

Intro Visual Recognition

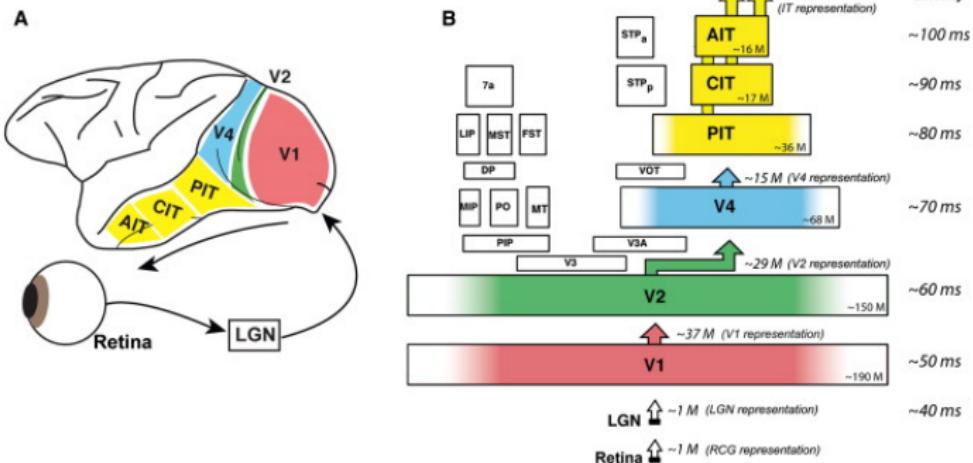
Alvaro Soto

Computer Science Department (DCC), PUC

what_do_you_see() ?



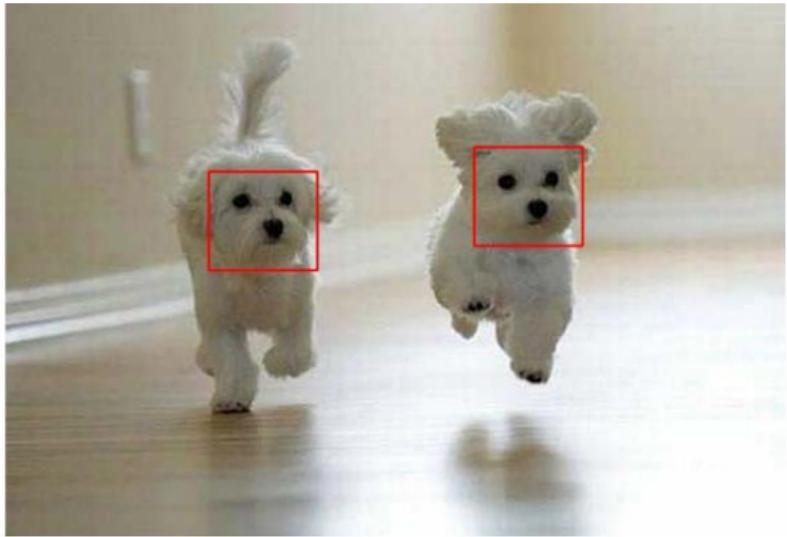
The amazing visual cortex



Computer Vision: Template Matching



? →

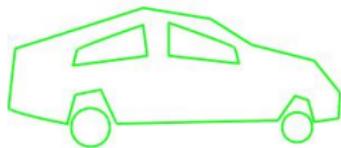


Computer Vision: Problems

- Changes in illumination
- Pose variations
- Scale variations
- Occlusions
- Changes in view-point
- ...

Computer Vision: Geometric Models

Model of Car



Image



Computer Vision: Problems

- Changes in illumination
- Pose variations
- Scale variations
- Occlusions
- Changes in view-point
- Intra-class variations
- Deformations
- Background clutter
- ...

