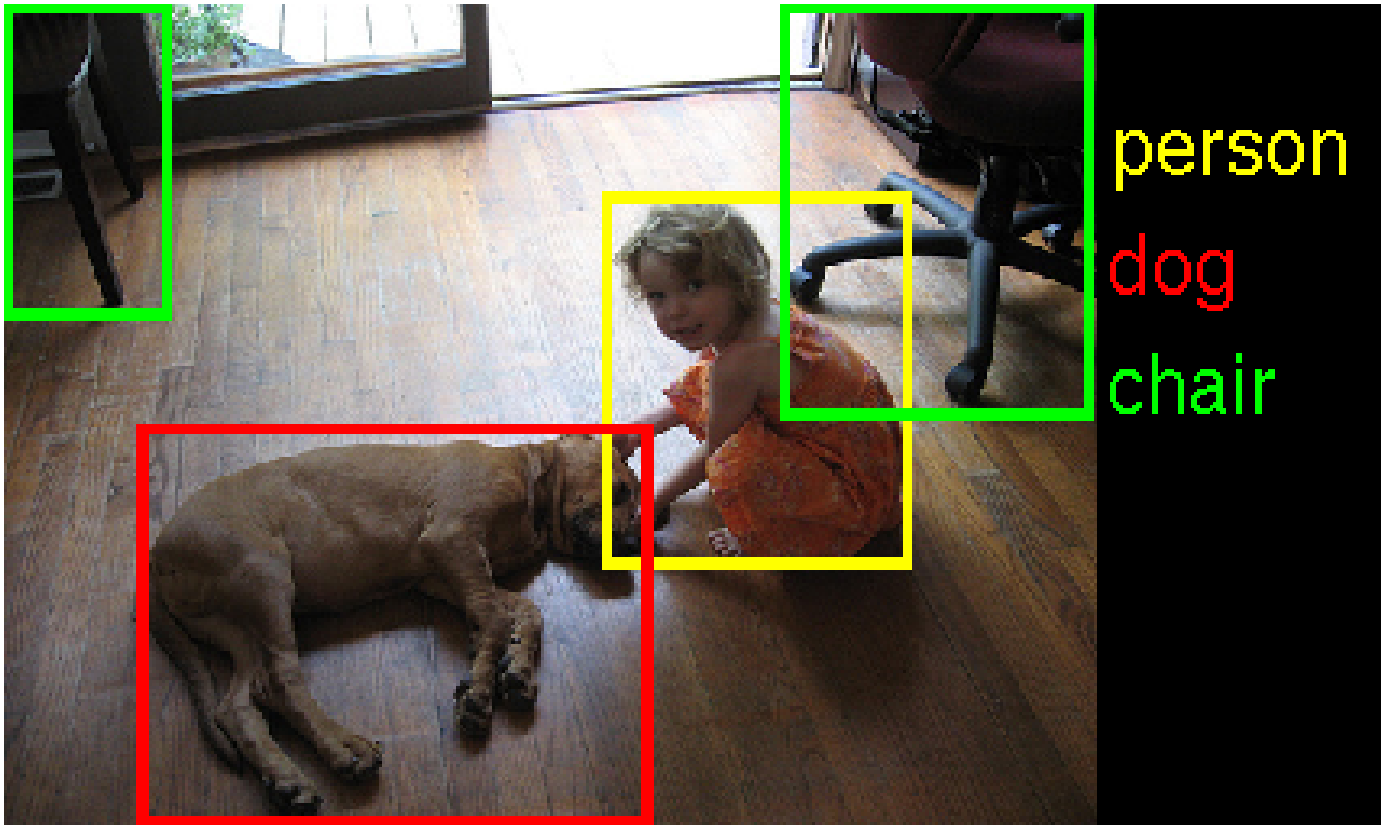
Overview of Computer Vision

CENTER FOR
Brains
Minds+
Machines

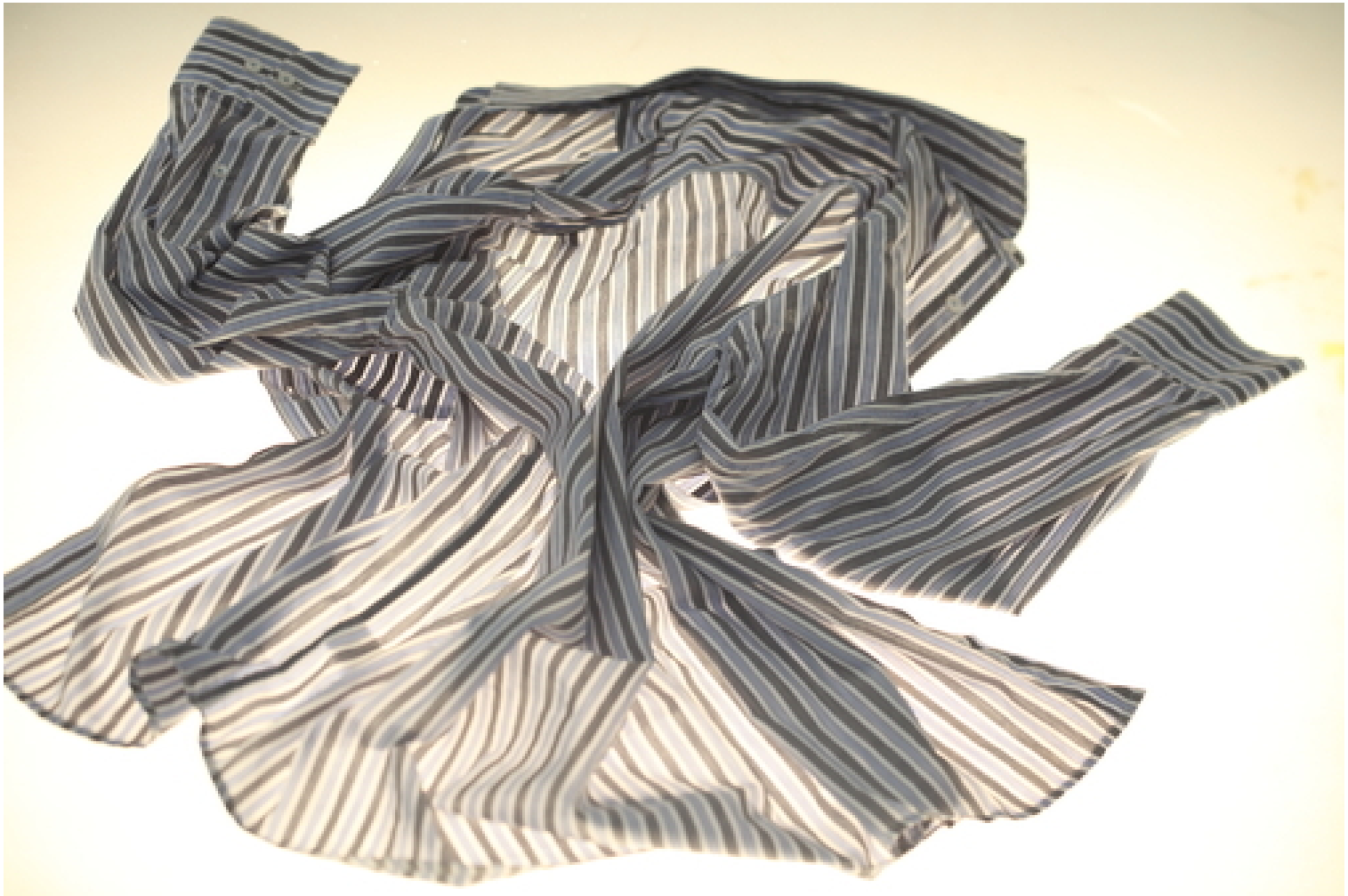
Andrei Barbu

Computer Vision

Computer Vision



Human vision



Microsoft

CaptionBot

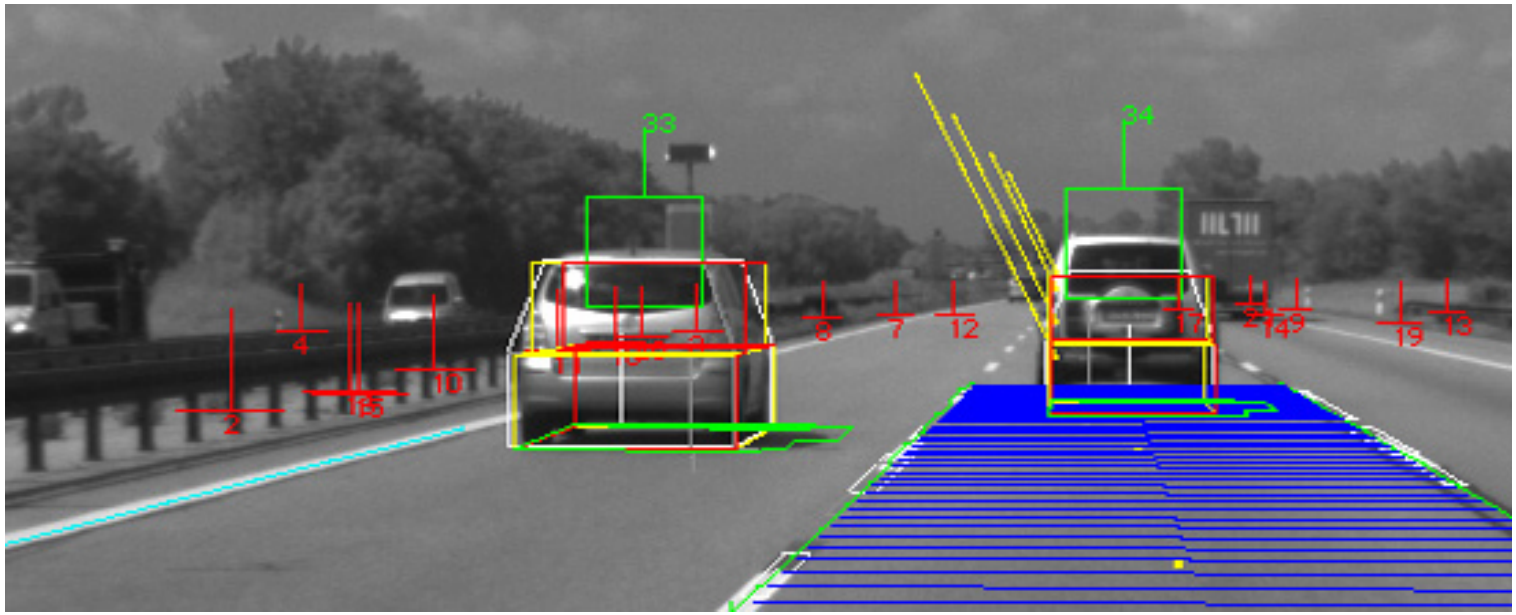
I think it's a hand holding a half eaten apple.

How did I do?

