CSE 175 / COGS 125 :
Introduction to Artificial Intelligence
An Introduction to Computer Vision

Mohammad K. Ebrahimpour

PhD Student of Electrical Engineering and Computer Science
University of California, Merced

November 2018

About me

http://graduatestudent.ucmerced.edu/mebrahimpour

Mohammad K. Ebrahimpour

PhD Student,
Electrical Engineering and Computer Science,
University of California
5200 N. Lake Road
Merced, CA, 95343
Phone: +1 (209) 228-4101
E-mail: *mebrahimpour* [@] ucmerced [DOT] edu

About me

I recieved my B.Sc. and M.Sc. degrees in computer engineering and artificial intelligence from Shahid Bahonar University of Kerman,Iran (SBUK) in 2013 and 2015, respectively. Since then, I've been working toward my Ph.D. degree under supervision of Prof. David C. Noelle in the computational cognitive neuroscience lab (CCNL) at University of California in Merced. I am interested in deep learning problems as well as their applications in computer vision. Mainly, my research has been focused on object localization and object detection on high resolution images. I am also interested in other machine learning problems like dimensionality reduction and ensemble learning. During my M.Sc. degree I have done research in the field of combinatorial optimization with its application on feature selection. You may take a look at my brief CV.
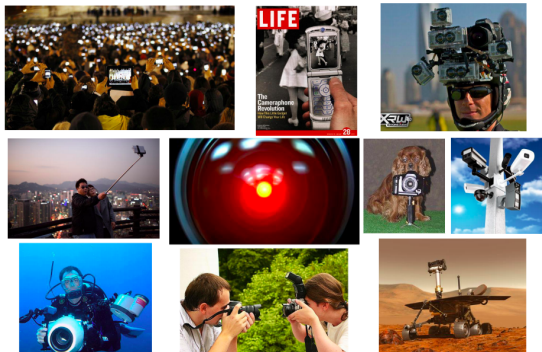
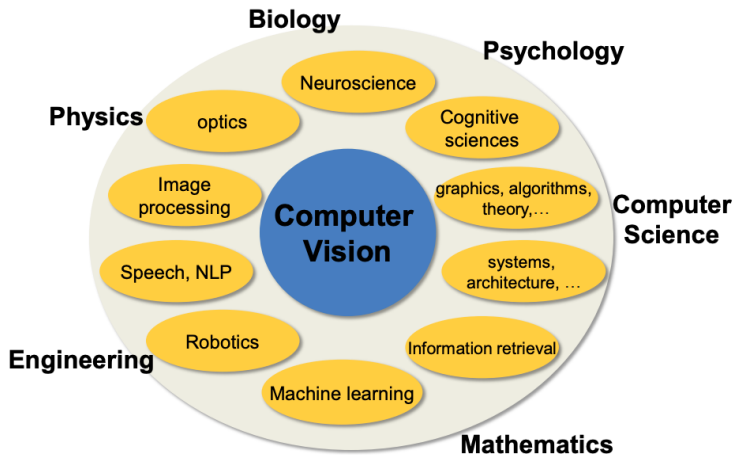### Research

My research interests include

- Object Detection
- Selective Attention
- Deep Learning
- Machine Learning
- Feature Selection
- Ensemble Learning

Welcome to Computer Vision!

- Vision is the fastest growing field in AI
- Cisco claimed that in 2018 more than 85% of internet data would pixels related
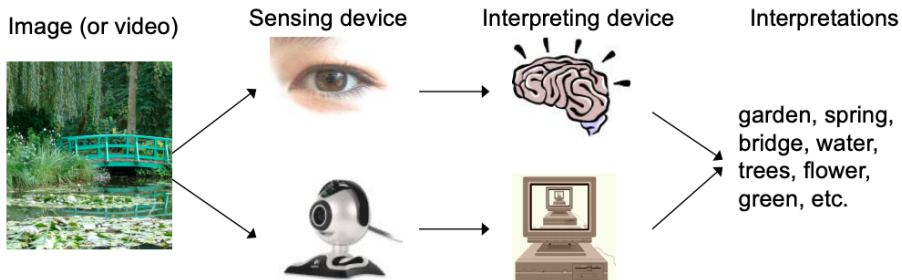- In every minute 150 hours of videos will be uploaded through the Youtube

Computer vision is highly interdisciplinary field

What is (Computer) Vision?

