Course: Deep Learning

Unit 2: Computer Vision

Introduction to Computer Vision

Luis Baumela

Universidad Politécnica de Madrid

Introduction to Computer Vision

1. What is Computer Vision?
   - Related problems
   - Areas of interest
   - Applications
   - Deep Learnig in Computer Vision

2. Object recognition
   - Why is it difficult?

3. Miscelaneous
   - Overview of unit 2

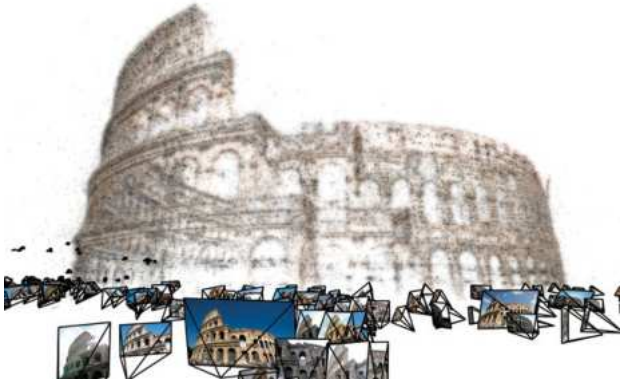
Computer Vision

- What is computer vision?

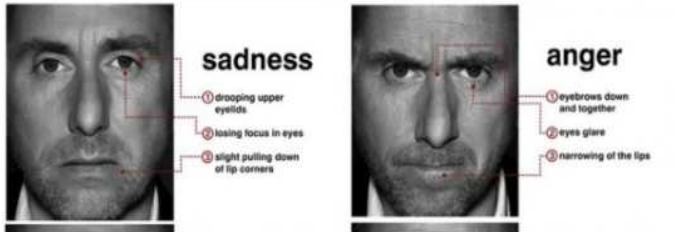
Make machines that "see"; obtain information from images.