★★★★★

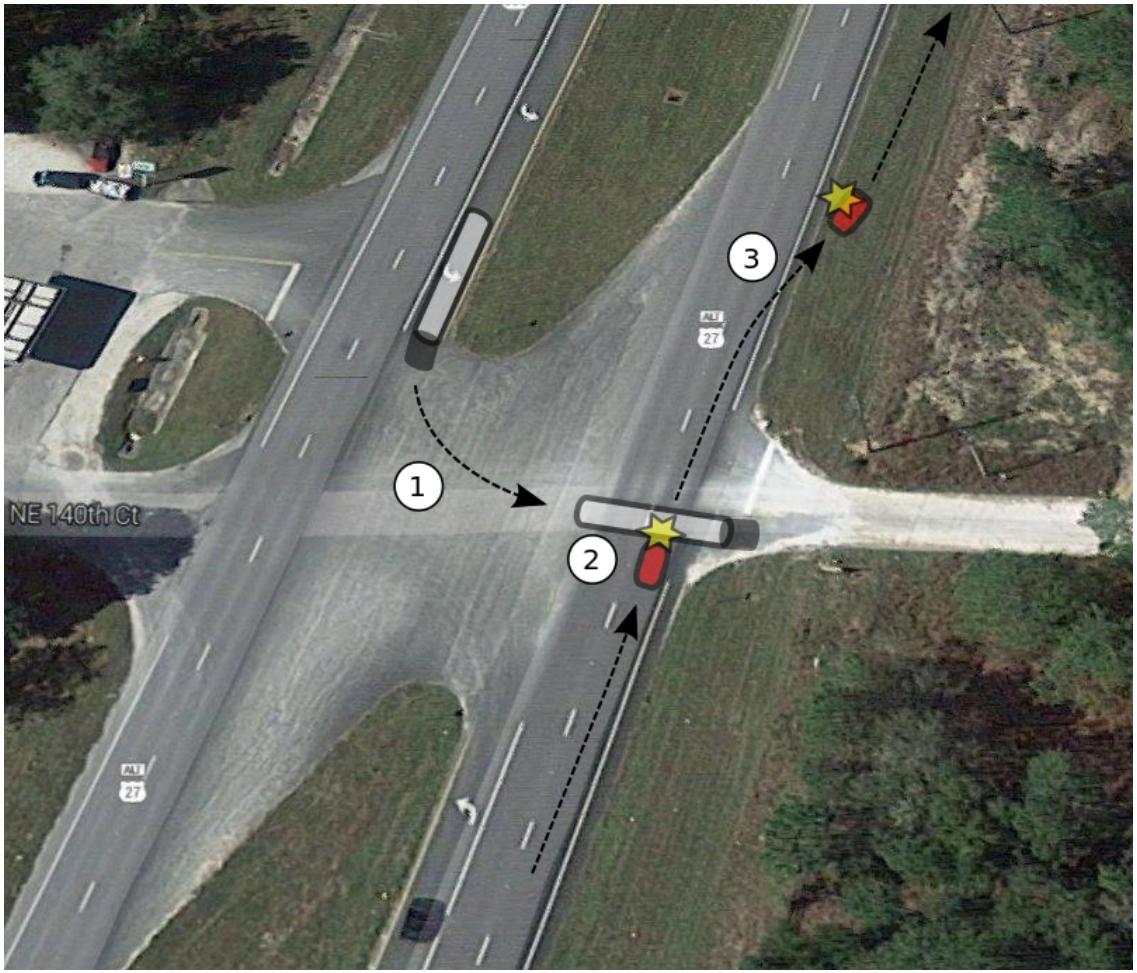
Captionbot June 2017

Car Vision



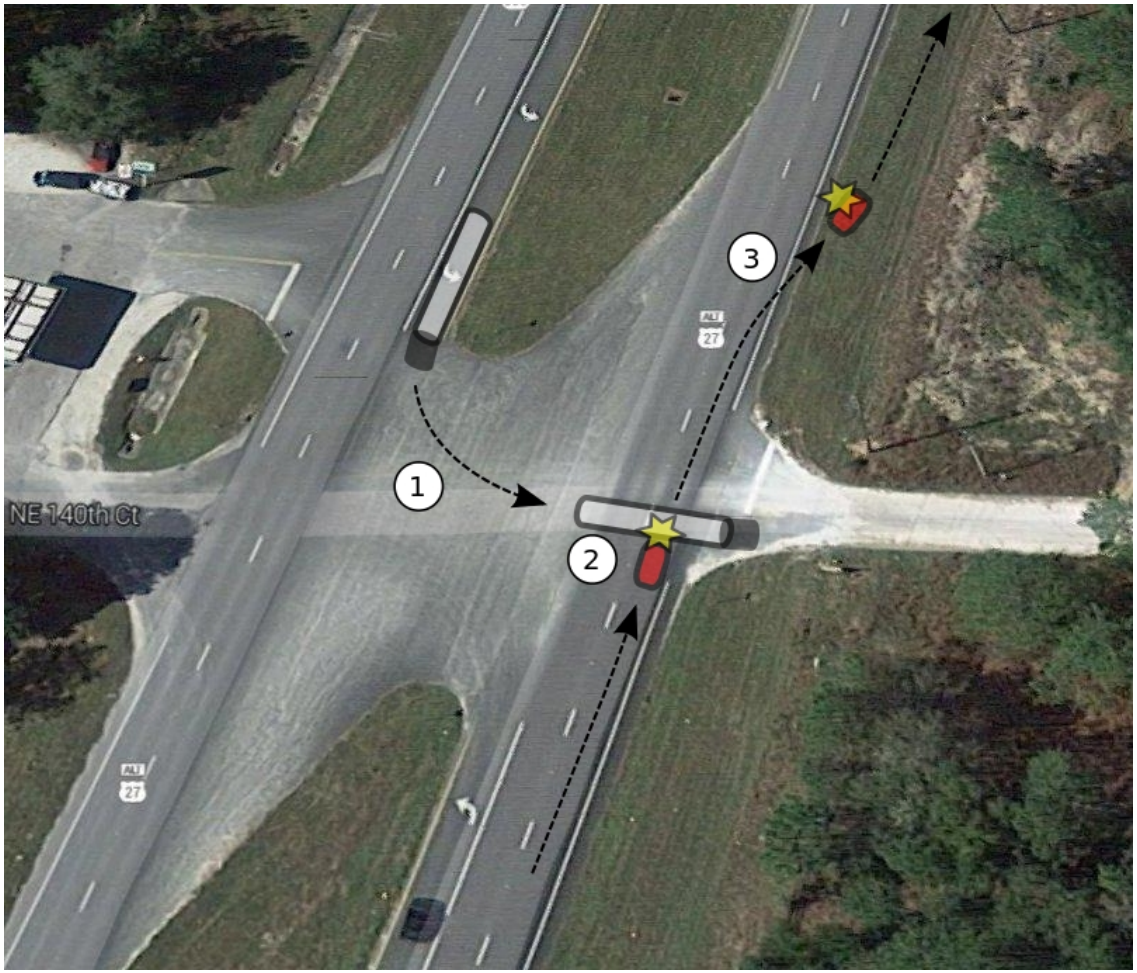
Computer vision failures

Computer vision failures



Computer vision failures

NHTSA: Neither Autopilot nor the driver noticed the white side of the tractor-trailer against a brightly lit sky, so the brake was not applied.



Vision

Vision

Vision

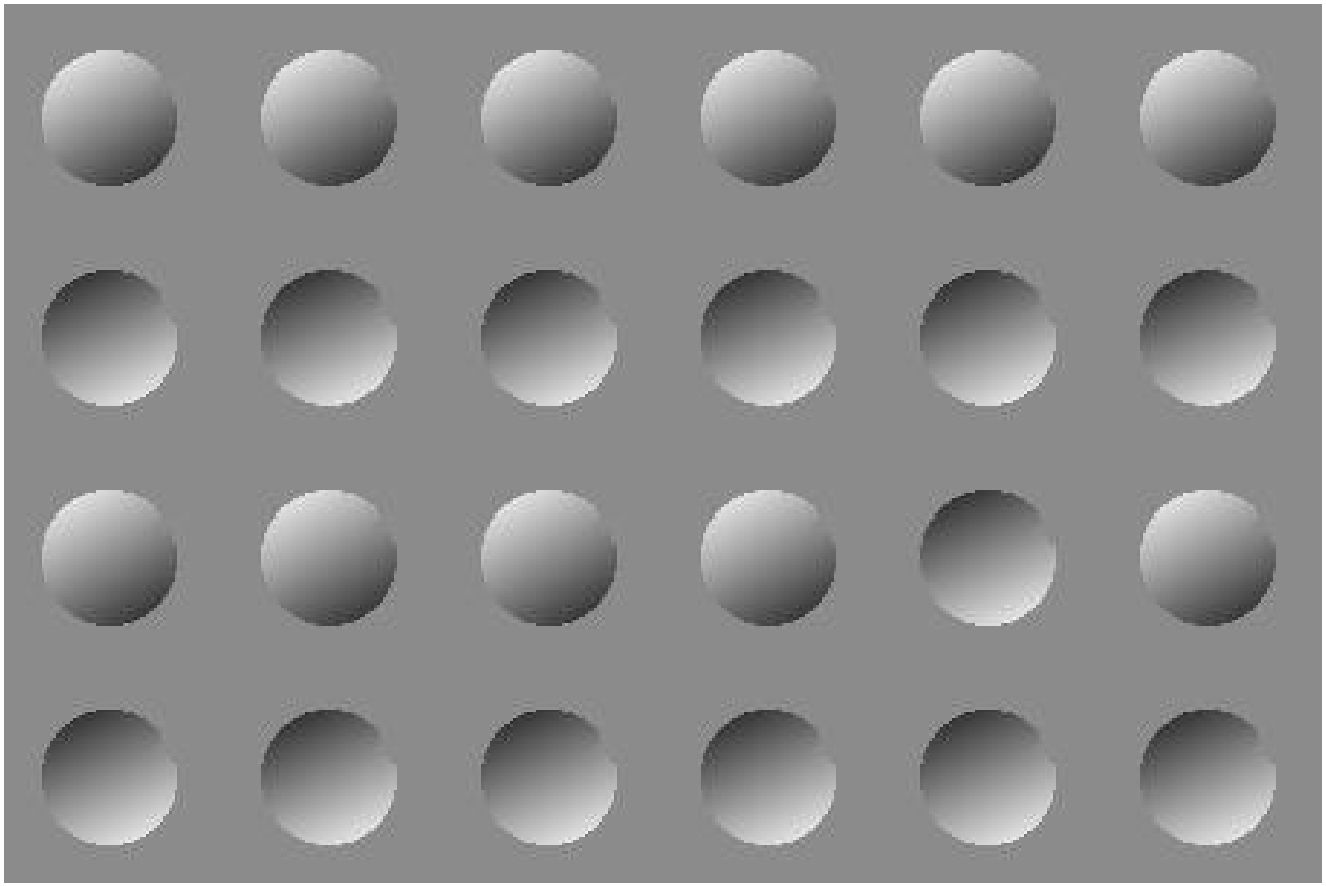
I can't really describe the picture but I do see indoor, table, room.
Microsoft CaptionBot, May 6 2018