Image (or video)   Sensing device   Interpreting device   Interpretations

garden, spring, bridge, water, trees, flower, green, etc.

- To bridge the gap between pixels and "meaning"

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

How does vision work in the biological brain?
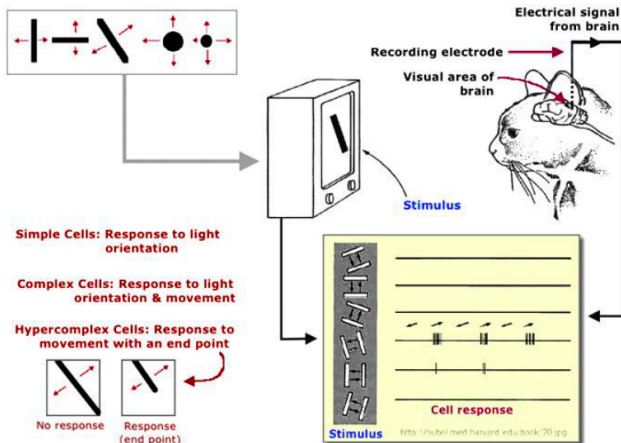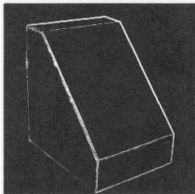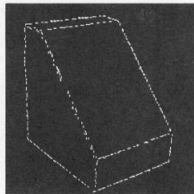

Hubel & Wiesel, 1959

Block world

Larry Roberts, 1963
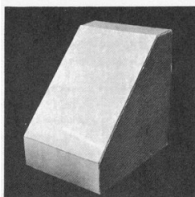
(a) Original picture.
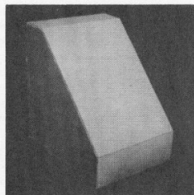
(b) Computer display of picture (reflected by mistake).

(c) Differentiated picture.

(d) Feature points selected.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group   July 7, 1966
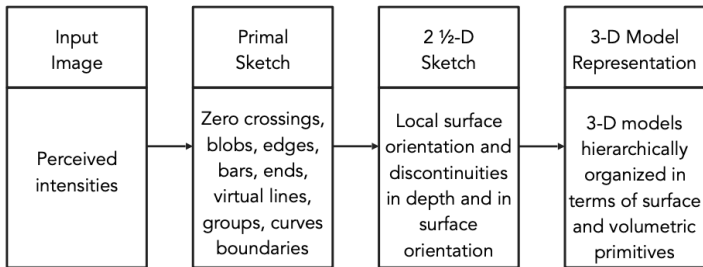Vision Memo. No. 100.

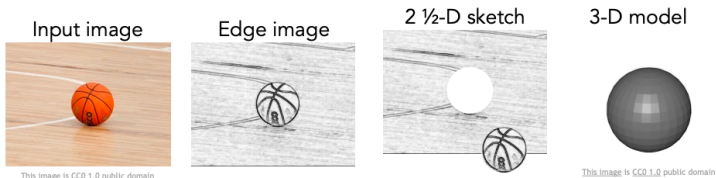THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers
effectively in the construction of a significant part of a visual system.
The particular task was chosen partly because it can be segmented into
sub-problems which will allow individuals to work independently and yet
participate in the construction of a system complex enough to be a real
landmark in the development of "pattern recognition!".