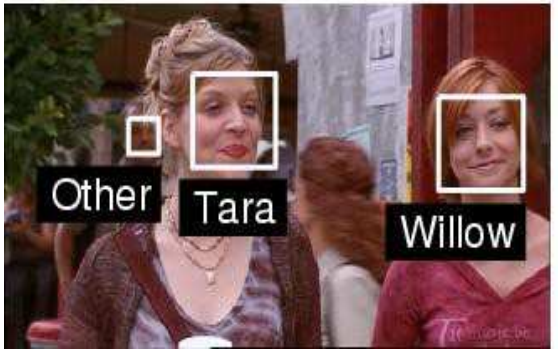
Recognize places

Reconstruction

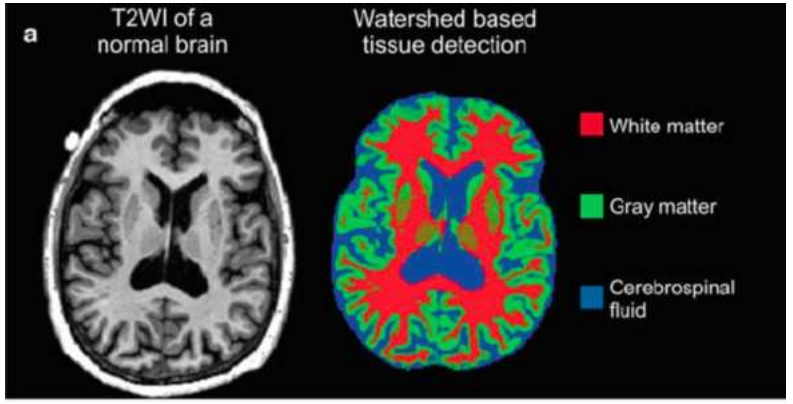


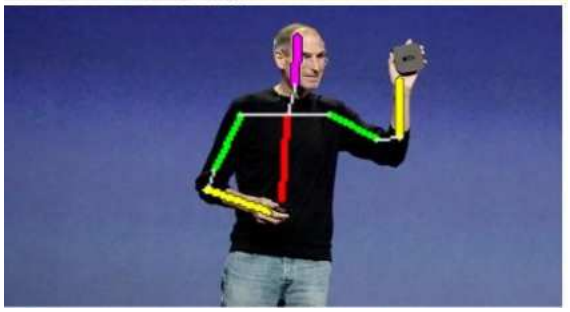
Look at people





Medical imaging



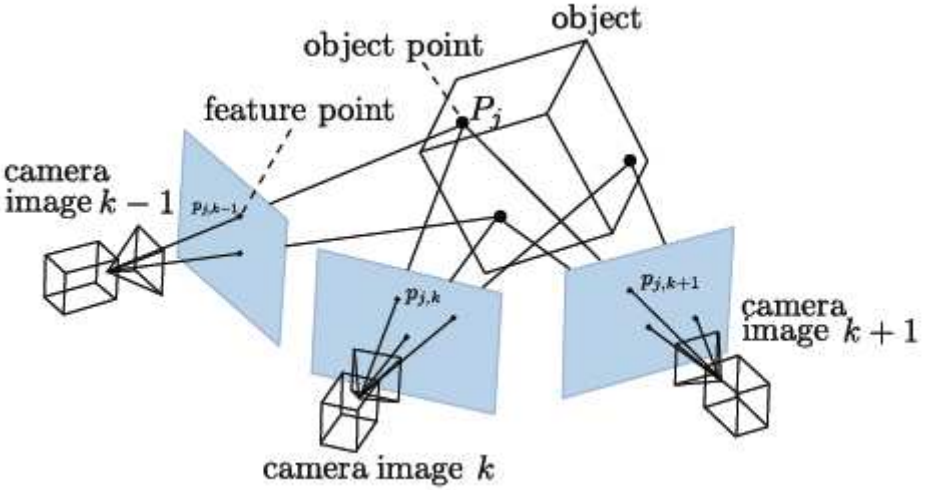


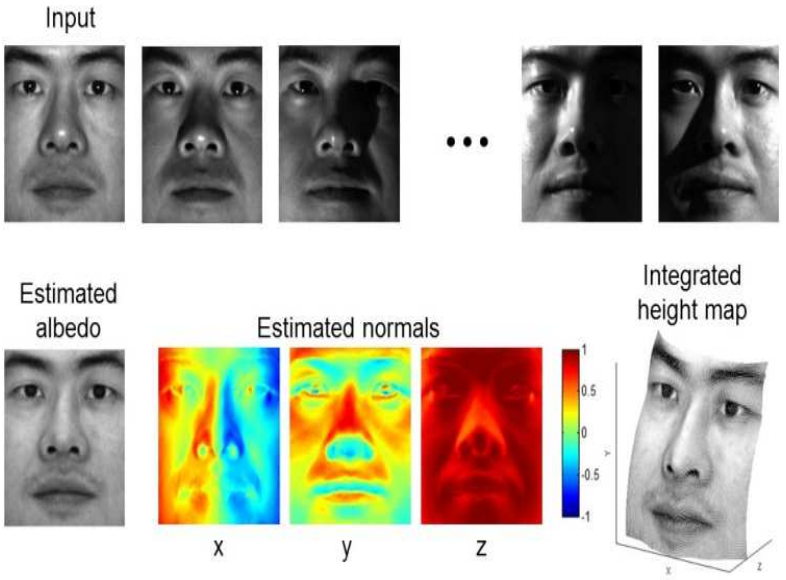
Computer Vision

- What is computer vision?

Related problems:

- **Reconstruccion**
  - Shape from X (estéreo)
  - Shadows and shading
  - Shape fron defocus



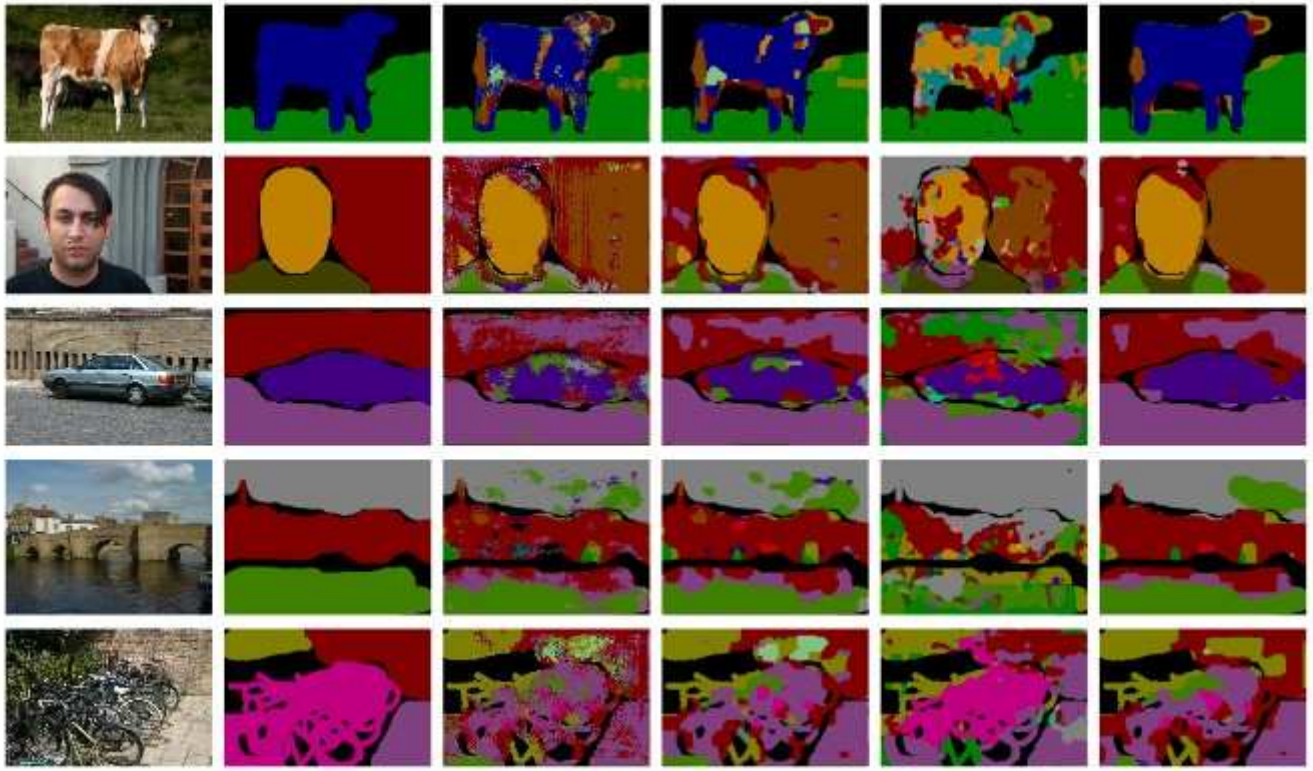


Computer Vision

- What is computer vision?

Related problems:

- Reconstruccion
- **Segmentation**

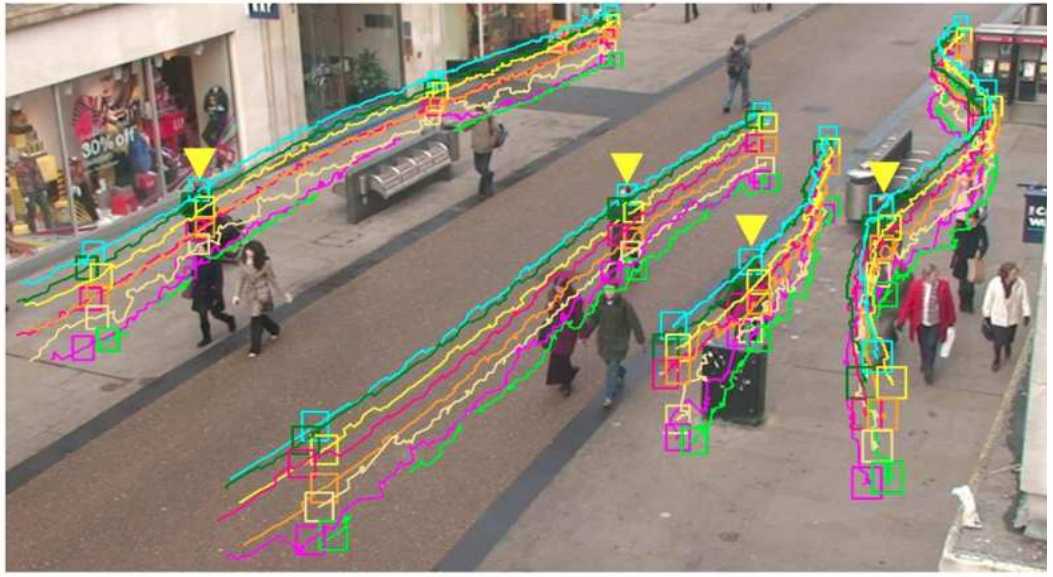


Computer Vision

- What is computer vision?