Vision

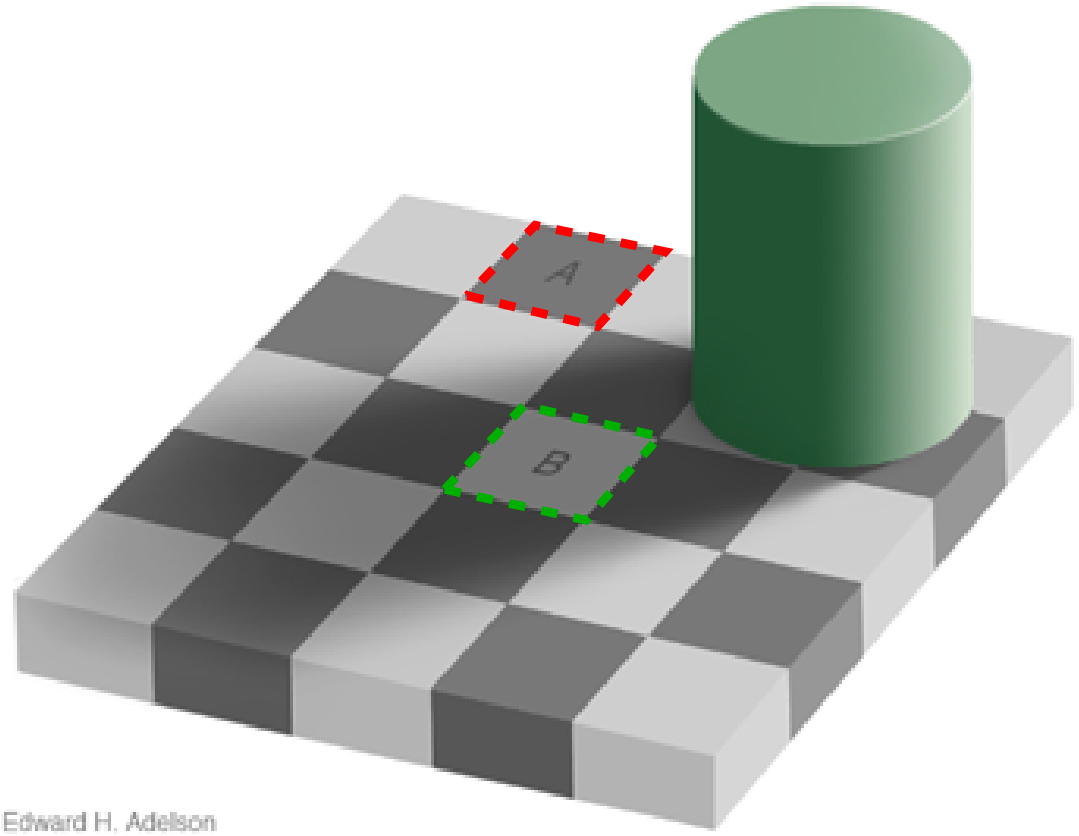
Vision

illumination

Illumination

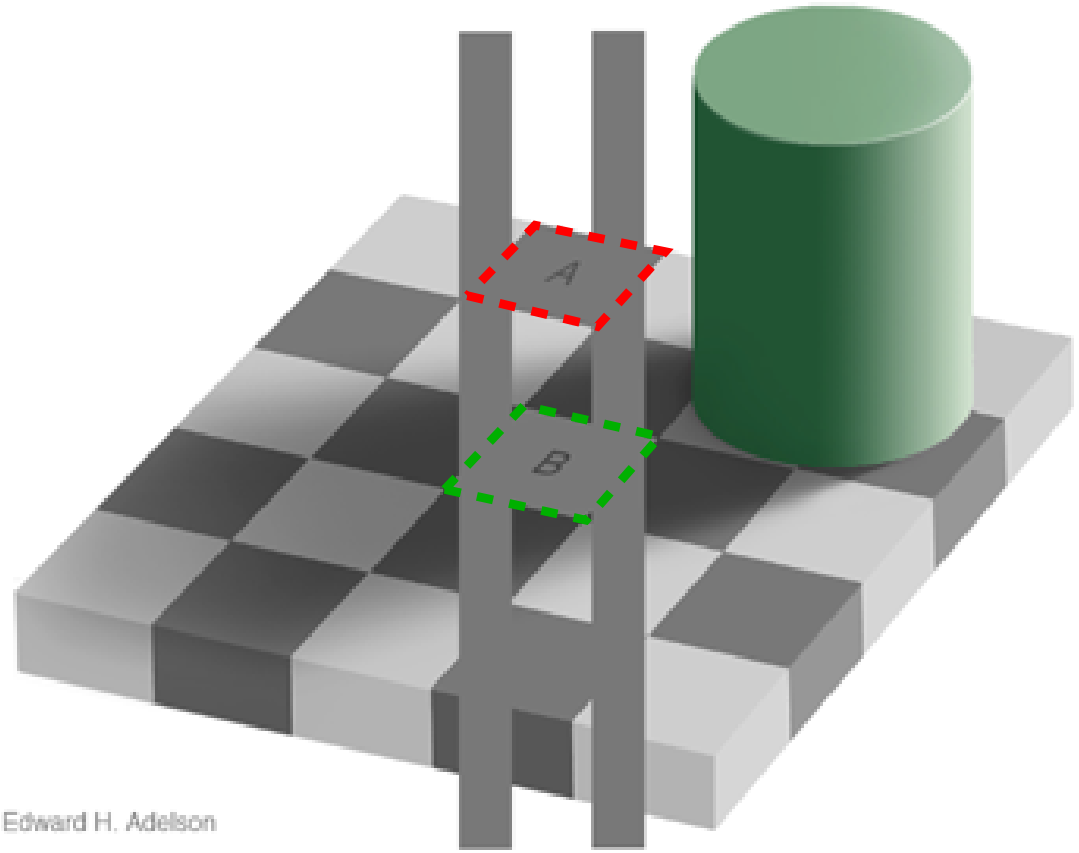


Illumination



Edward H. Adelson

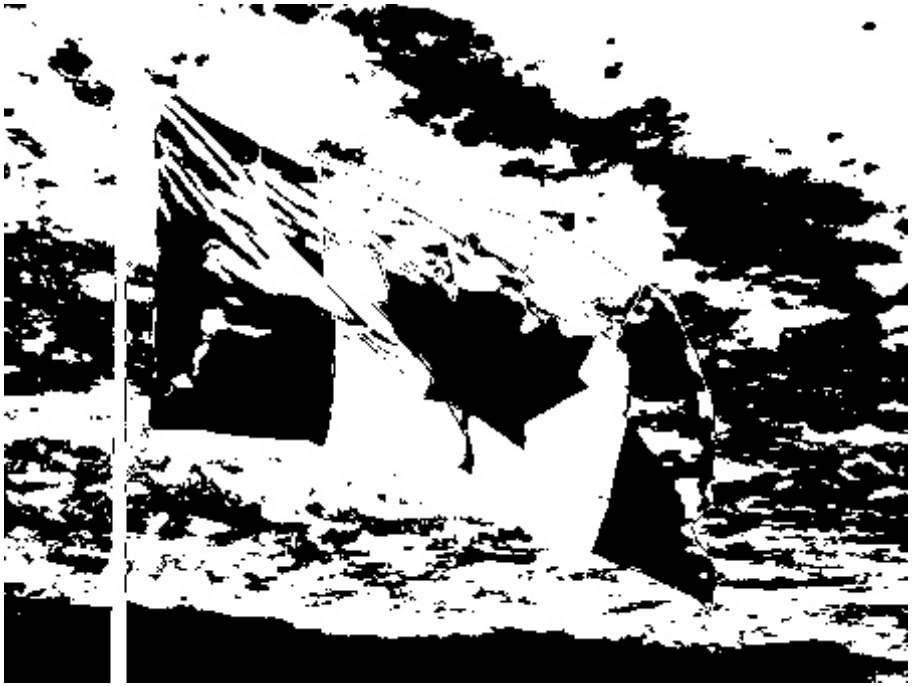
Illumination



Vision

illumination

color



Vision

illumination

color

attention

Vision

illumination

color

attention

scene and object organization

Vision

illumination

color

attention

scene and object organization

object categorization

Vision

illumination

color

attention

scene and object organization

object categorization

object shape and structure

Vision

illumination

color

attention

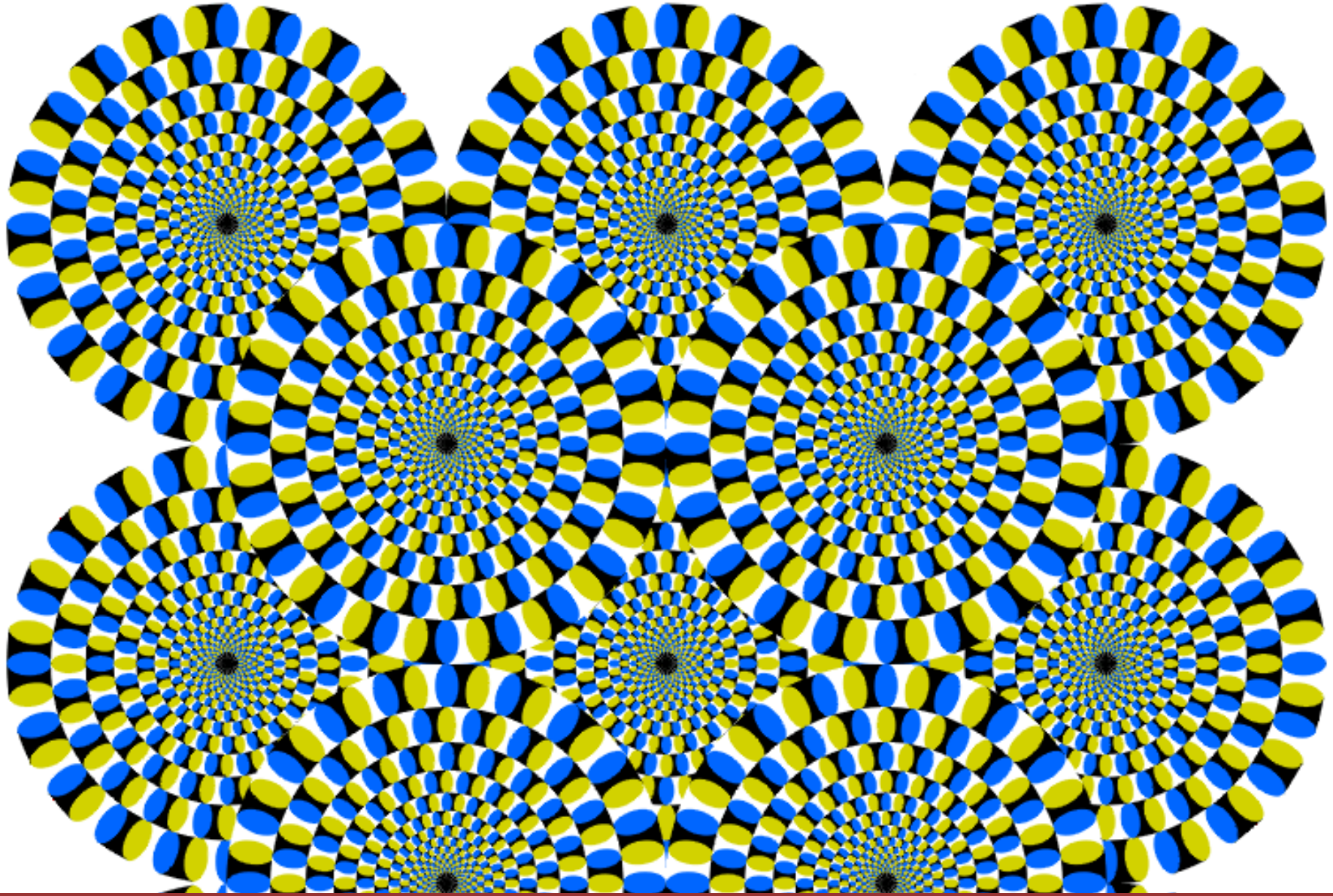
scene and object organization

object categorization

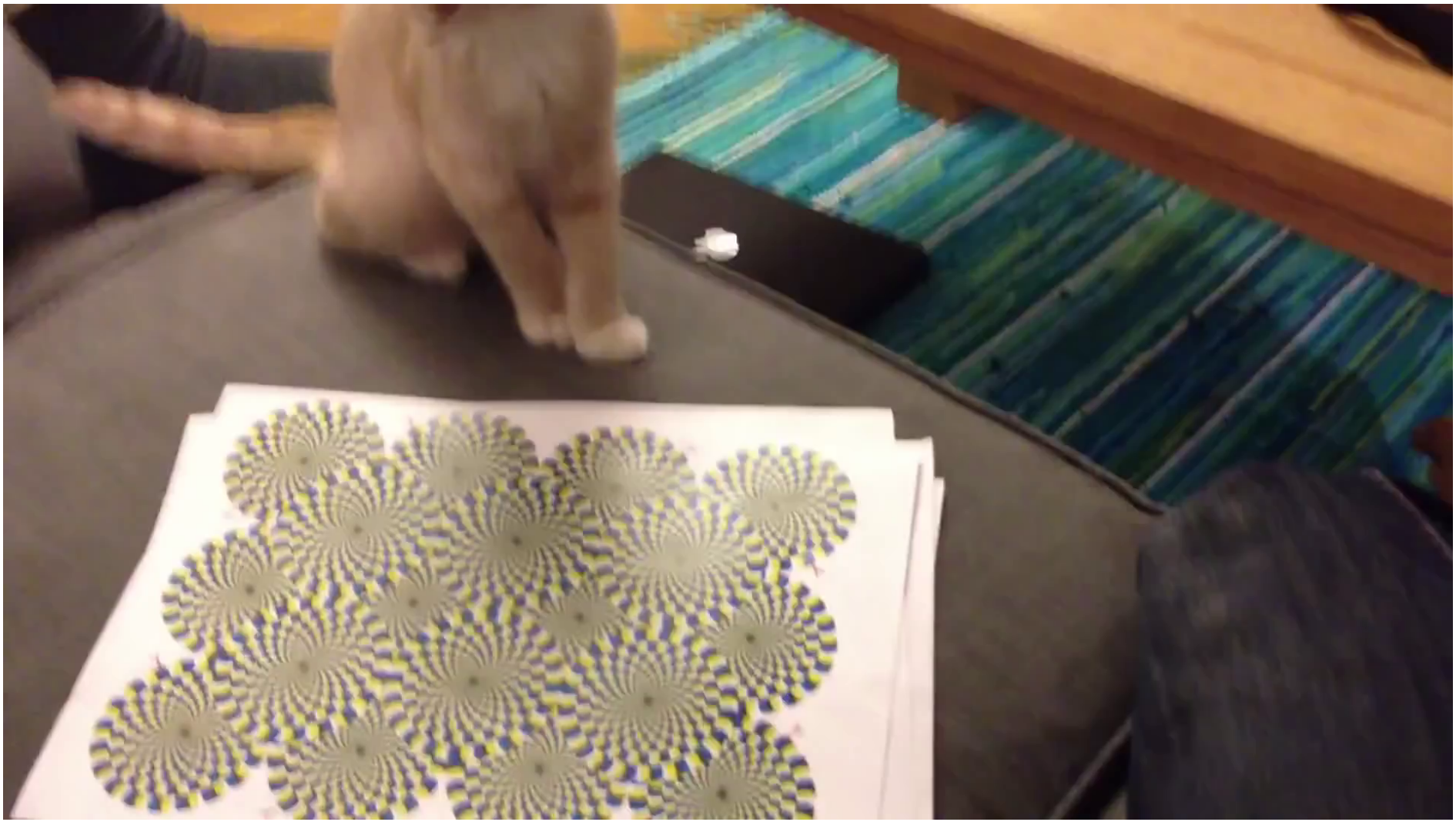
object shape and structure

motion

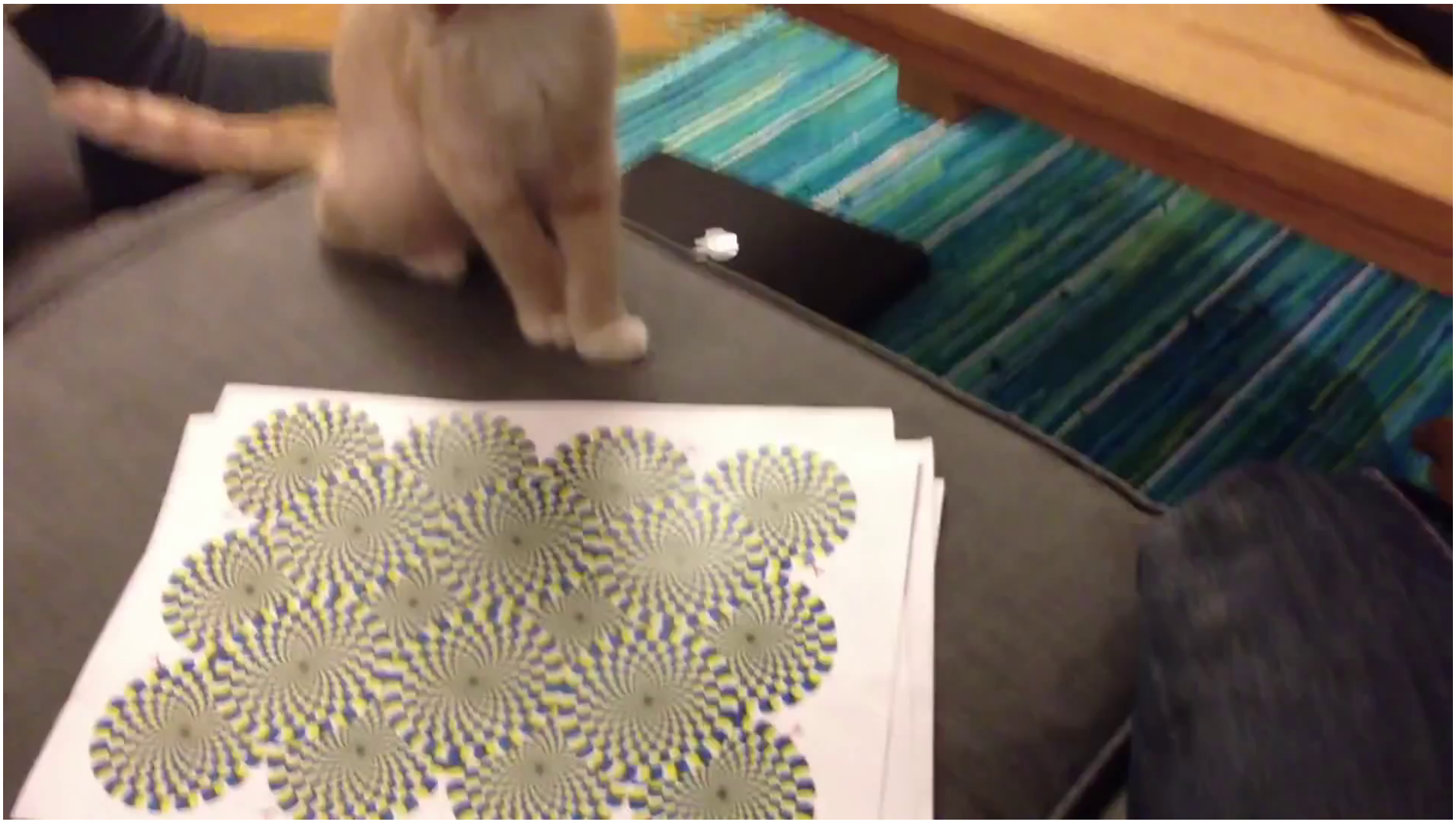
Motion



Motion



Motion



Vision

illumination

color

attention

scene and object organization

object categorization

object shape and structure

motion

Vision

illumination

color

attention

scene and object organization

object categorization

object shape and structure

motion

actions

Vision

illumination

color

attention

scene and object organization

object categorization

object shape and structure

motion

actions

memory

Vision

illumination

color

attention

scene and object organization

object categorization

object shape and structure

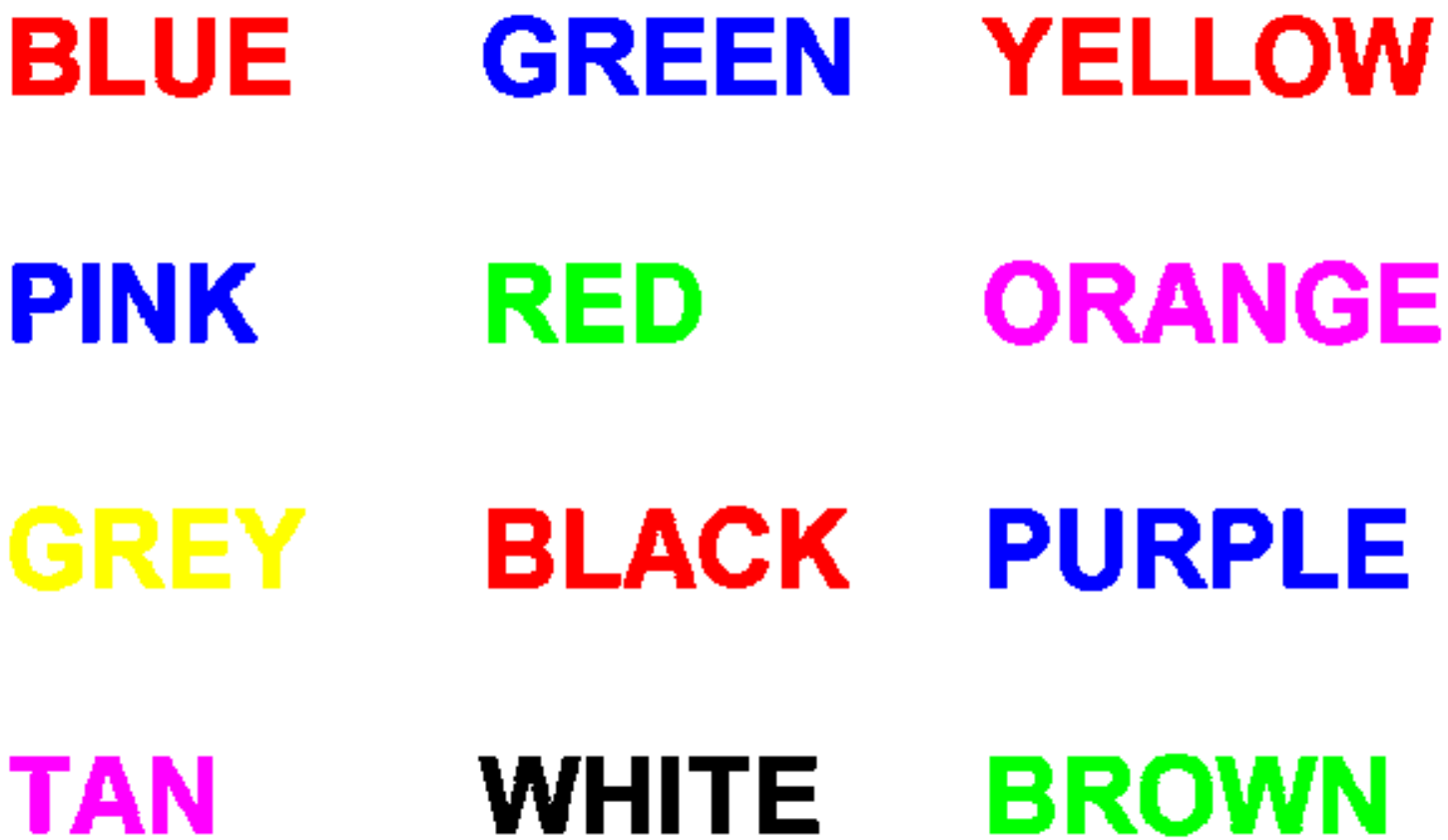
motion

actions

memory

social interactions

Color & Language



The origins of CV: image processing

The origins of CV: image processing

blur

The origins of CV: image processing

blur

sharpening

The origins of CV: image processing

blur

sharpening

denoising

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

compression

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

compression

white balance

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

compression

white balance

etc.