Another peace of vision– Vision is hierarchical

David Marr, 1970's

Stages of Visual Representations

Input image   Edge image   2 ½-D sketch   3-D model

This image is CC0 1.0 public domain   This image is CC0 1.0 public domain

Input Image	Primal Sketch	2 ½-D Sketch	3-D Model Representation
Perceived intensities	Zero crossings, blobs, edges, bars, ends, virtual lines, groups, curves boundaries	Local surface orientation and discontinuities in depth and in surface orientation	3-D models hierarchically organized in terms of surface and volumetric primitives

Stages of Visual Representation, David Marr, 1970s

Active Vision

Active vision is particularly important to cope with problems like occlusions, limited field of view and limited resolution of the camera.
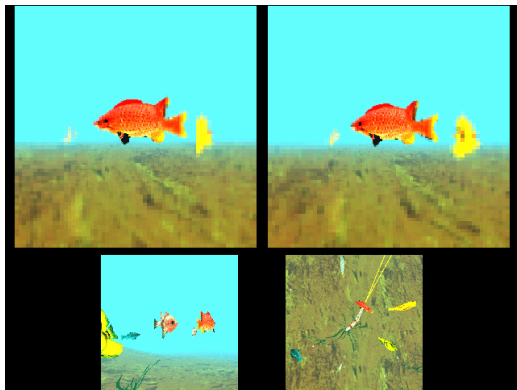
Computer Vision in 80's
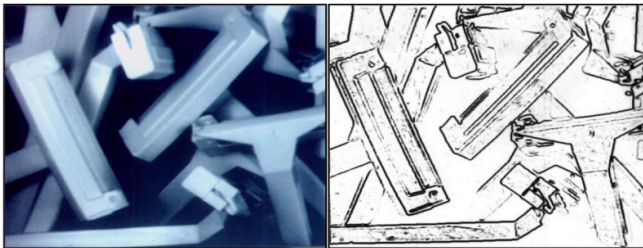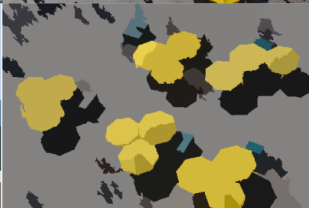

David Lowe, 1987

Colorful images in 90's

Normalized Cut (Shi & Malik, 1997)

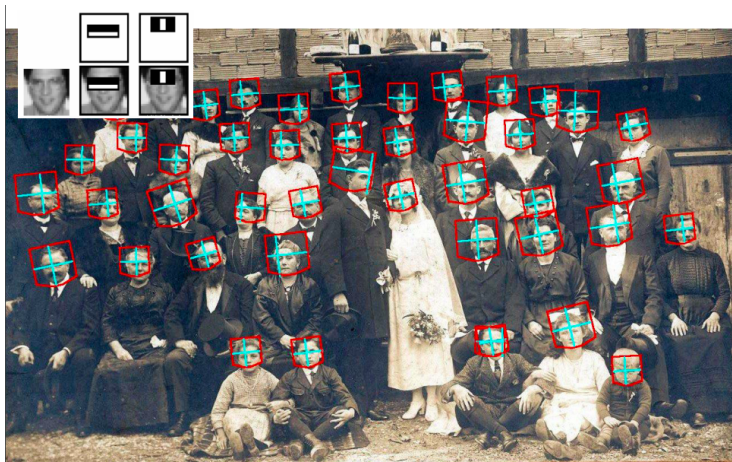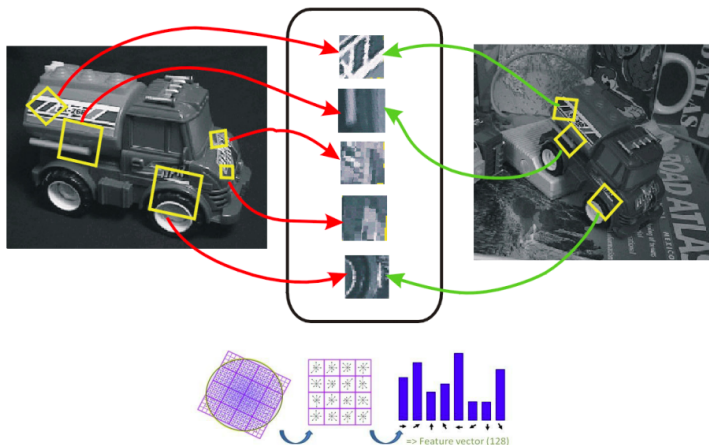# Vision in early 2000 - Face Detection!