Related problems:

- Reconstruccion
- Segmentation
- **Tracking**



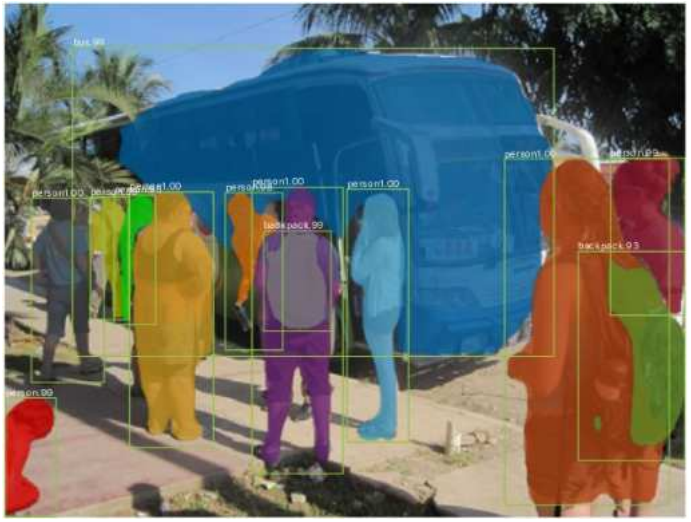
Computer Vision

- What is computer vision?

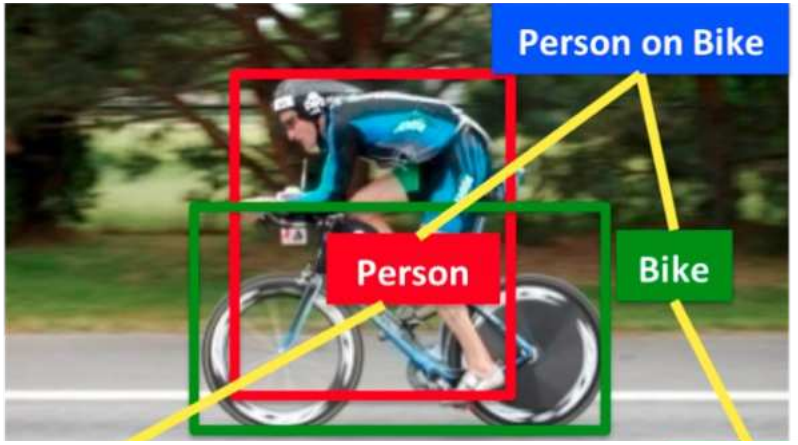
Related problems:

- Reconstruccion
- Segmentation
- Tracking
- **Recognition**

Image classification



Instance segmentation



Object detection/localization

Computer Vision

- What is computer vision?

Related problems:

- Reconstruccion
- Segmentation
- Tracking
- Recognition
- **Interpretation**

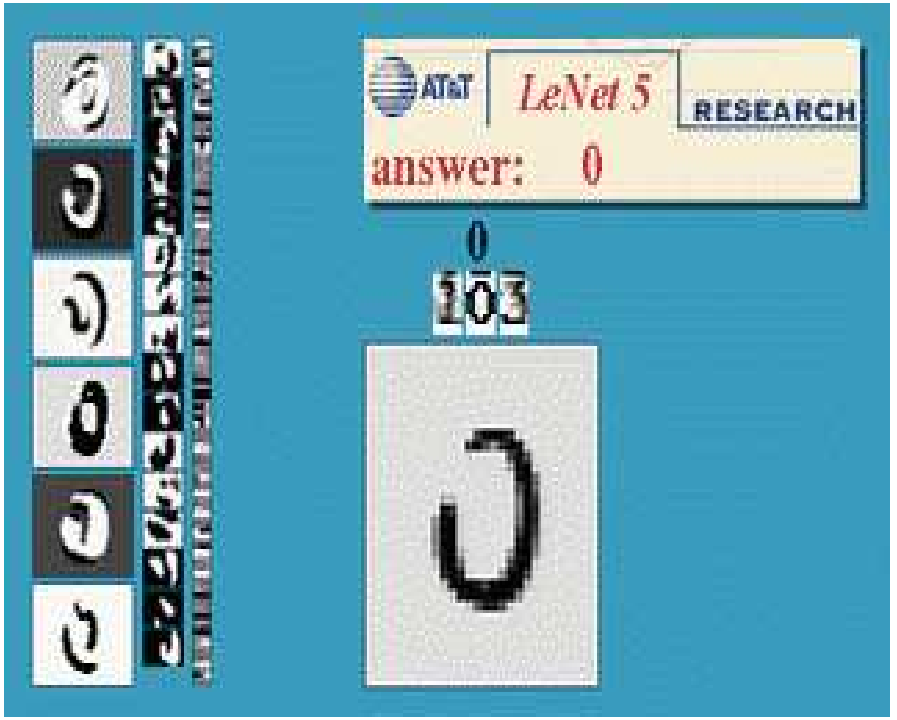
"man in black shirt is playing guitar."