1970
1980
1990
2000
Digital image processing
Blocks world, line labeling
Generalized cylinders
Pictorial structures
Stereo correspondence
Intrinsic images
Optical flow
Structure from motion
Image pyramids
Scale-space processing
Shape from shading, texture, and focus
Physically-based modeling
Regularization
Markov Random Fields
Kalman filters
3D range data processing
Projective invariants
Factorization
Physics-based vision
Graph cuts
Particle filtering
Energy-based segmentation
Face recognition and detection
Subspace methods
Image-based modeling and rendering
Texture synthesis and inpainting
Computational photography
Feature-based recognition
MRF inference algorithms
Category recognition
Learning

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

compression

white balance

etc.

The origins of CV: image processing

blur

sharpening

denoising

morphological operations

contrast enhancement

correcting distortion

compression

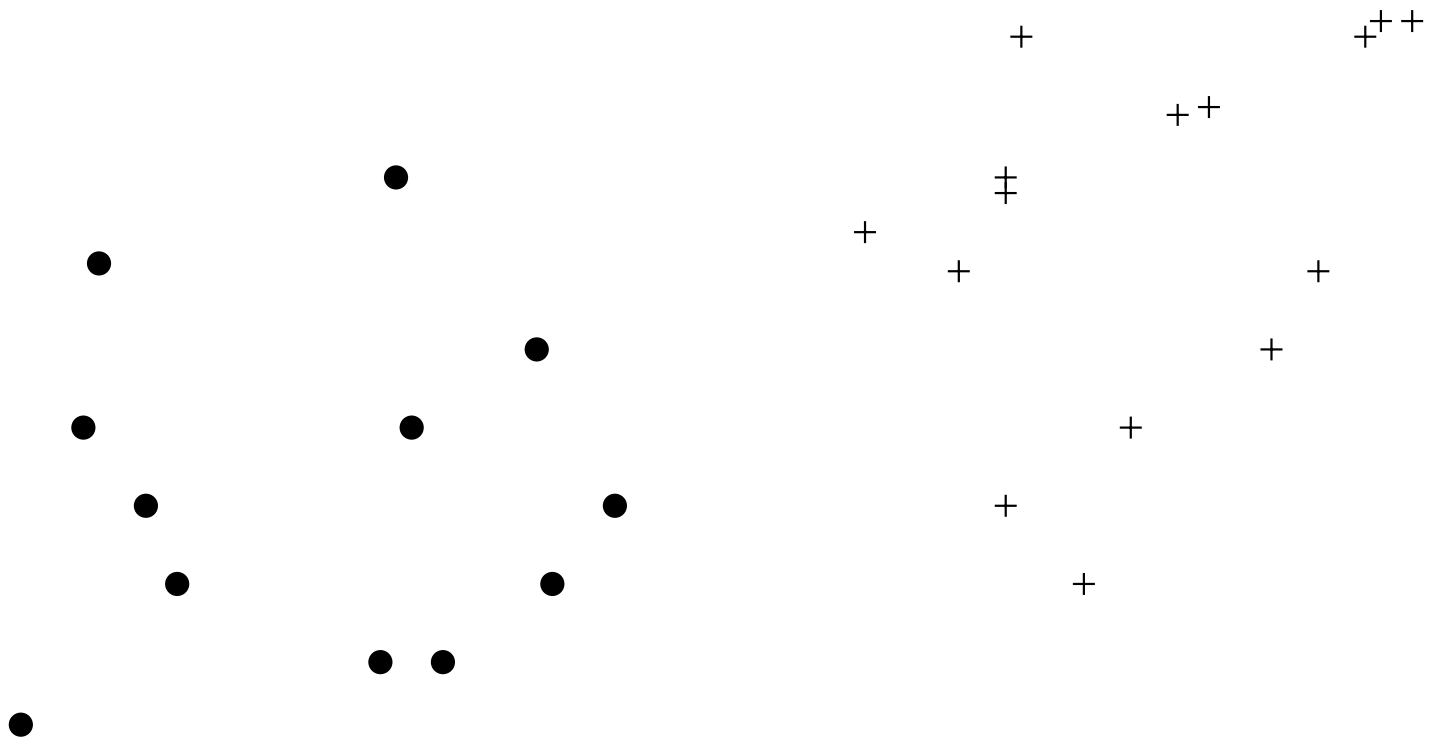
white balance

etc.

