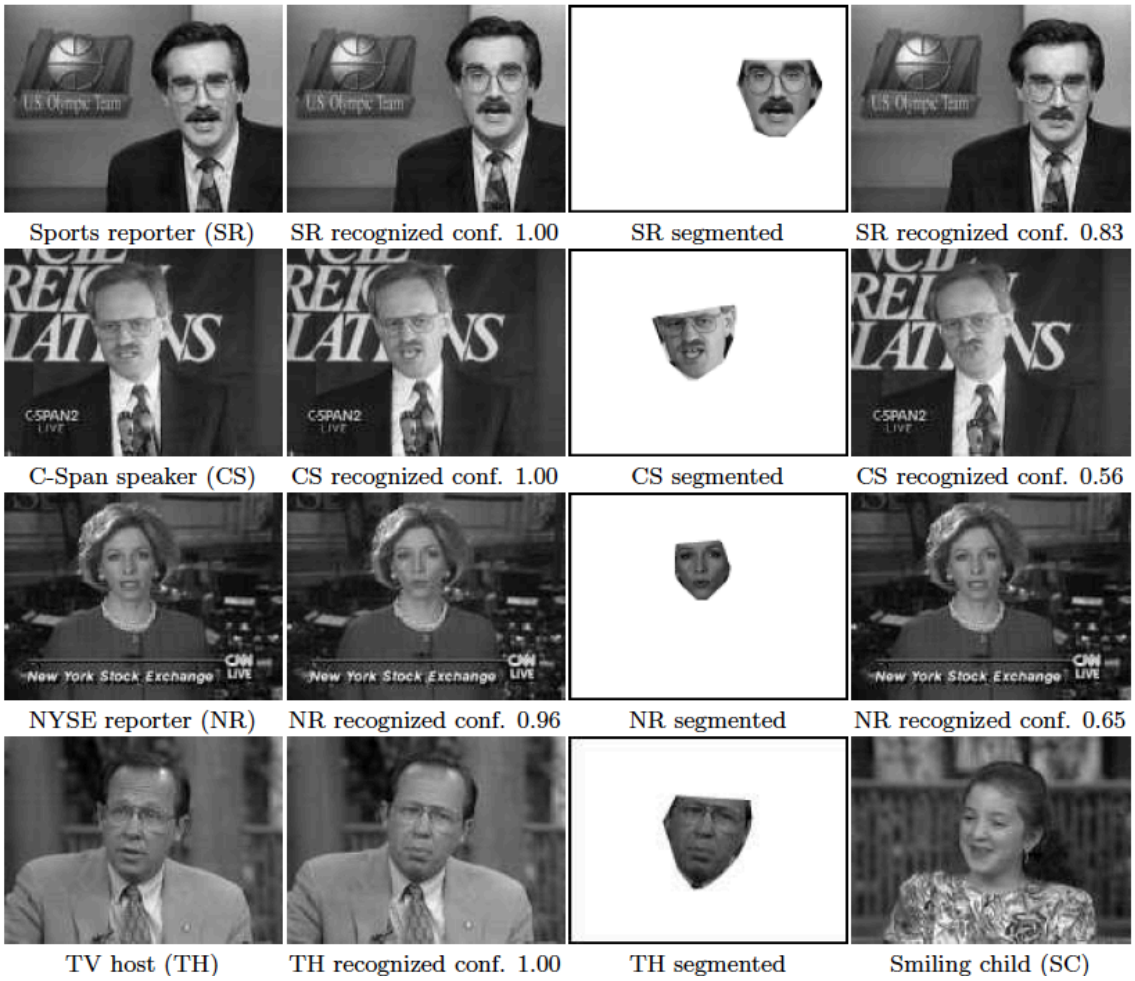
Chair learned Recognized conf. 0.68 Segmented



Monitor learned Recognized conf. 0.81 Segmented

Detect, Recognize, and Segment (3)

Detect, Recognize, and Segment (4)



Limitations of Cresceptron (and later ImageNet) Overcome by the AIML Contest

- The brute-force scans (e.g., left-to-right, top-to-bottom, large-to-small), replaced by general-purpose autonomous attention. No master map!
- The convolution that means sensory-only is replaced by sensorimotor internal representations. Max-pooling is gone!
- The image-resolution based deep-cascade of layers is replaced by an autonomously-generated hierarchy of sensorimotor skills.
- Sensory Hebbian learning in Cresceptron (not error back-prop in ImageNet) becomes optimal sensorimotor Hebbian learning in the sense of maximum likelihood.

AI: Symbolic School vs Connectionist School

Symbols are logic and clean.

Artificial neural networks are analogical and scruffy.

- *Marvin Minsky, 1991*

(Artificial) neural networks do not abstract well.