"construction worker in orange safety vest is

Computer Vision

- Applications

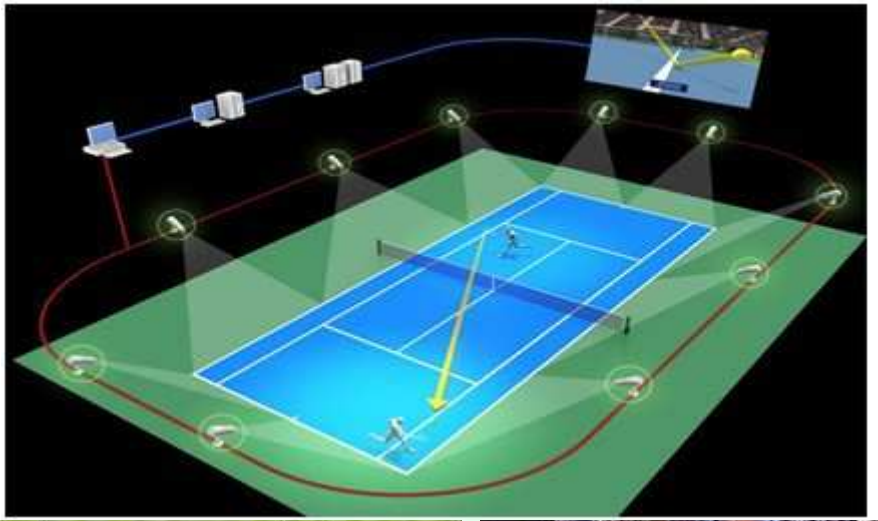
Optical Character Recognition (OCR)
Reading checks and forms

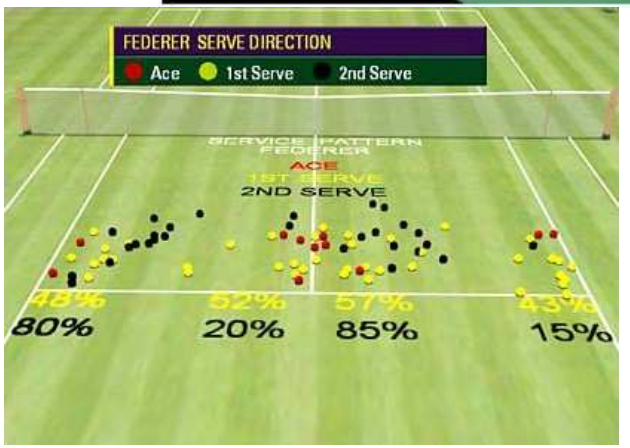


OCR, LeNet, LeCun 1989

Computer Vision

- Applications
  3D reconstruction. Hawk Eye





http://www.hawkeyeinnovations.co.uk/page/sports-officiating/tennis

Computer Vision

- Applications

  Movies (graphics animation, augmented reality)
  Avatar, Rise of the planet of the apes, ….

Computer Vision

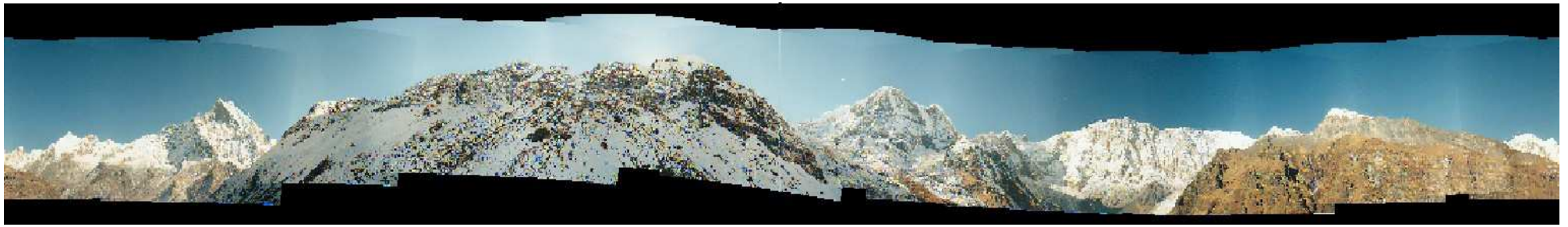
- Applications

  Movies (graphics animation, augmented reality)
  Avatar, Rise of the planet of the apes, ….