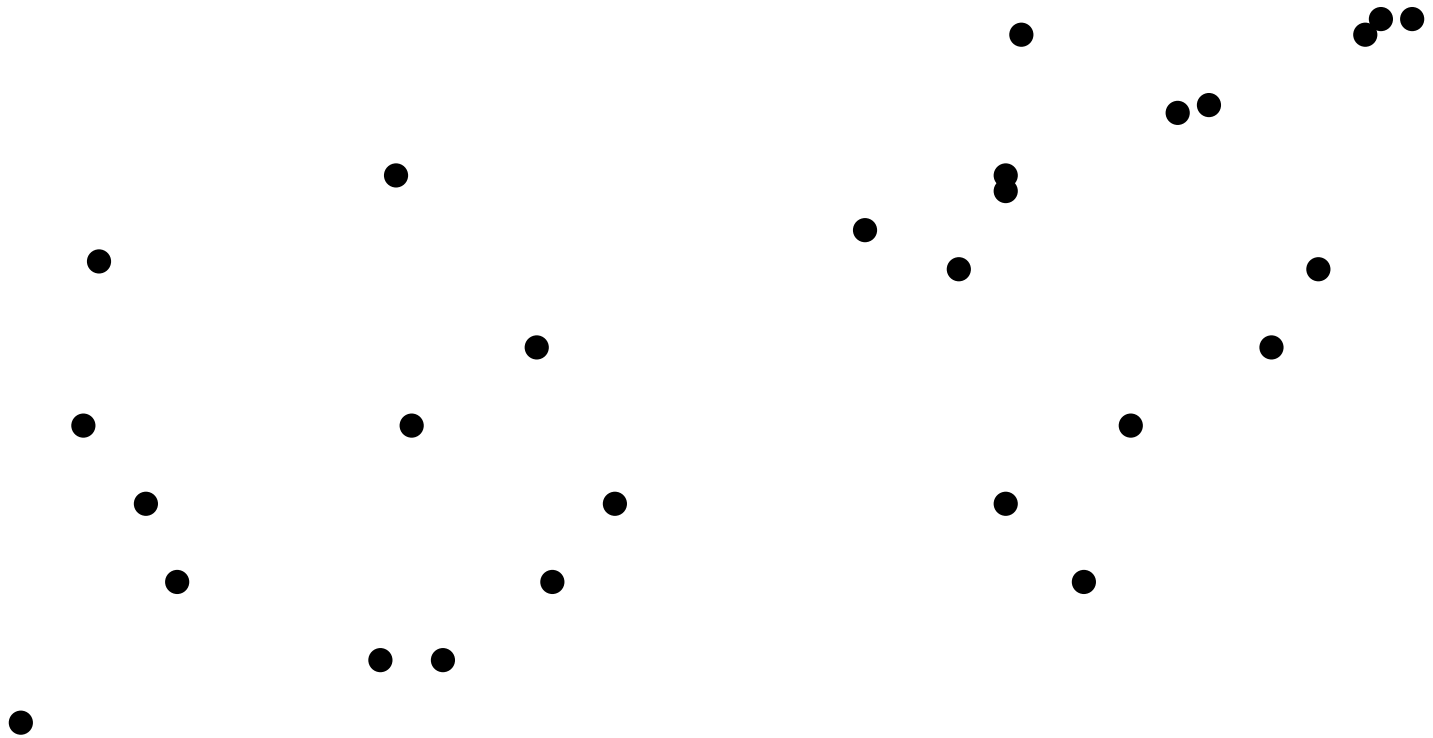
Baby machine learning

Classifiers

Classifiers

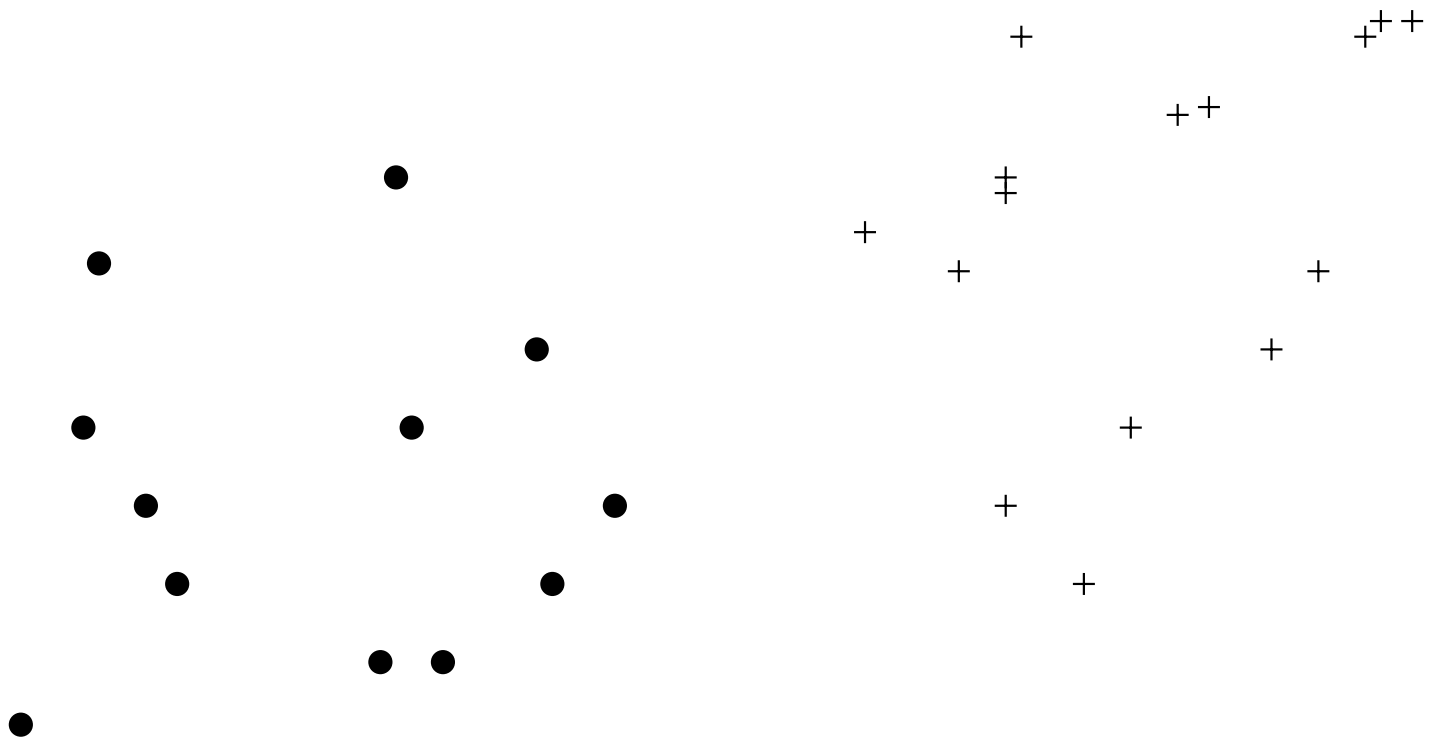


Classifiers

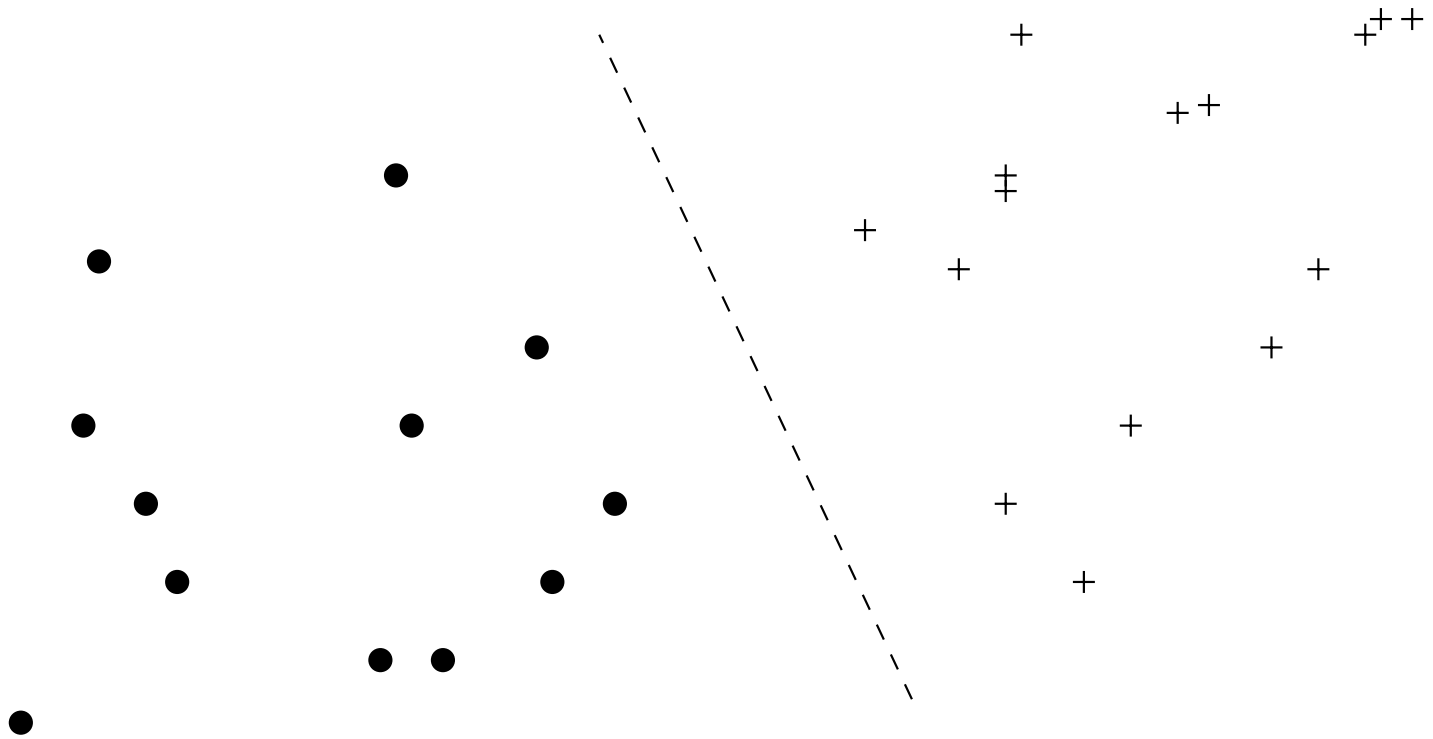


Linear classifiers

Linear classifiers

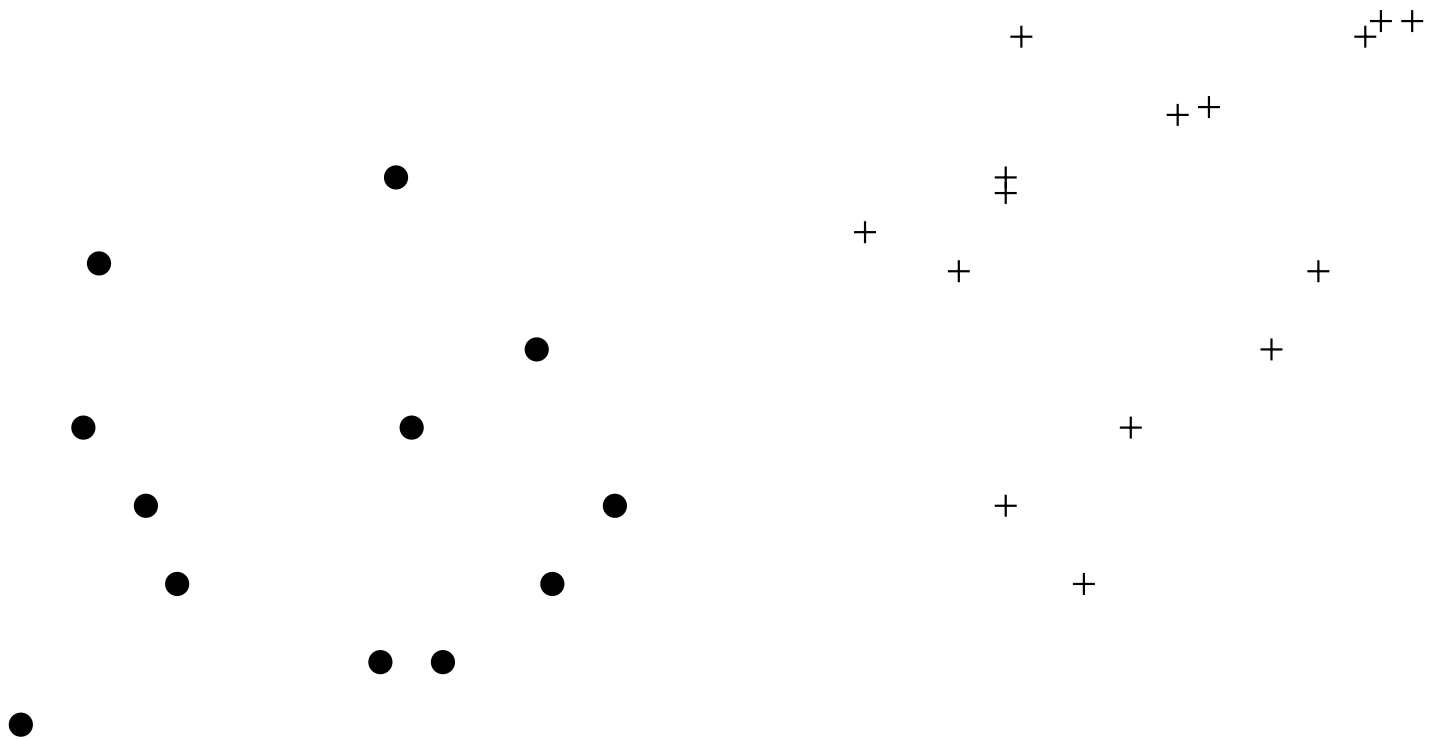


Linear classifiers

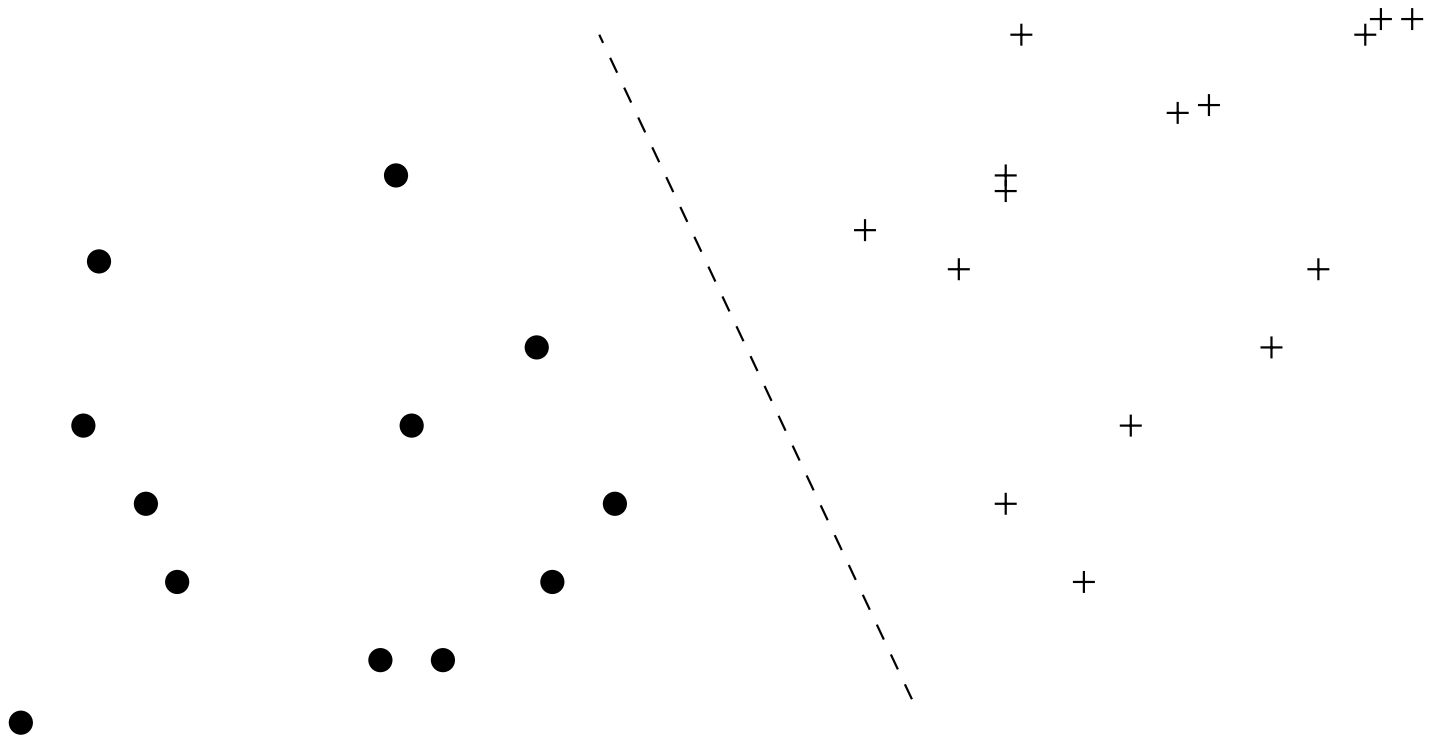


Linear classifiers

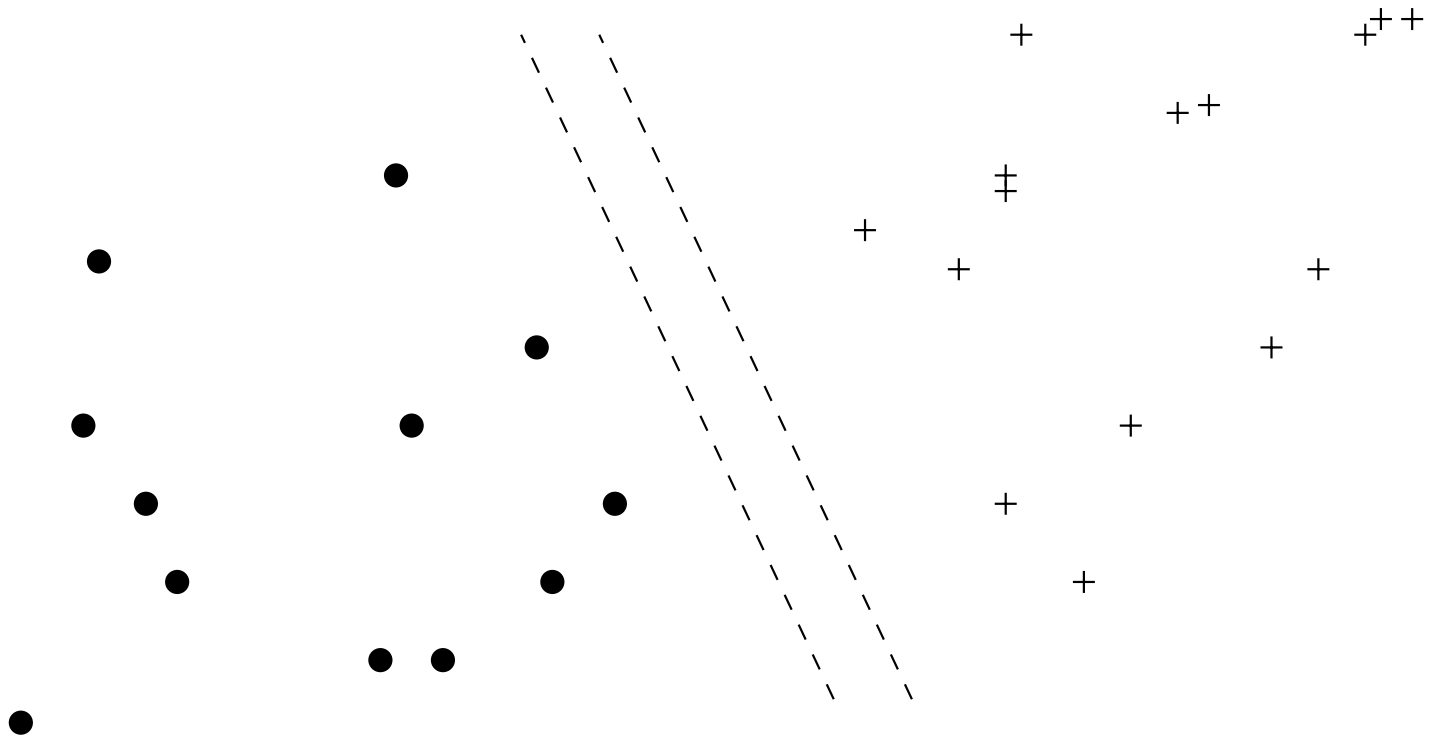
Linear classifiers



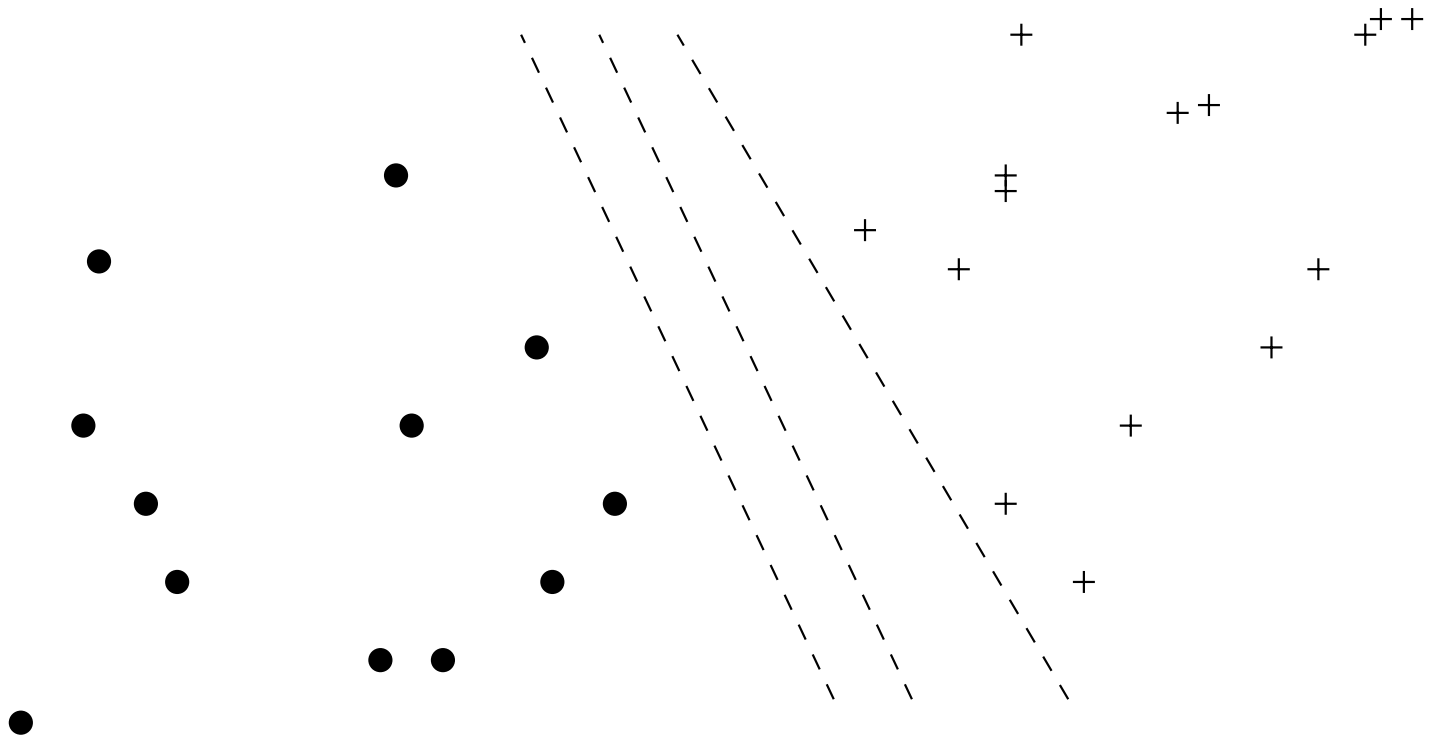
Linear classifiers



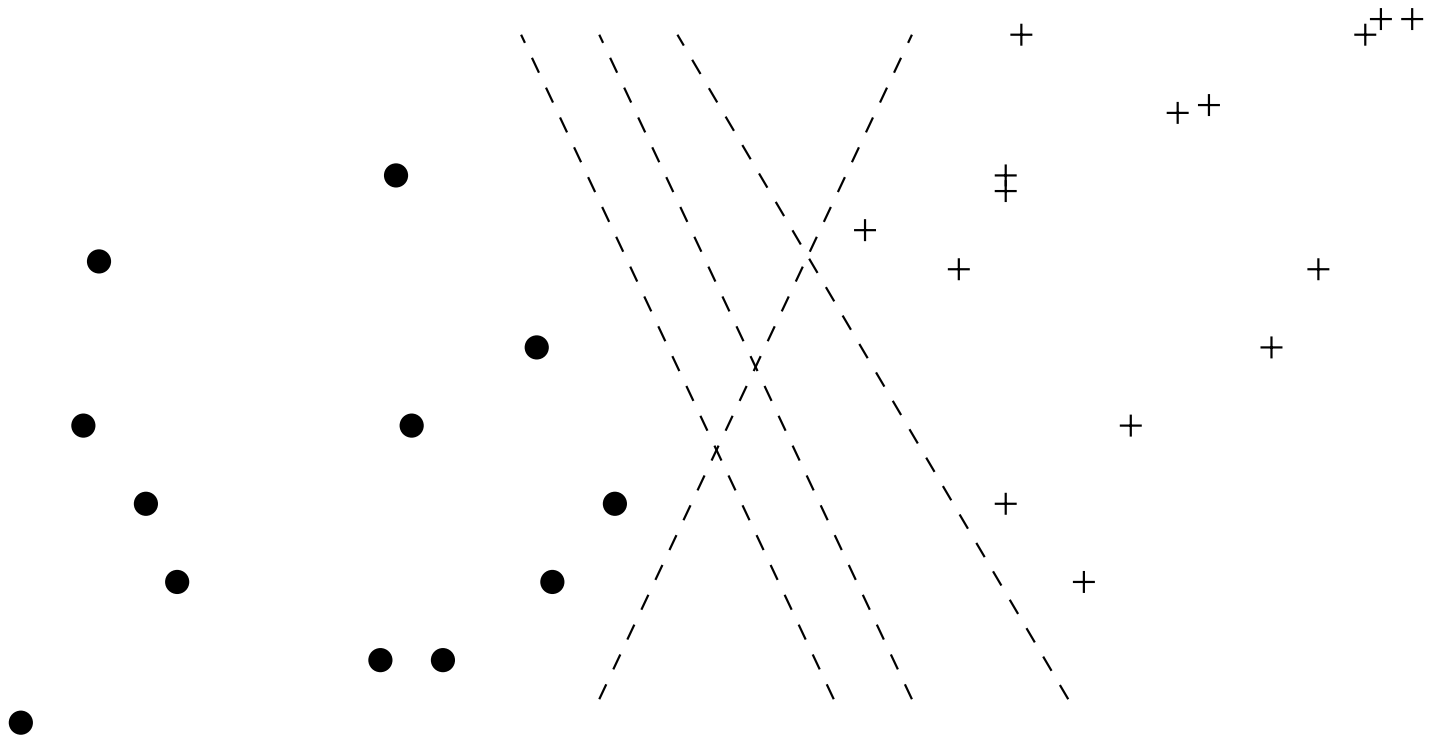
Linear classifiers



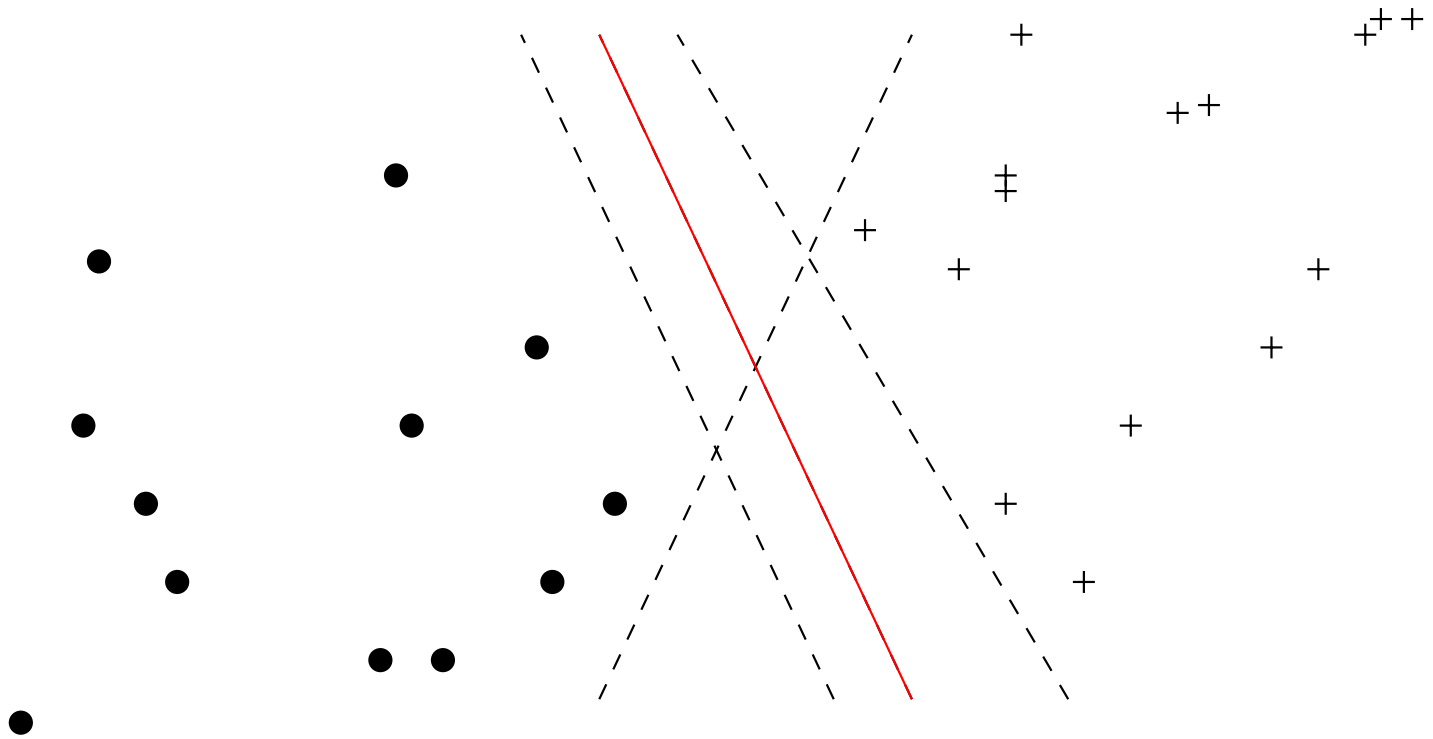