Viola & Jones, 2001 Fuji Film implemented this approach in their digital camera in 2006!
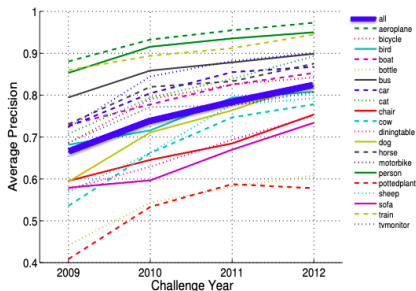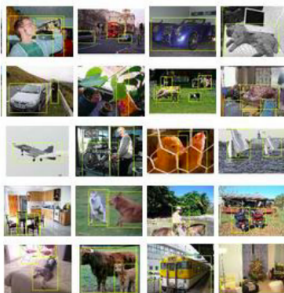
Feature Extraction

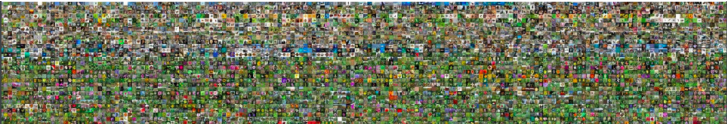"SIFT" & Object Recognition, David Lowe, 1999

**PASCAL Visual Object Challenge
(20 object categories)**
[Everingham et al. 2006-2012]

World is bigger than 20 categories

IM🔶GENET

www.image-net.org

22K categories and **14M** images

- Animals
  - Bird
  - Fish
  - Mammal
  - Invertebrate
- Plants
  - Tree
  - Flower
- Food
- Materials
- Structures
- Artifact
  - Tools
  - Appliances
  - Structures
- Person
- Scenes
  - Indoor
  - Geological Formations
- Sport Activities

Deng, Dong, Socher, Li, Li, & Fei-Fei, 2009

Image Net Challenge as Computer Vision Olympics

IM**A**GENET **Large Scale Visual Recognition Challenge**

The Image Classification Challenge:
1,000 object classes
1,431,167 images
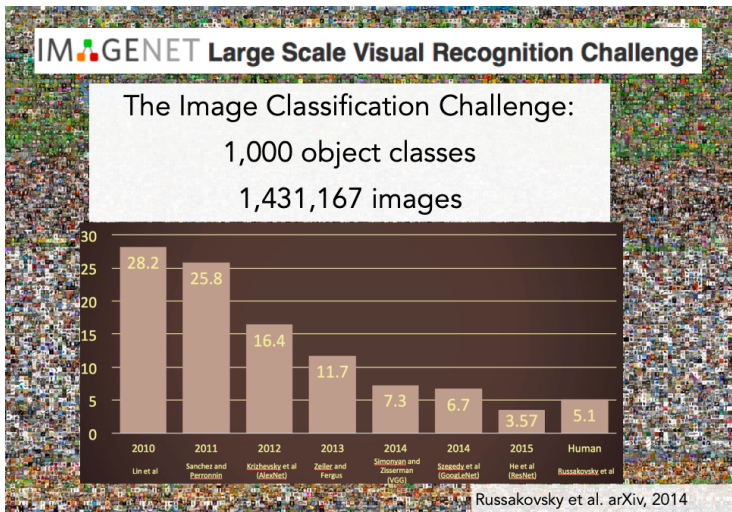
Output:
Scale
T-shirt
Steel drum
Drumstick
Mud turtle
✔

Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle
✗

Russakovsky et al. arXiv, 2014

Surpassing human performance!

IM·GENET **Large Scale Visual Recognition Challenge**

The Image Classification Challenge:
1,000 object classes
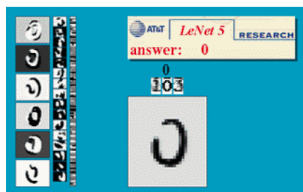1,431,167 images

Russakovsky et al. arXiv, 2014

Applications of Computer Vision (Optical Character Recognition)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software

Digit recognition, AT&T labs

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition
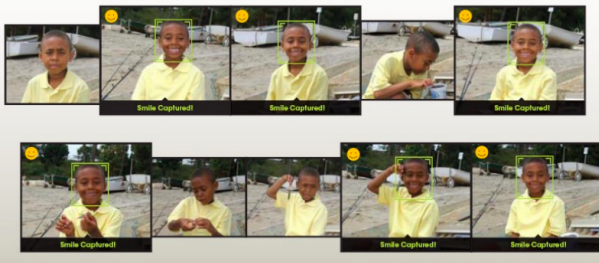
Source: S. Seitz

Face detection

- **Many digital cameras now detect faces**
  - Canon, Sony, Fuji, ...