- Michael Jordan, IJCNN 2011

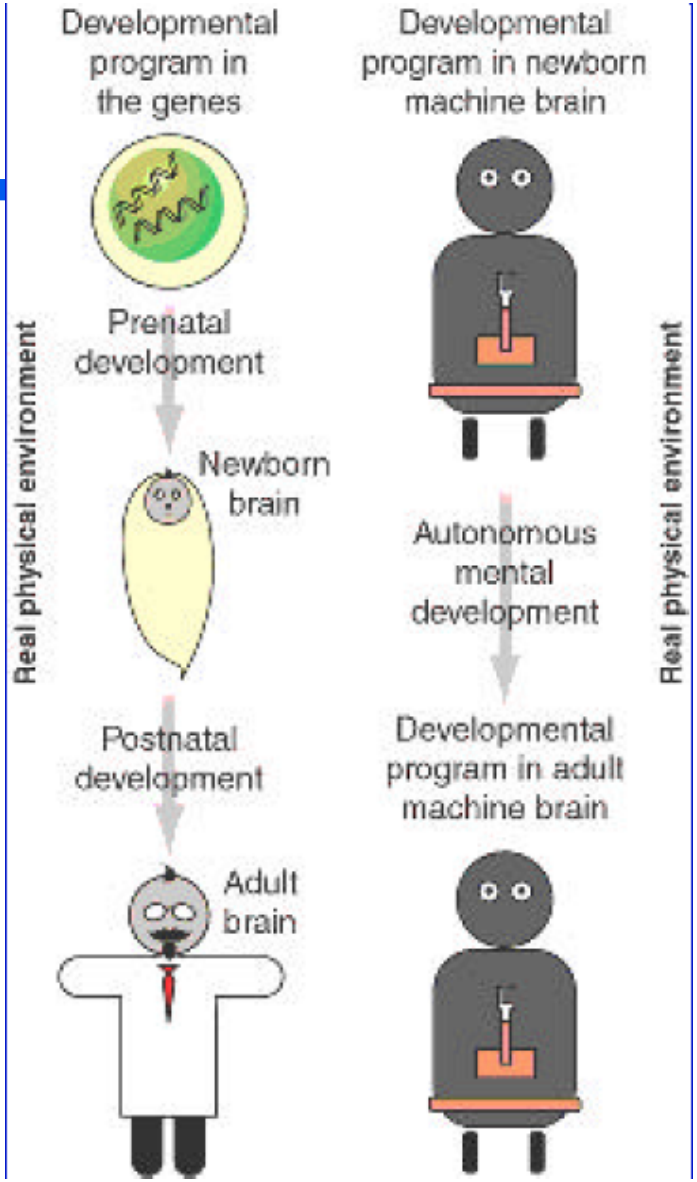
Uniqueness of the AIML Contest

- **Task in-dependence:**
  - **An open number of,**
  - **an open kind of, and**
  - **an unknown set of tasks are trained and tested in "lifetime" learning.**
- **Modality in-dependence:**
  - **Sensory modality: vision, audition, and text**
  - **motor modality: declarative and non-declarative, muscles**

Why 1st Ind.?

- Autonomous Mental Develop (AMD)
- Task-nonspecific
- "Genome-like" Developmental Program (only about 2 pages long)
- IEEE ICDL Conferences
- IEEE Transactions on AMD
- WWN-1 through WWN-9

Weng et al. *Science* 2001



Experiments: Where-What Networks (WWNs)