Linear classifiers



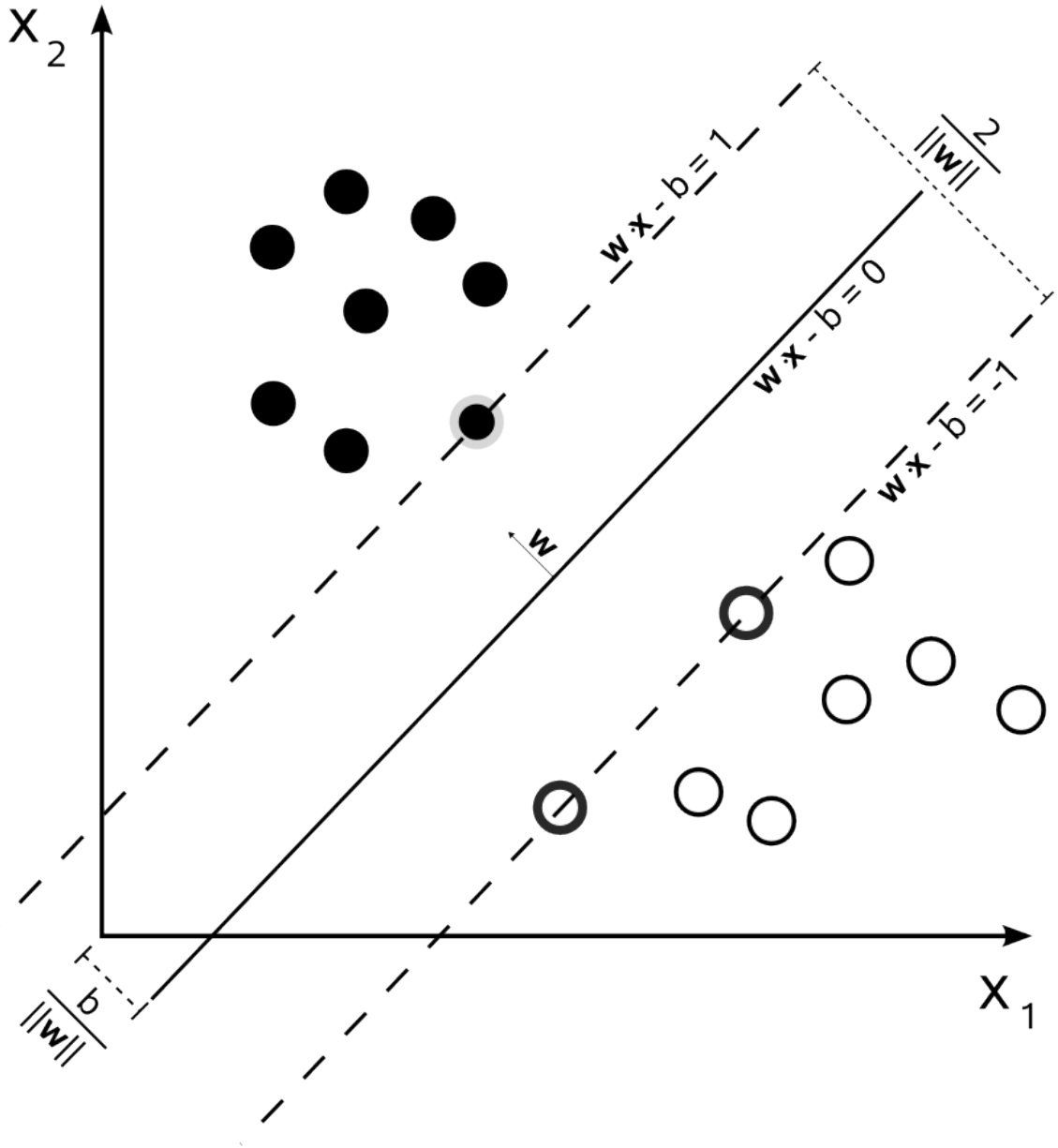
Linear classifiers



Linear classifiers



Support Vector Machines



Support Vector Machines

Support Vector Machines

$$\{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\, y_i \in \{-1, 1\}\}_{i=1}^{n}$$

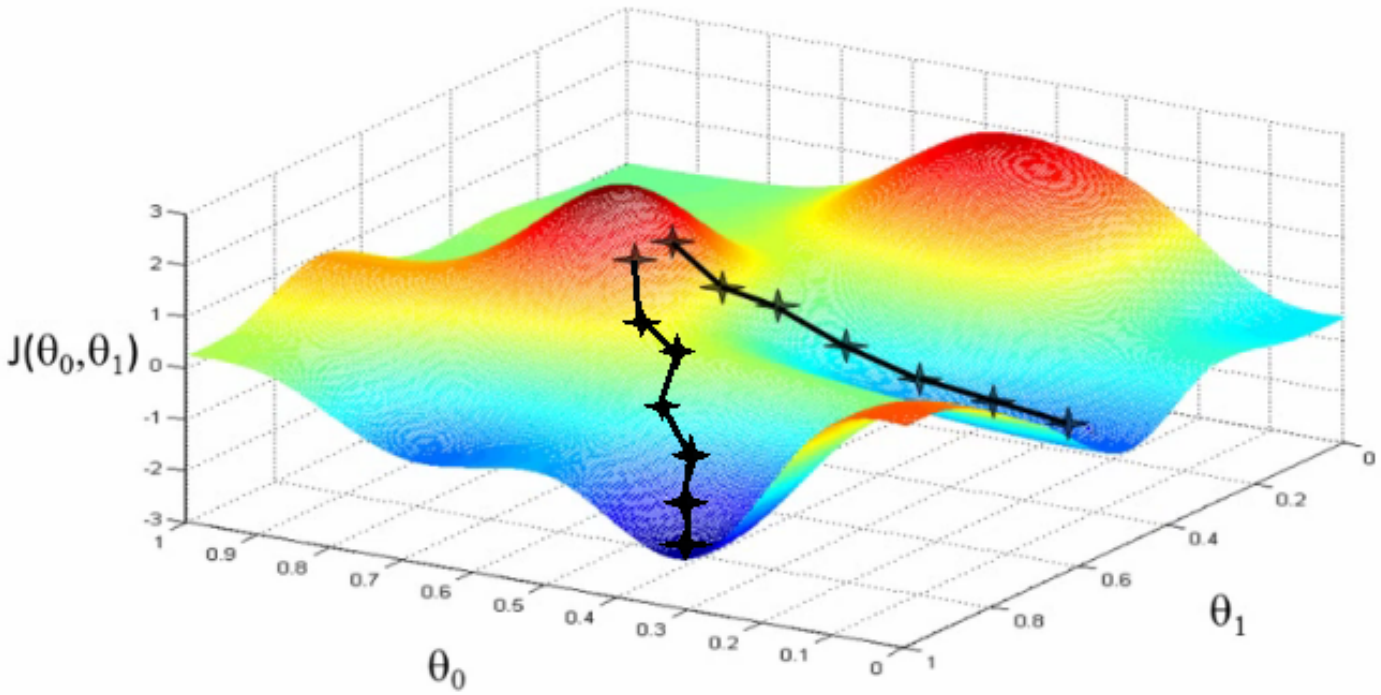
Support Vector Machines

$$\{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^p,\, y_i \in \{-1, 1\}\}_{i=1}^{n}$$

$$\arg\min_{(\mathbf{w},b)} \frac{1}{2}\|\mathbf{w}\|^2$$
such that for each point i
$$y_i(\mathbf{w} \cdot \mathbf{x_i} - b) \geq 1$$