Computer Vision

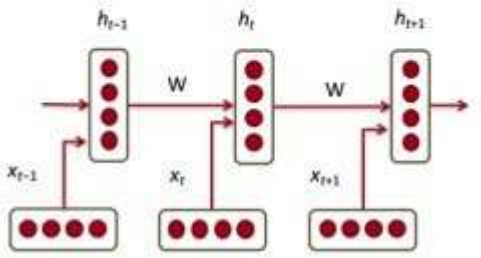
- Applications

Computational photography

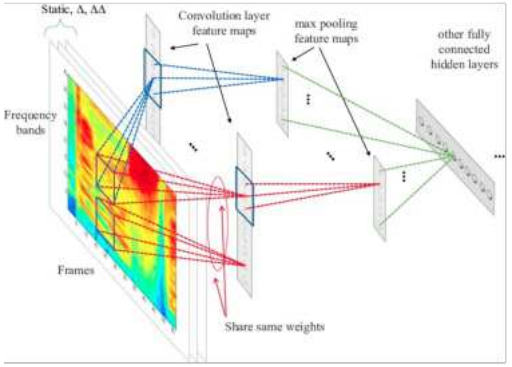


Deep Learning and perception

- In the past perception modalities used different technologies



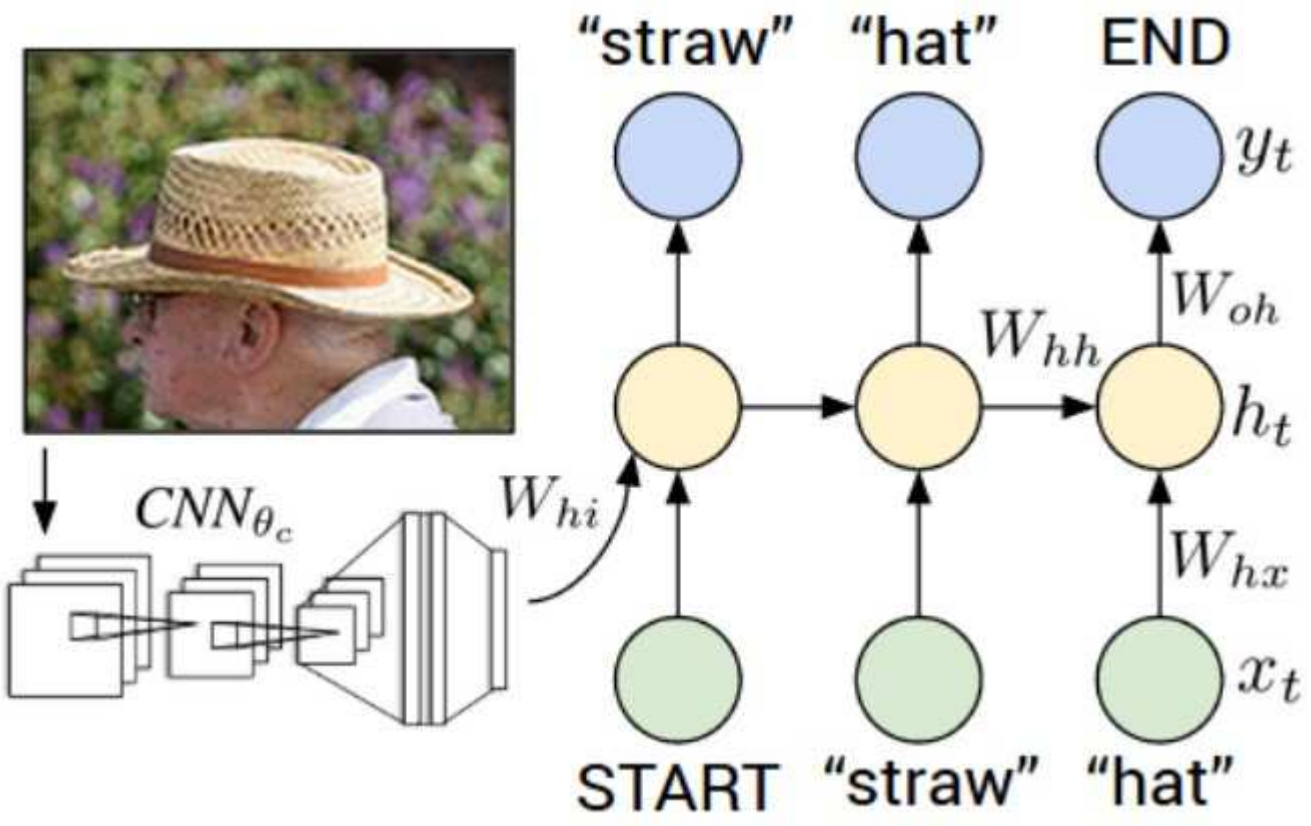
- Deep learning is the state of the art for
  - Computer vision
  - Speech recognition
  - Natural language



- DL provides a common ground to address perception
- Opens the way for ultimodal communication

Computer Vision

- Deep Learning results
  Image to text



Karpathy, Deep Visual-Semantic Alignments for Generating Image Descriptions. CVPR, 2015.

Computer Vision

- Deep Learning results
  Image to text

"man in black shirt is playing guitar."

"construction worker in orange safety vest is working on road."

"two young girls are playing with lego toy."

"boy is doing backflip on wakeboard."

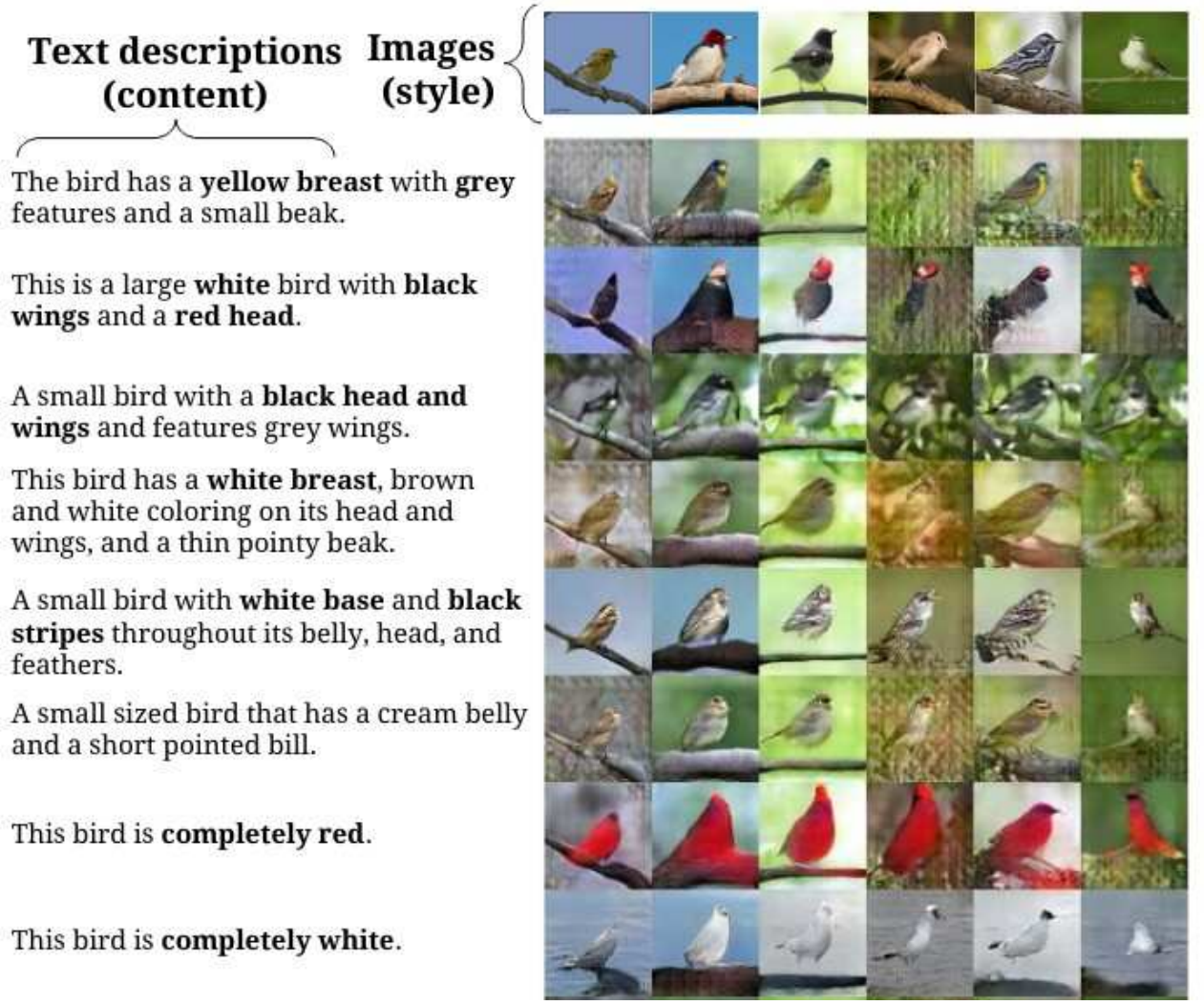
Karpathy, Deep Visual-Semantic Alignments for Generating Image Descriptions. CVPR2015.