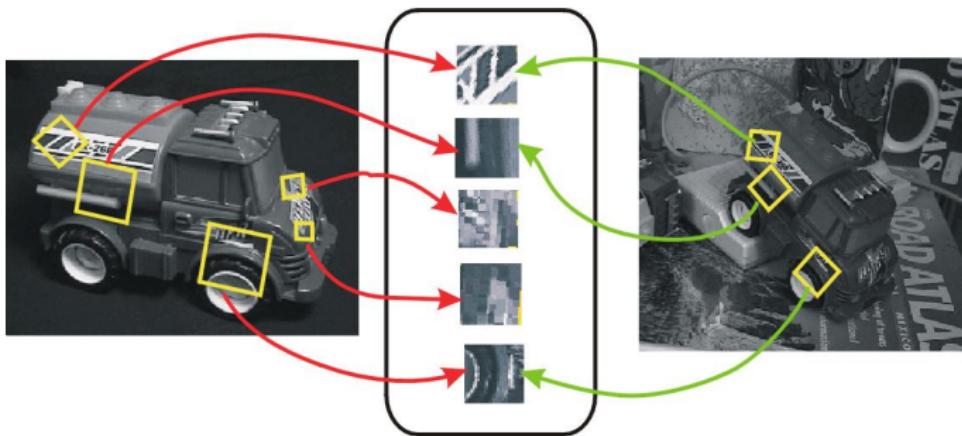
what_do_you_see() ?

Computer Vision + Machine Learning

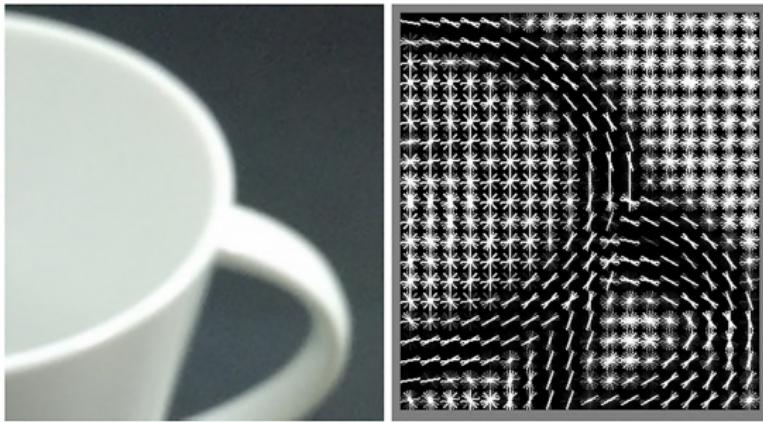
Handcrafted visual features + Classifier

Handcrafted Features

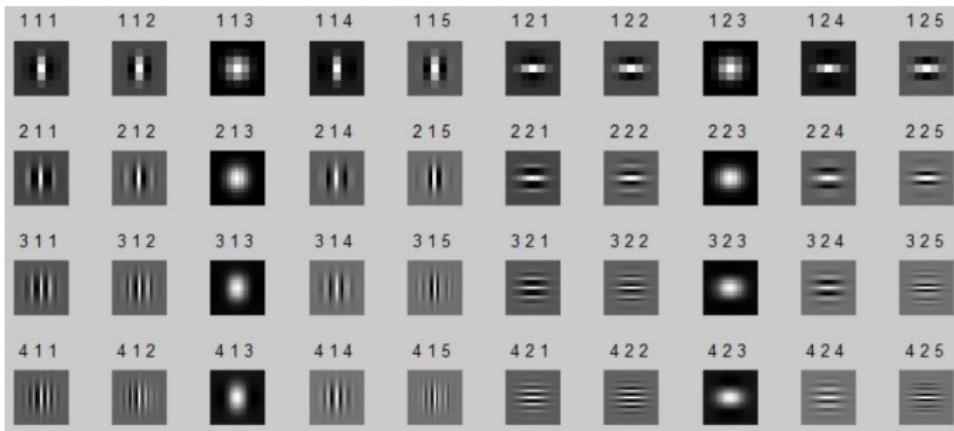
SIFT (D. Lowe, 1999)



HoG (Dalal and Triggs, 2005)