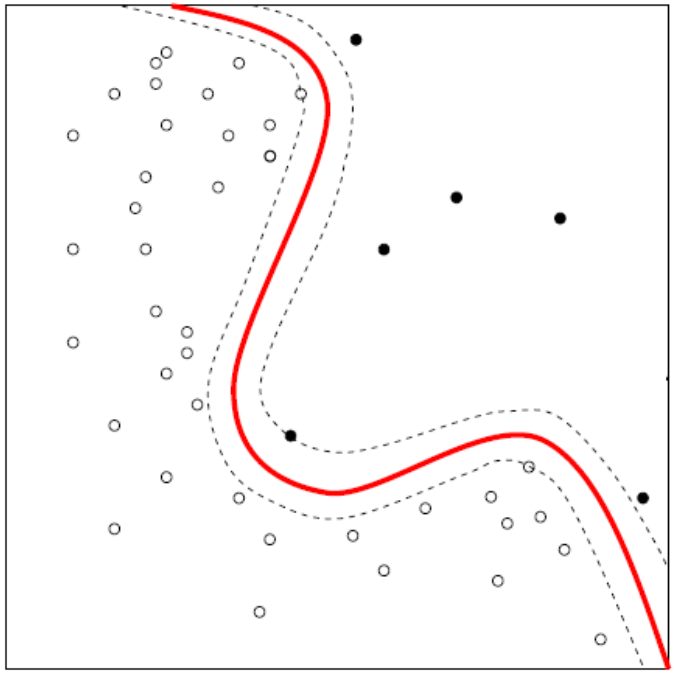
Gradient descent

Gradient descent

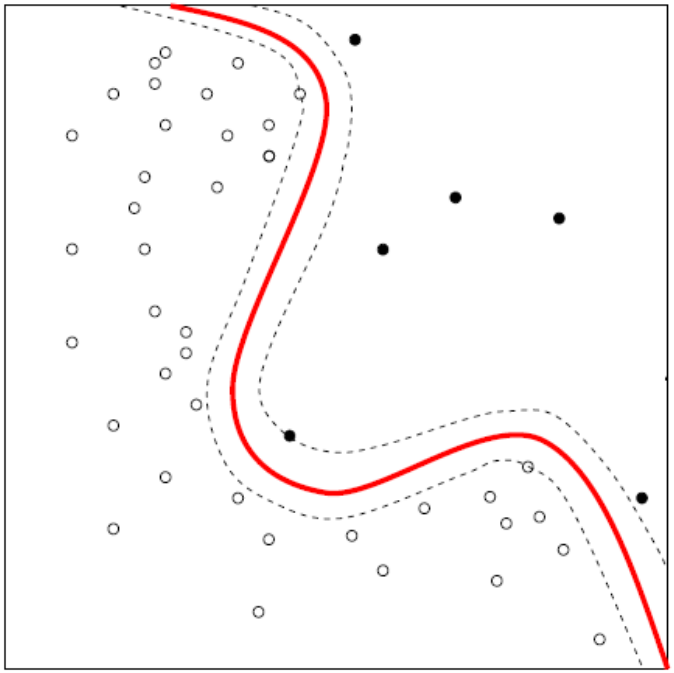


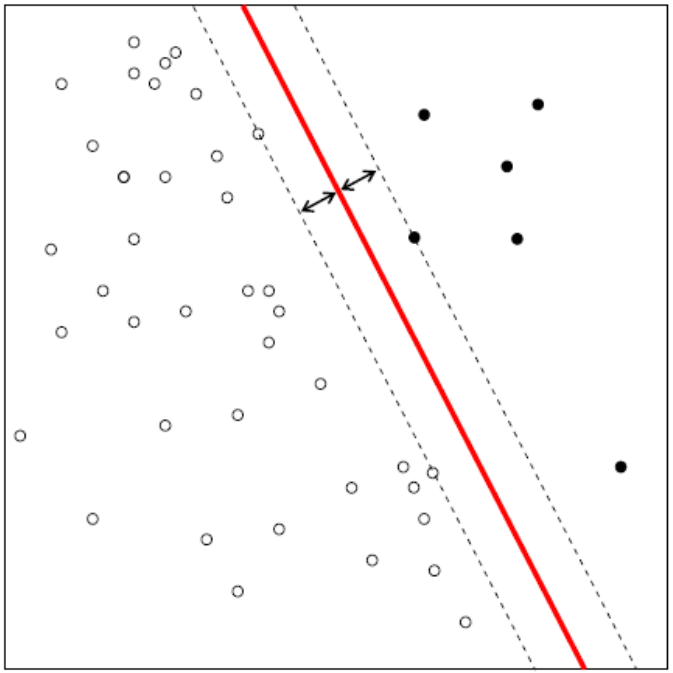
Linear classifiers and kernels

Linear classifiers and kernels



Linear classifiers and kernels





Computer vision

Convolution

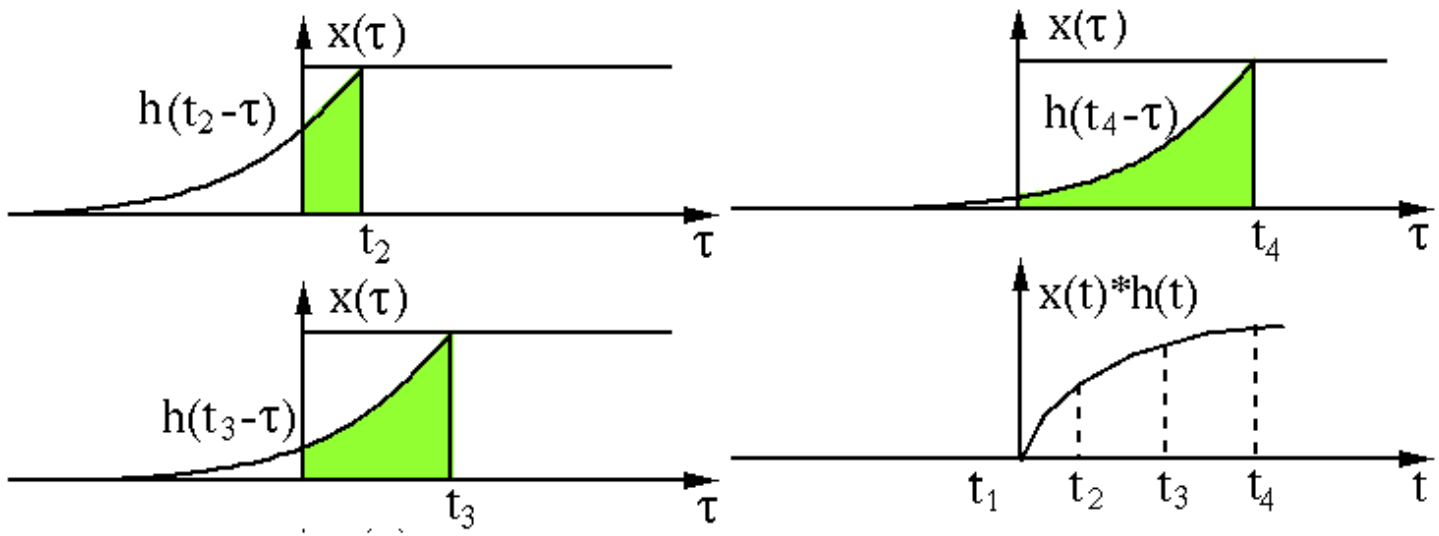
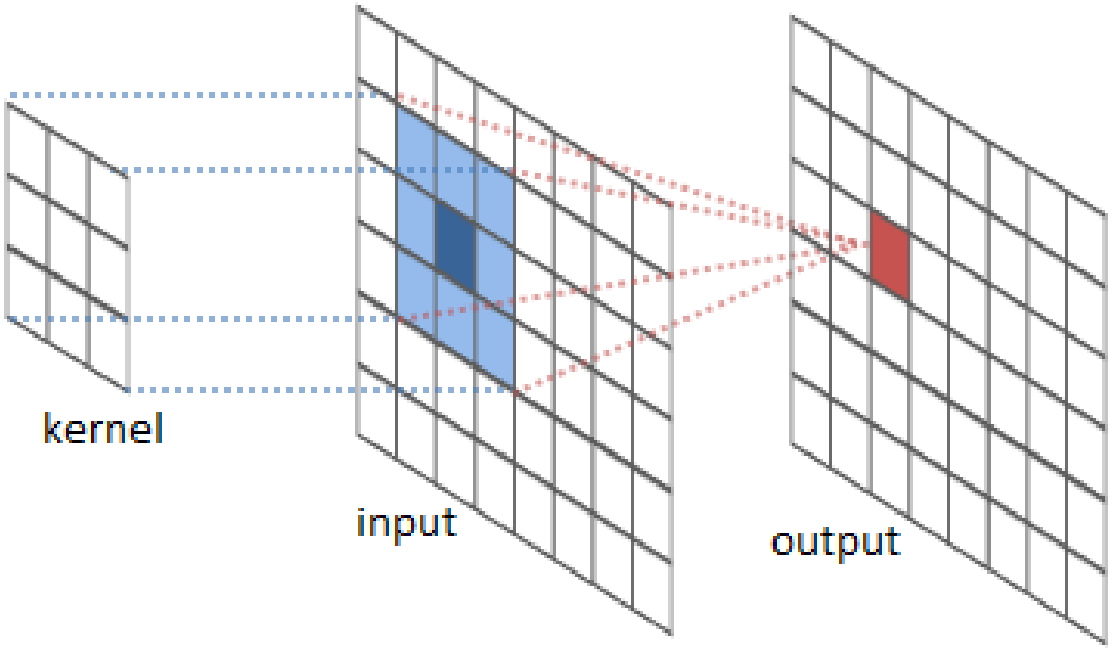
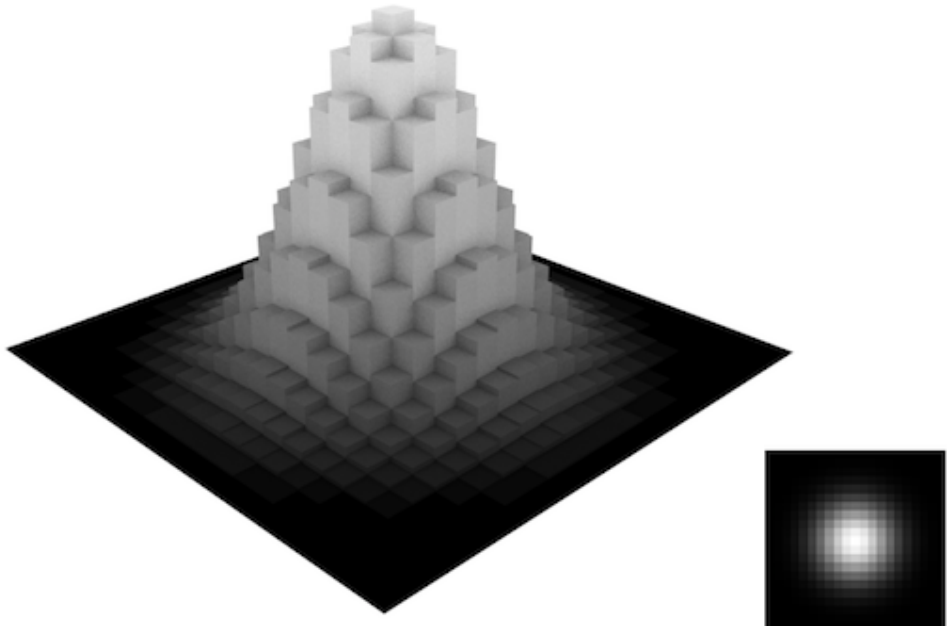


Image convolution



Gaussian blur kernel



Gaussian blur kernel

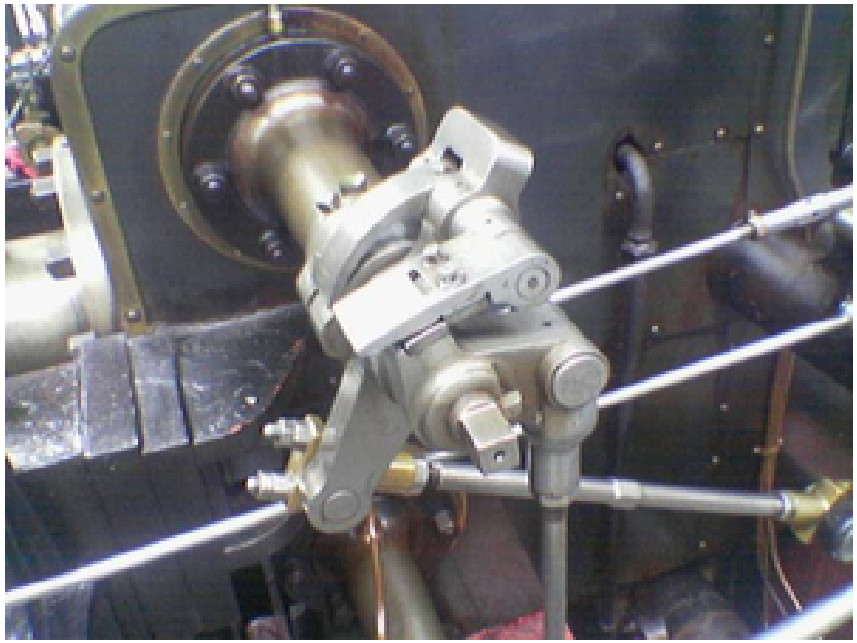
0	0	0	5	0	0	0
0	5	18	32	18	5	0
0	18	64	100	64	18	0
5	32	100	100	100	32	5
0	18	64	100	64	18	0
0	5	18	32	18	5	0
0	0	0	5	0	0	0

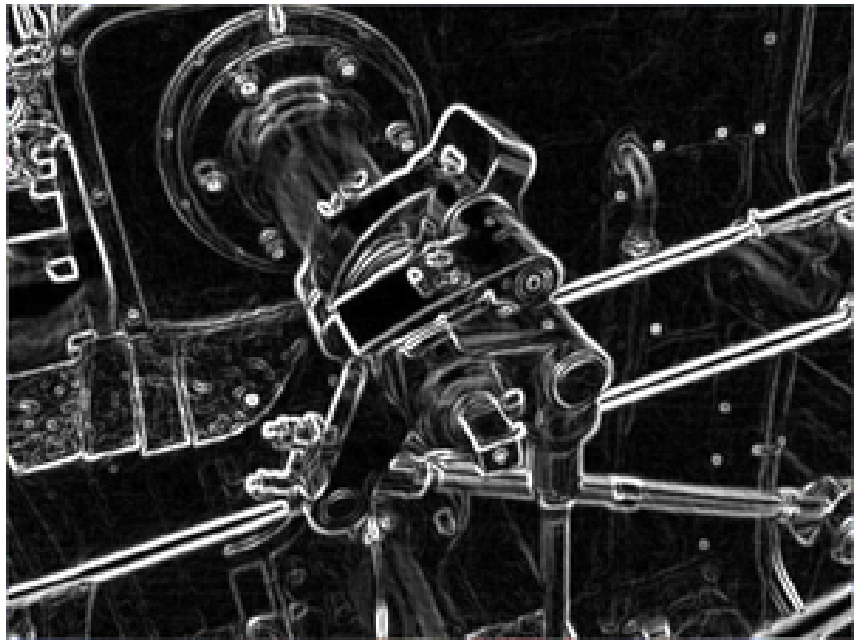
Gaussian blur

Early edge detection: Sobel

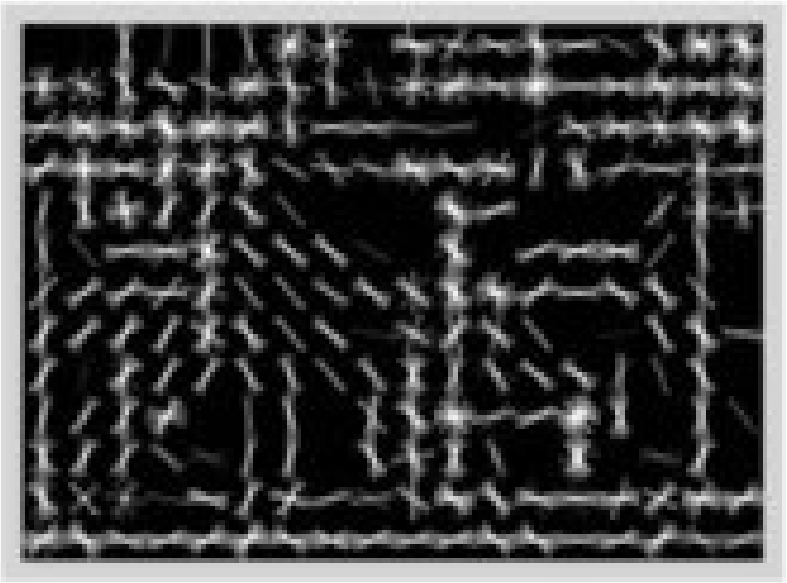
$$\begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \qquad \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$$