Computer Vision

- Deep Learning results
  Text to image

this magnificent fellow is almost all black with a red crest, and white cheek patch.

this white and yellow flower have thin white petals and a round yellow stamen

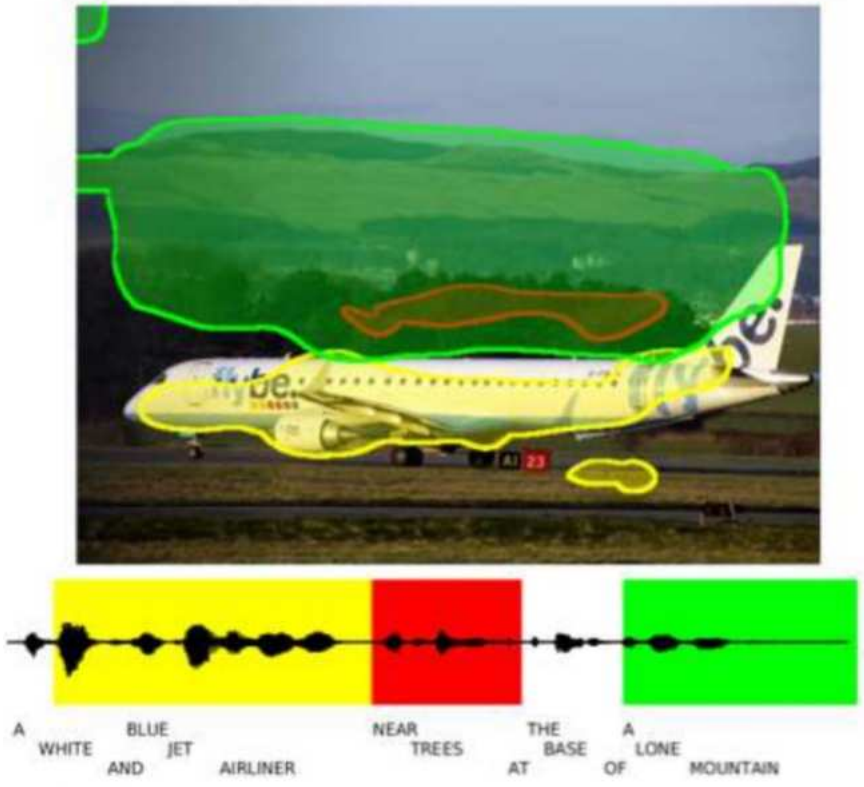


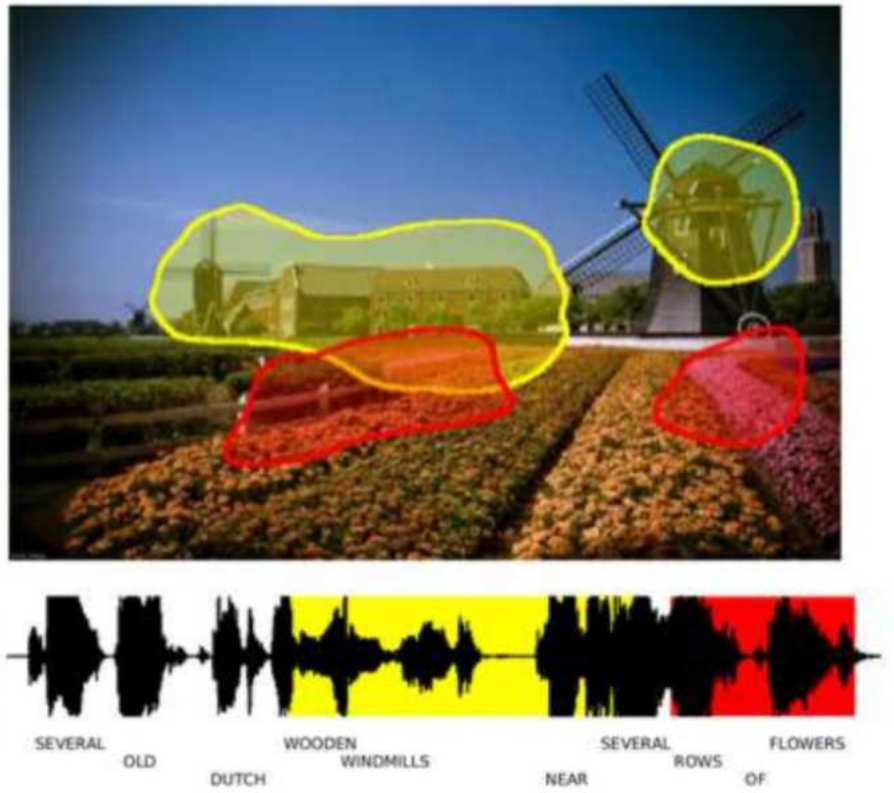
Reed, Generative Adversarial Text to Image Synthesis, ICML, 2016.

Computer Vision

- Deep Learning results

  Self learning semantic correspondences between speech and pixels with no annotated data