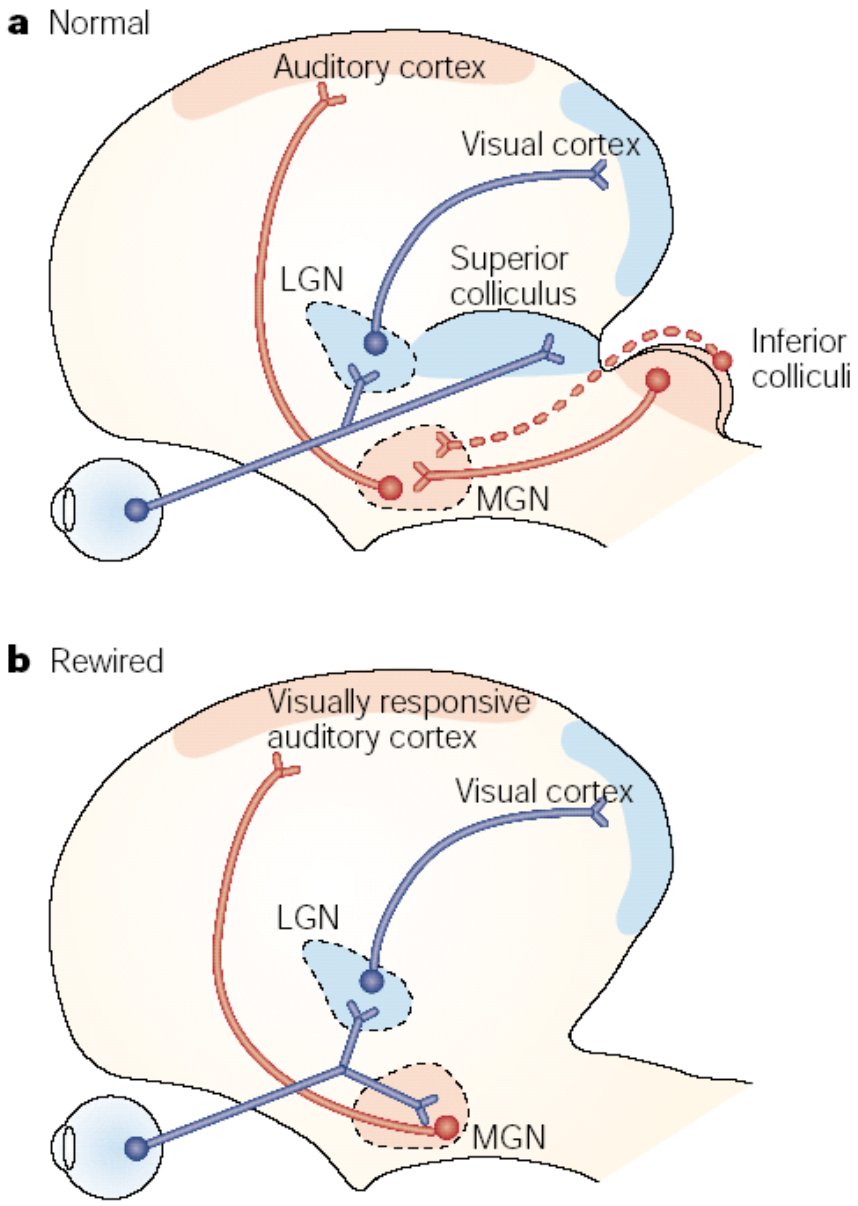
- WWN-1 (2008): single object; cluttered scenes, without pre-segmentation: from location to type (recognition task) and from type to location (detection task) using the same network for the two tasks
- WWN-2 (2010): add to above *free viewing*
- WWN-3 (2010): add to above *multiple objects*
- WWN-4 (2010): allowing bypass MM and PP
- WWN-5 (2011): add to above *scale*
- WWN-6 (2012): synaptogenic factors enable neurons to self-wire
- WWN-7 (2013): add to above *multiple parts of each object*
- WWN-8 (2013): multi-modality (left-eye right eye), to appear BigData 2015
- WWN-9 (2015): relation of objects (A-B group, A plays B, etc.)
- Texty: natural language and knowledge acquisition from natural text

8 Requirements for Practical Learning

1. Environmental openness: cluttered environments
2. High dimensional sensing (e.g., video cameras are necessary)
3. Completeness in internal representation for each age group
4. Online
5. Real time speed
6. Incremental:
   for each fraction of second (e.g., 10-30Hz)
7. Perform while learning
8. Scale up to large memory

Why 2nd Ind.?

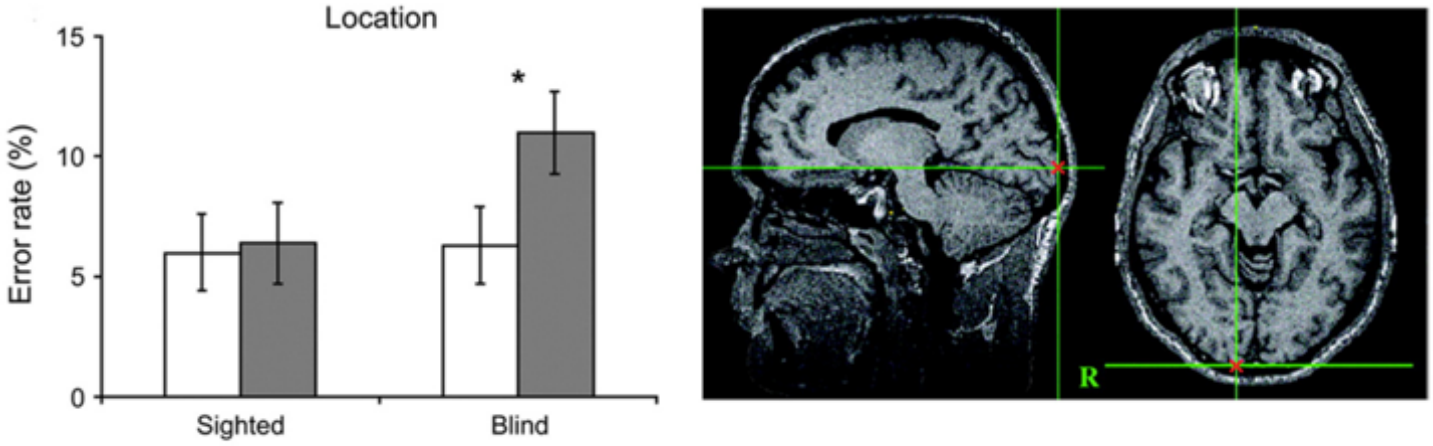
- Ferrets rewired early in life
- "See" using the "sound" zone



Sur, Angelucci and Sharm, *Nature* 1999

Why 2nd Ind.? Human

- Trans Magnetic Stimulation (TMS) to the occipital area (normal visual area) hampers the early blind for
  - Sound localization
  - Verbal memory
  - Braille identification



Collignon et al., (2007)

Thank you
for your attention