Source: S. Seitz

Applications of Computer Vision (Smile Detection)

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.
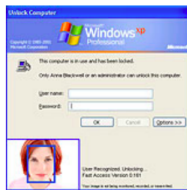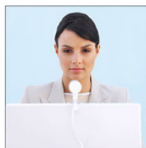
Sony Cyber-shot® T70 Digital Still Camera

Source: S. Seitz

Applications of Computer Vision (Biometrics)

Fingerprint scanners on many new laptops, other devices

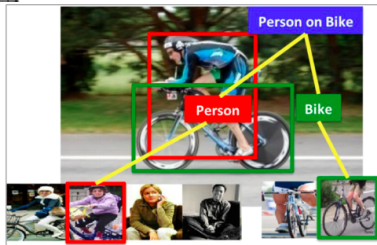Face recognition systems now beginning to appear more widely
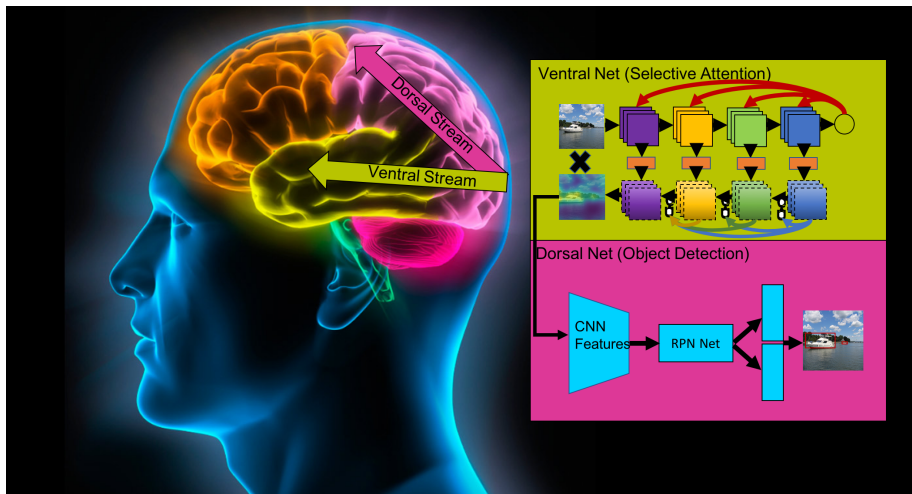http://www.sensiblevision.com/

Source: S. Seitz

Visual Recognition Problems

- Object detection
- Action classification
- Image captioning
- …

My Research
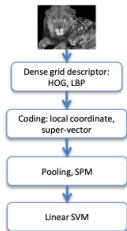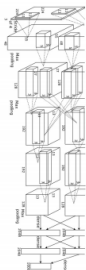
Deep Learning Models

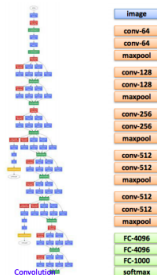| **Year 2010** | **Year 2012** | **Year 2014** | | **Year 2015** |
| NEC-UIUC | SuperVision | GoogLeNet | VGG | MSRA |

[Lin CVPR 2011]   [Krizhevsky NIPS 2012]   [Szegedy arxiv 2014]   [Simonyan arxiv 2014]

Computer Vision Goal

PT = 500ms

Some kind of game or fight. Two groups of two men? The man on the left is throwing something. Outdoors seemed like because i have an impression of grass and maybe lines on the grass? That would be why I think perhaps a game, rough game though, more like rugby than football because they pairs weren't in pads and helmets, though I did get the impression of similar clothing. maybe some trees? in the background. (Subject: SM)

Fei-Fei, Iyer, Koch, Perona, JoV, 2007

We Are FAR From the Ultimate Goal!

Computer Vision
Technology
Can Better Our Lives