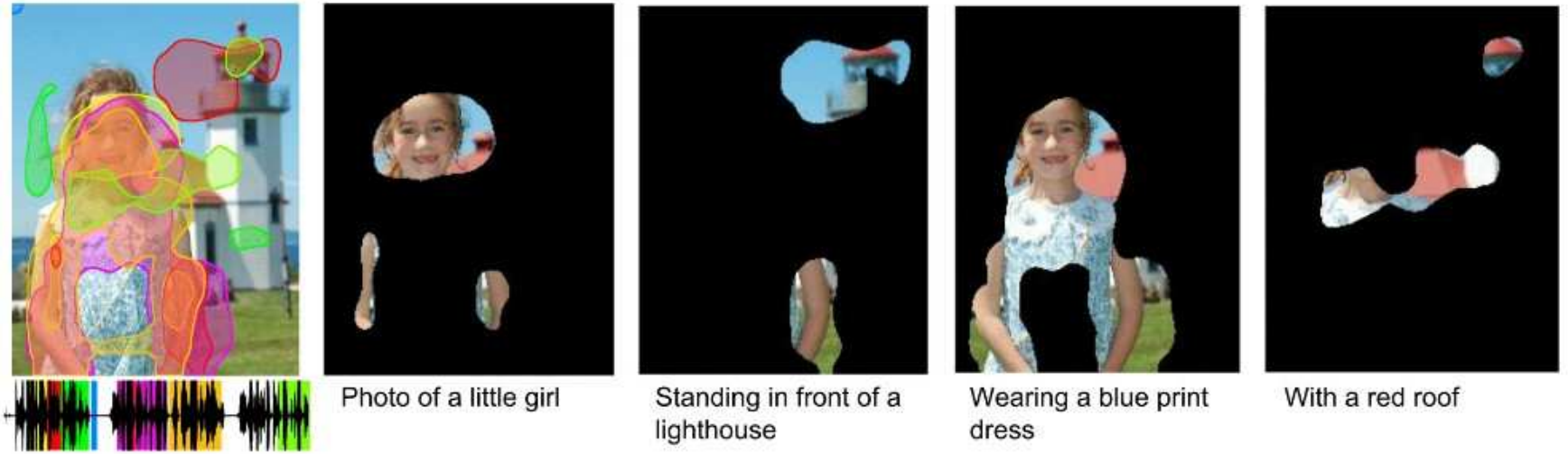


Harwath, Jointly Discovering Visual Objects and Spoken Words from Raw Sensory Input. ECCV18.

Computer Vision

- Deep Learning results

  Self learning semantic correspondences between speech and pixels with no annotated data

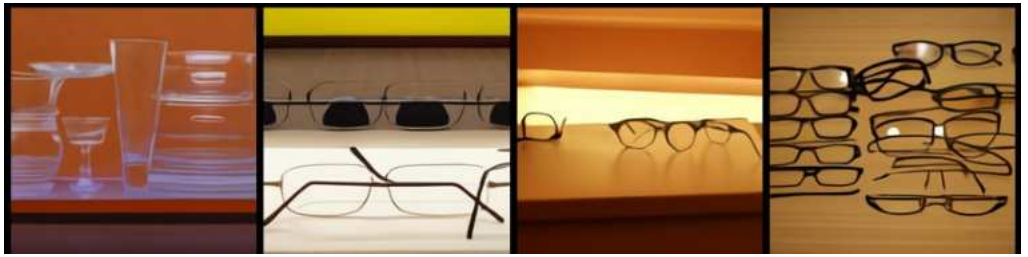


Harwath, Jointly Discovering Visual Objects and Spoken Words from Raw Sensory Input. ECCV18.

Computer Vision

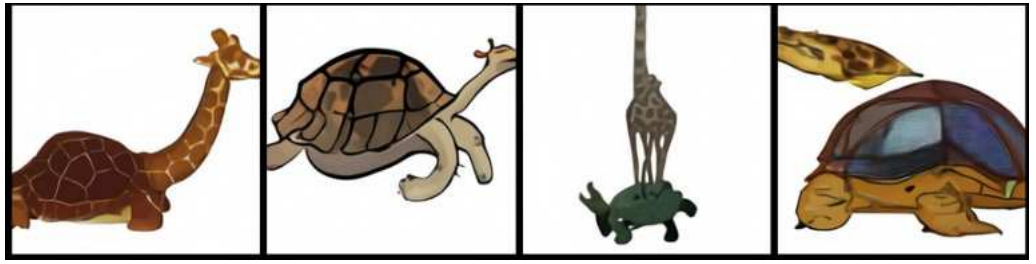
- Deep Learning results.
  DALL-E

  Image generation from text description, trained from 250M image-text pairs from the Internet



A collection of glasses sitting on a table

An arcmchair in the shape of an avocado



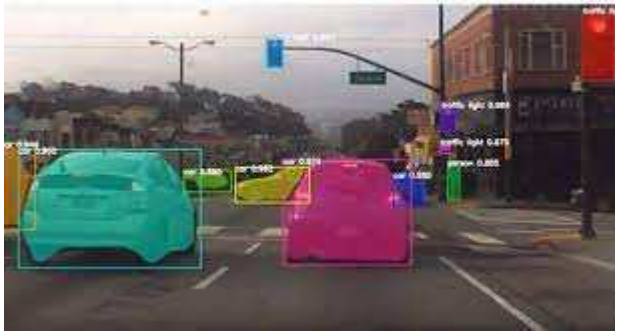
A jiraffe made of turtle

Computer Vision

- Deep Learning applications

Automated retail

Augmented reality



Self-driving

Human pose

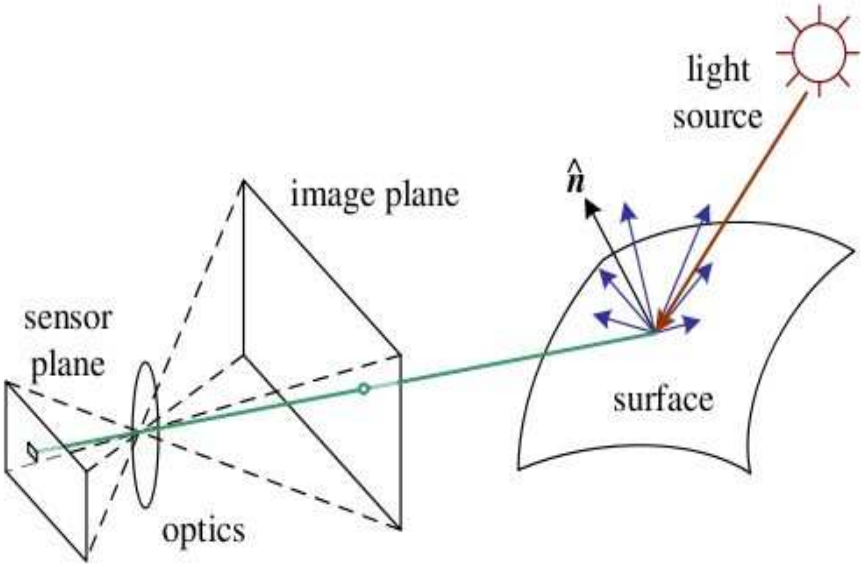
Face recognition

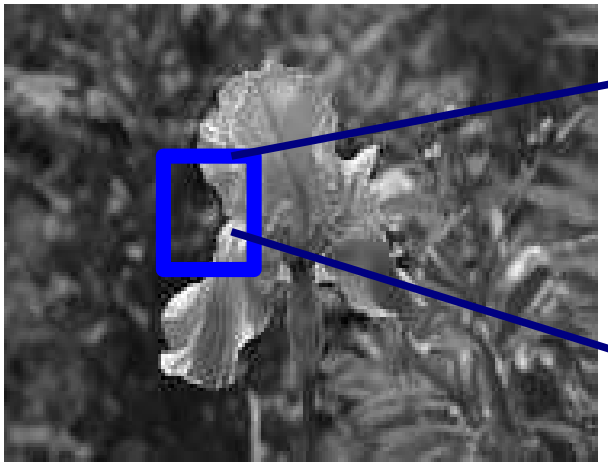
Driver monitoring

Computer Vision

- Why is vision hard?