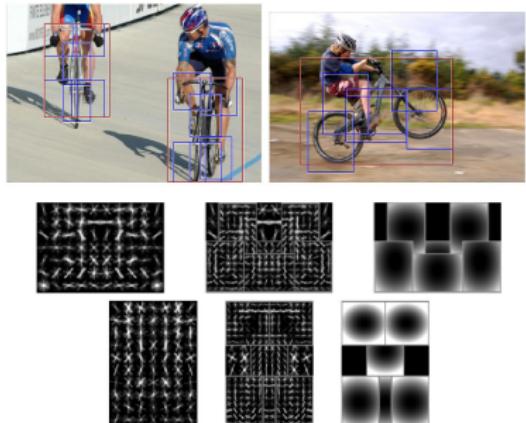
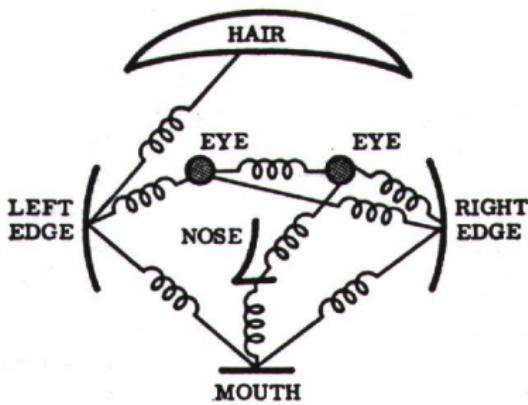


Gabor Filters

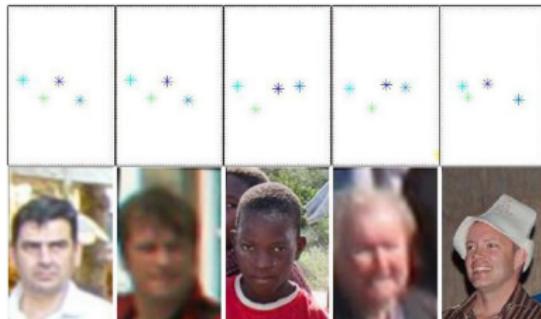


Part Based Models (hierarchical models)

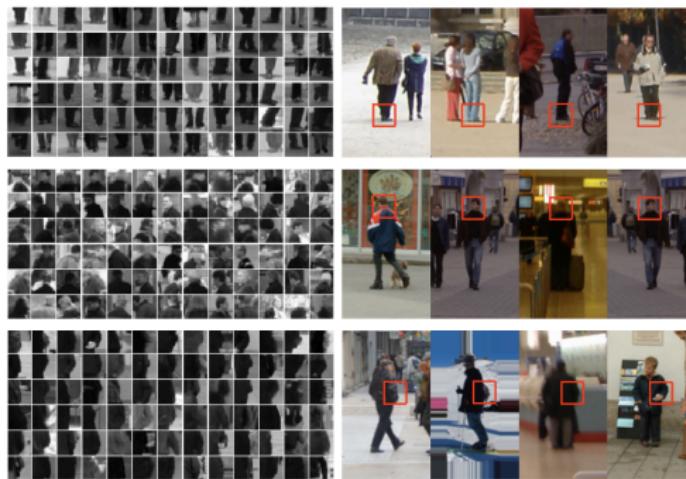
Deformable Part Model (Felzenszwalb et al., 2008)



Poselets (Bourdev et al., 2010)



K-Classifiers (Soto et al., 2012)



K-Classifiers

For $j=1:K$
{

$$\min_{\{w_j, \beta_j\}} \frac{1}{2} ||w_j||_2^2 + C \sum_{i=1}^{N_j} \xi_i$$

subject to :

$$(w_j^T x_i + \beta_j) \geq 1 - \xi_i, \quad \forall i = 1, \dots, N_j$$

}

K-Classifiers

For $j=1:K$

$$\min_{\{\xi_i, y_i: y_i \neq 0\}_{i=1}^N} \min_{\{w_j, \beta_j\}} \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^{N_j} \xi_i$$

subject to :

$$(w_j^T x_i + \beta_j) \geq 1 - \xi_i, \quad \forall i = 1, \dots, N, j = 1, \dots, K : y_i = j,$$
$$-(w_j^T x_i + \beta_j) \geq 1 - \xi_i, \quad \forall i = 1, \dots, N, j = 1, \dots, K : y_i \neq j,$$