Early edge detection: Sobel

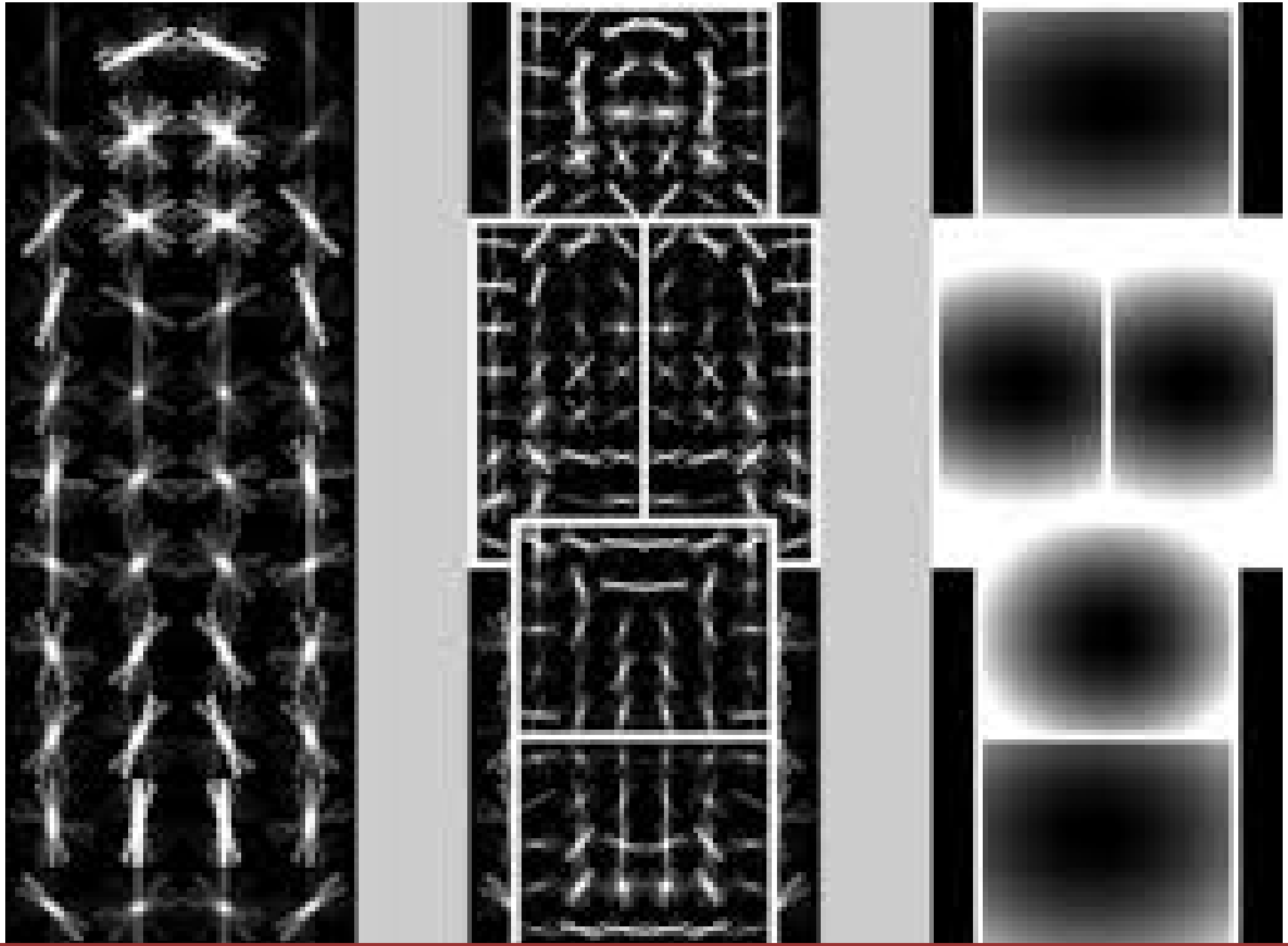




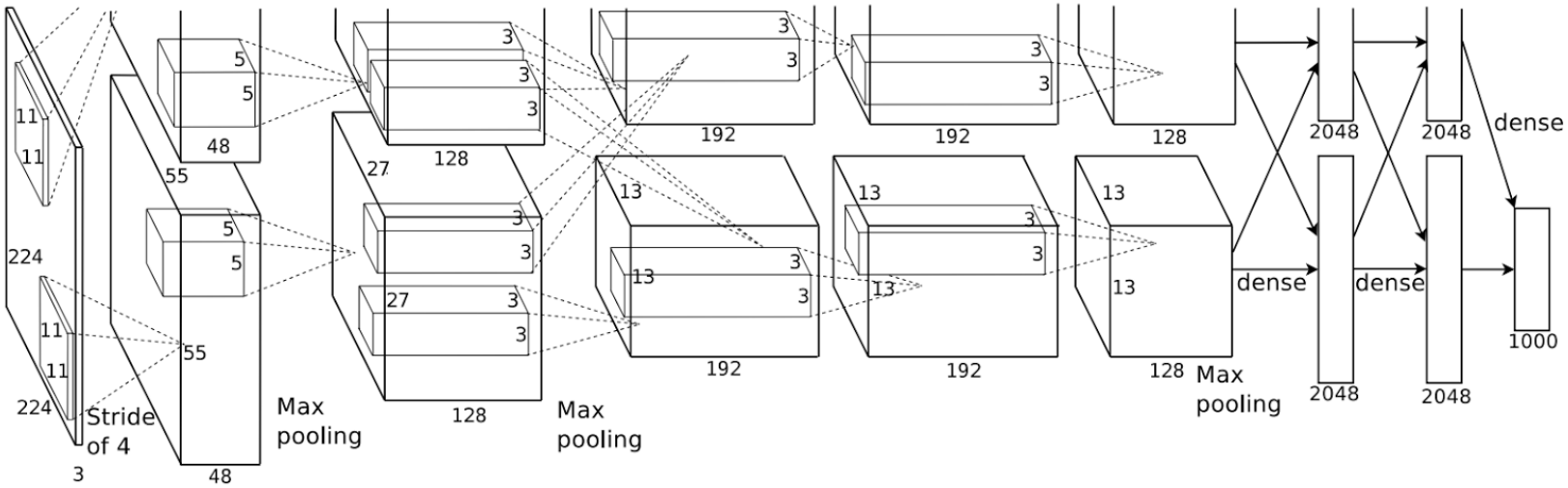
Object detection: HoG



Object detection: DPM



Object detection: Deep learning



Object detection

Object detection

image as input

Object detection

image as input

propose fixed-size regions that are likely to have objects

Object detection

image as input

propose fixed-size regions that are likely to have objects

feed each through a network with *many* layers

Object detection

image as input

propose fixed-size regions that are likely to have objects

feed each through a network with *many* layers

linear classifier at the top

Object detection

image as input

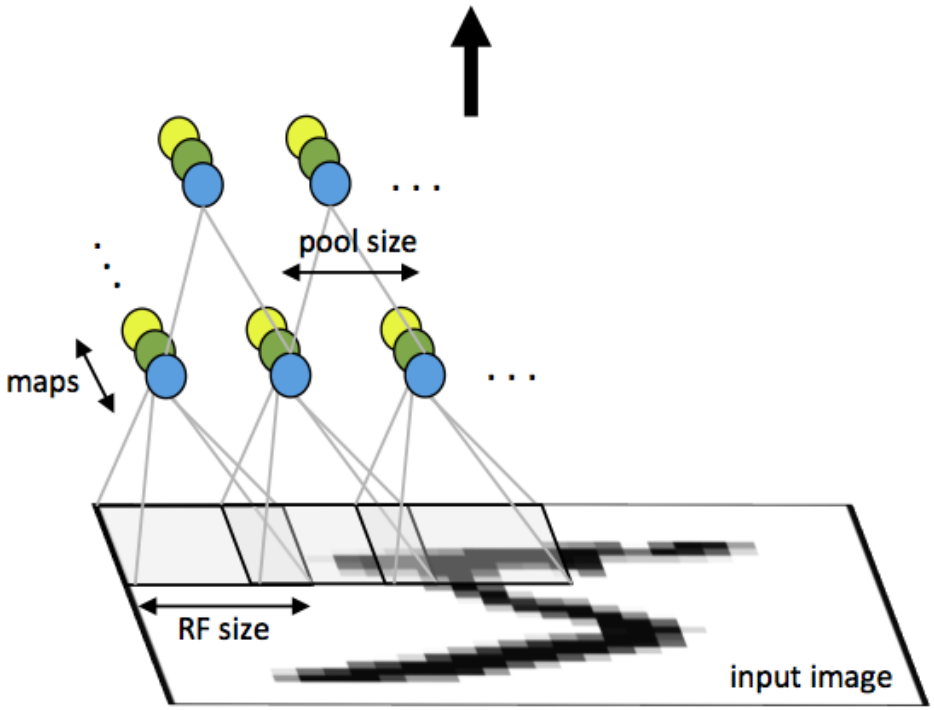
propose fixed-size regions that are likely to have objects

feed each through a network with *many* layers

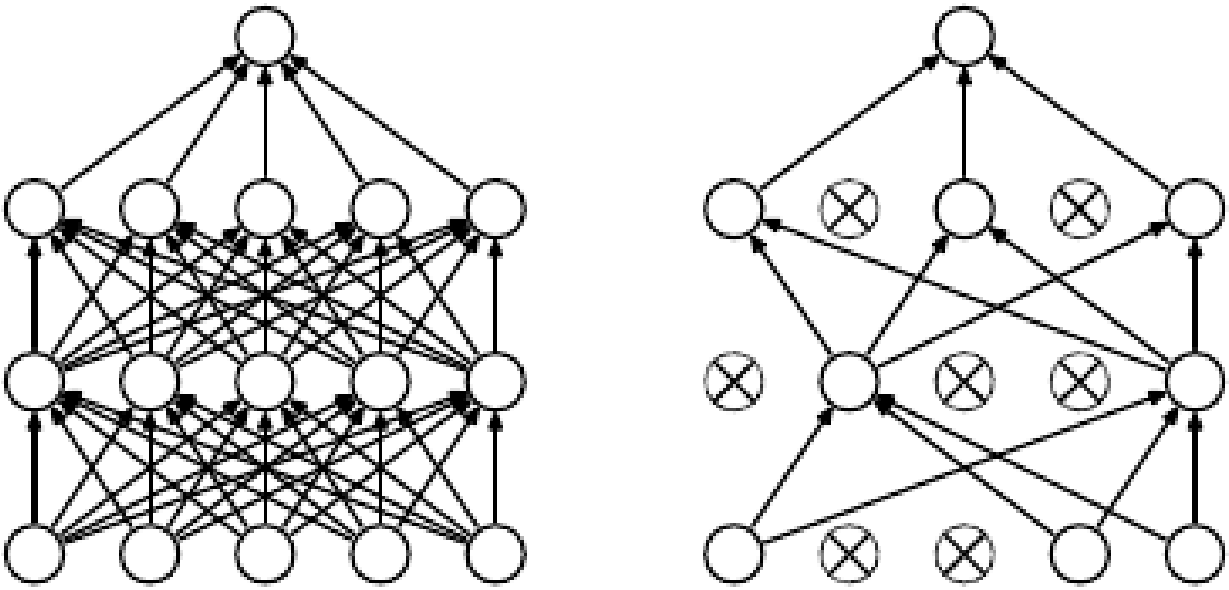
linear classifier at the top

but first . . . some missing pieces

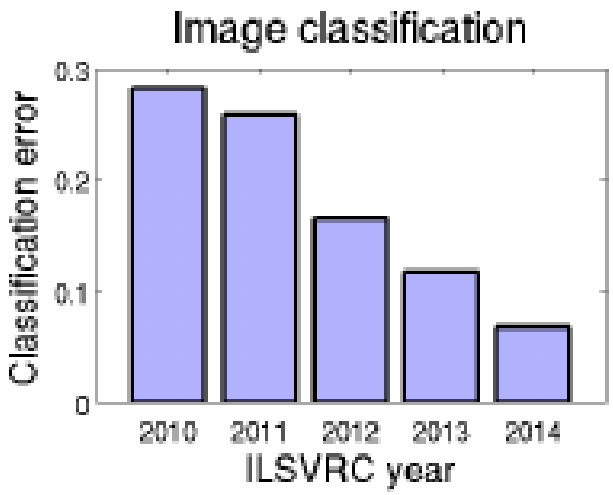
Convolution and pooling

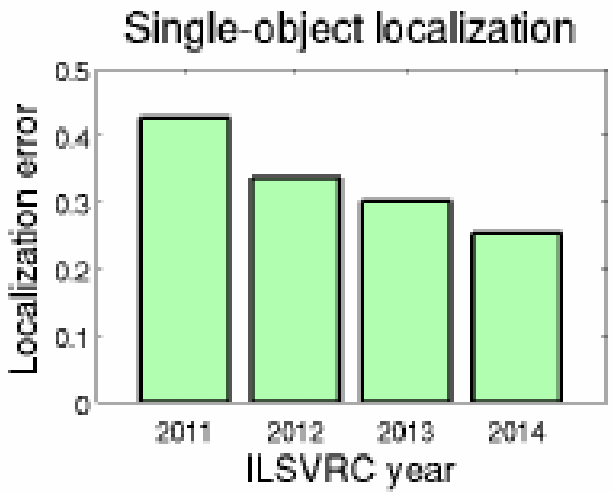


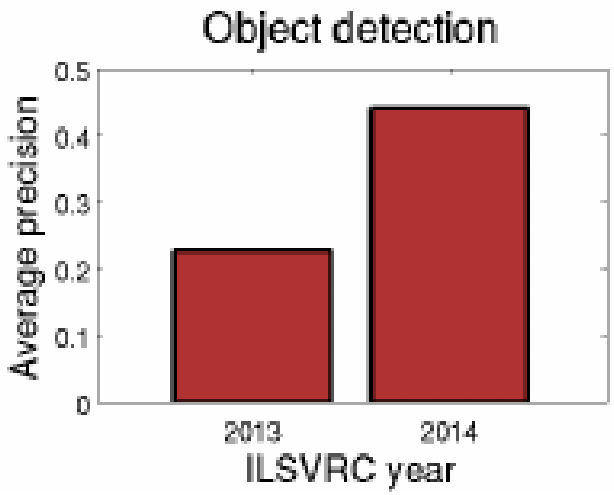
Dropout



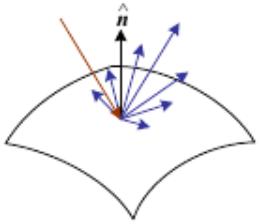
Performance







Object detection in action



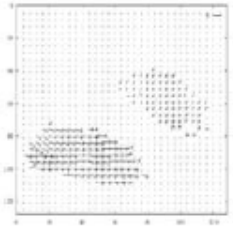
2. Image Formation

3. Image Processing

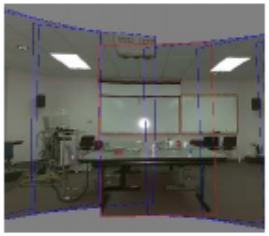
4. Features

5. Segmentation

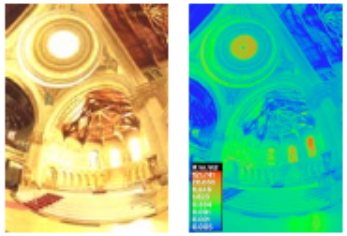
6-7. Structure from Motion



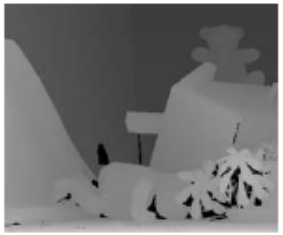
8. Motion



9. Stitching



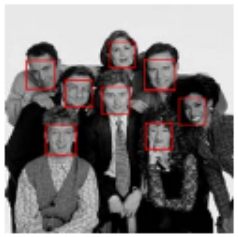
10. Computational Photography



11. Stereo

12. 3D Shape

13. Image-based Rendering



14. Recognition

Szeliski, Computer Vision; 2010

Vision . . .

Vision . . .