It is an *inverse problem*





45	60	98	127	132	133	137	133
46	65	98	123	126	128	131	133
47	65	96	115	119	123	135	137
47	63	91	107	113	122	138	134
50	59	80	97	110	123	133	134
49	53	68	83	97	113	128	133
50	50	58	70	84	102	116	126
50	50	52	58	69	86	101	120

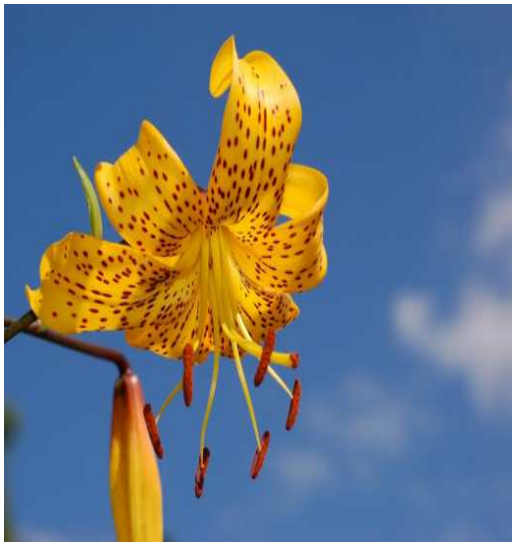
Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes







Object Recognition

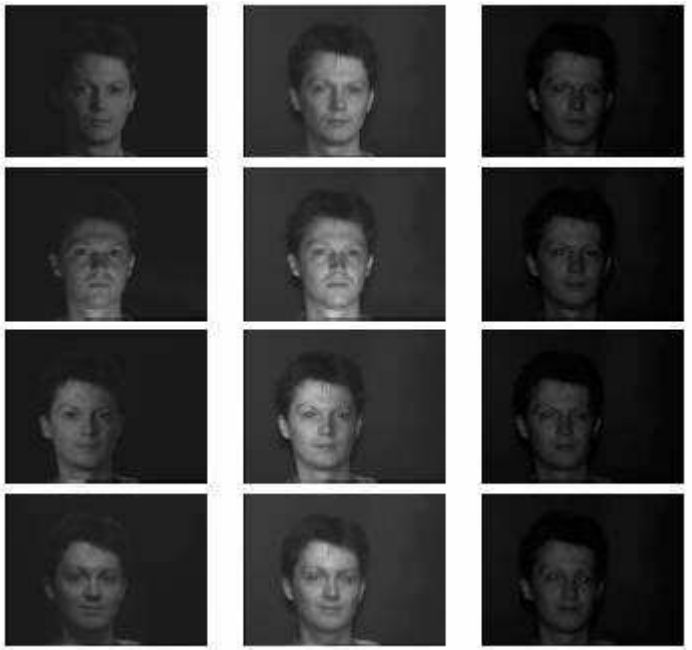
- Why is object recognition hard?
  - High variability of natural object classes

Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes

Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes
  - Lighting contrast, shadows, specularities

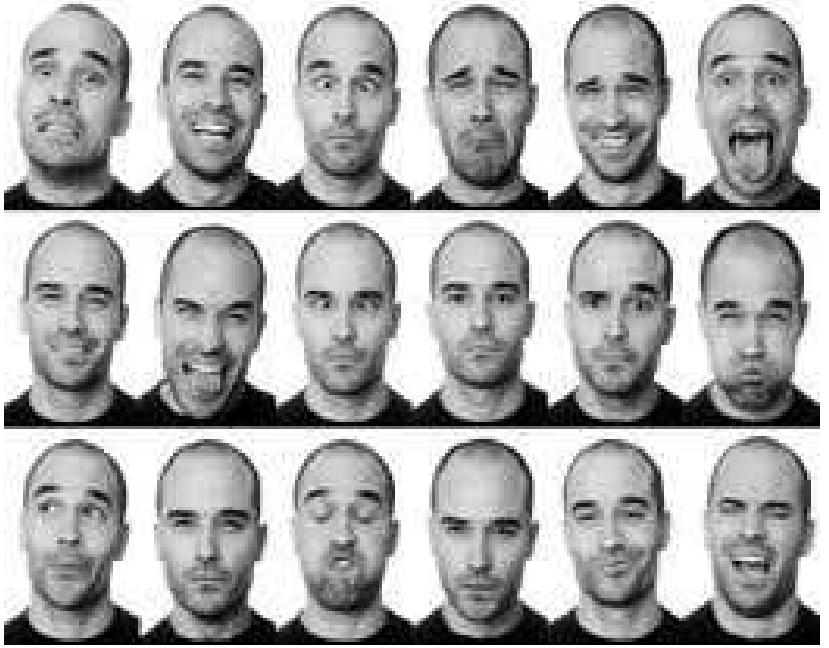


Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes
  - Lighting contrast, shadows, specularities
  - Geometric variability

Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes
  - Lighting contrast, shadows, specularities
  - Geometric variability
  - Deformation



Object Recognition

- Why is object recognition hard?
  - High variability of natural object classes
  - Lighting contrast, shadows, specularities
  - Geometric variability
  - Clutter, occlusion
  - Deformation

How is it addressed?

Overview of Unit 2:

1. Introduction to
   - computer vision,
   - image processing, convolution, object recognition,
2. Convolutional Neural Networks
3. Deep Architectures
   - AlexNet, VGG, LeNet, ResNet,…
   - MobileNet, Efficient Net, ...
4. Representation Learning
5. Computer Vision applications
   - Segmentation, Object Detection