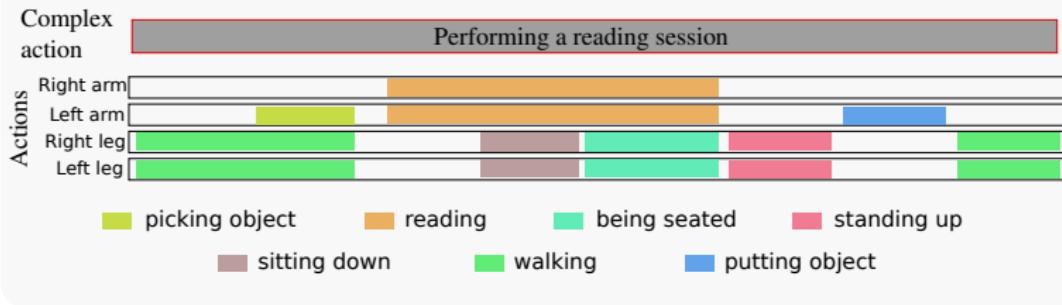
}

Hierarchical K-Classifiers (Lillo et al., 2014)

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$



Generated labels



Hierarchical K-Classifiers (Lobel et al., 2014)

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

Hierarchical K-Classifiers (Lobel et al., 2014)

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

Hierarchical K-Classifiers (Lobel et al., 2014)

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

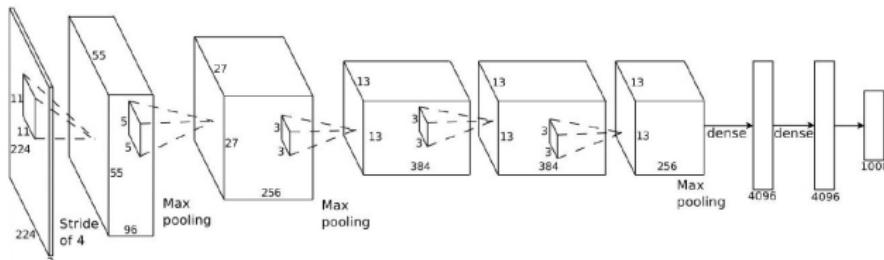
$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

Hierarchical K-Classifiers (Lobel et al., 2014)

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$

$$\min_{\{Z, W, \alpha\}} \frac{1}{2} \sum_{j=1}^L \|\alpha_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + \frac{1}{2} \sum_{j=1}^K \|w_j\|_2^2 + C \sum_{i=1}^M \xi_i$$



A Research Path

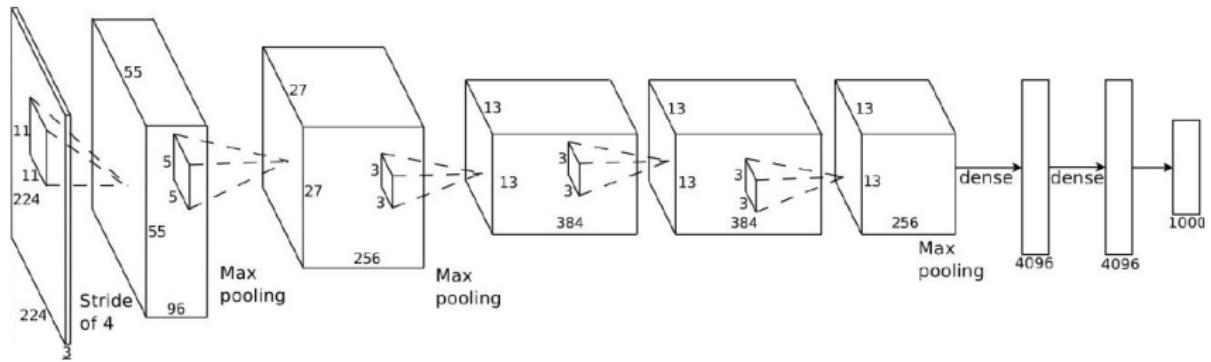
- Template matching
- Geometric models
- Machine learning: classifiers + handcrafted features
- Part based models (hierarchical model)
- Learning parts: poselets, KCL
- Learning parts and classifiers
- Learning features and everything: deep learning
- . . . ?

what_do_you_see() ?

Current Best Solution:

Deep Learning

Convolutional Neural Networks



CNNs and Learning of Semantic Spaces

First column: query images from ILSVRC-2010 test set.

Next columns: nearest neighbors in ImageNet using AlexNet FC6-space.

