

# Introduction to Robotics for cognitive science

**Dr. Andrej Lúčný**

**KAI FMFI UK**

**lucny@fmph.uniba.sk**

# Web page of the subject

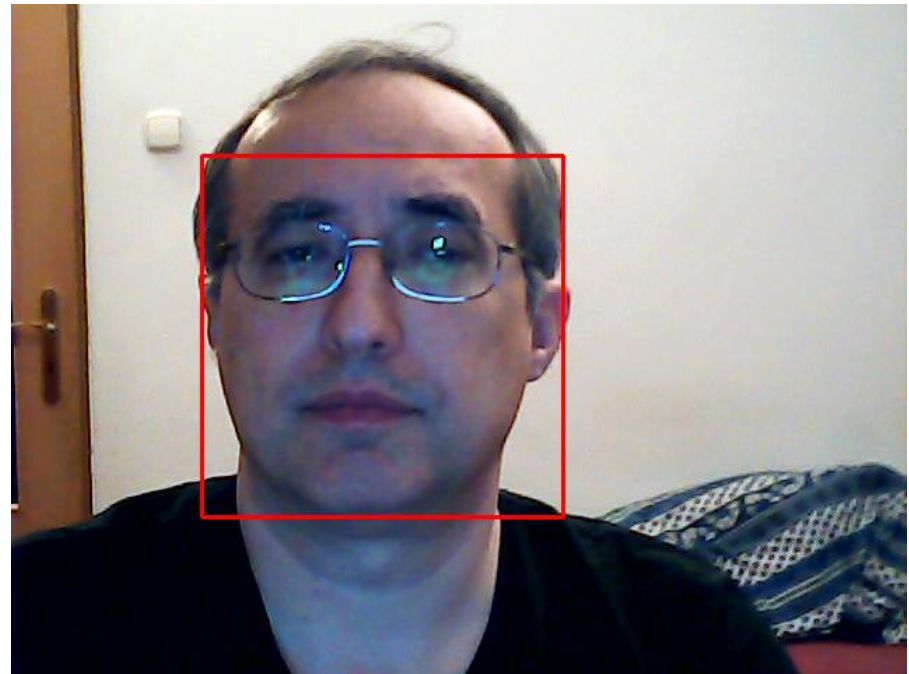
[www.agentspace.org/kv](http://www.agentspace.org/kv)



# More general object detectors

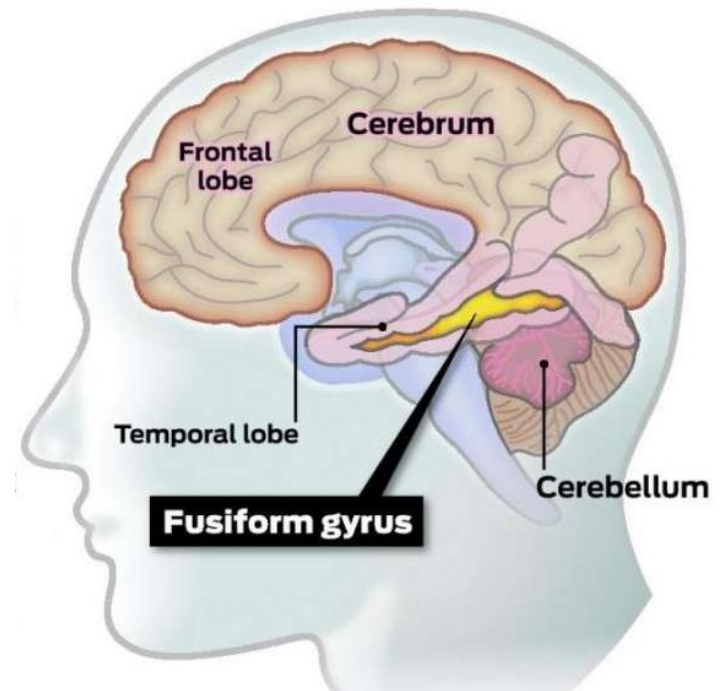
So far we have been detecting and tracking very particular objects, e.g. part of your face

Now we would like to detect a general object, e.g. face (any human face, regardless age, gender or race)



# Face detection in human brain

- Ability to recognize face is strongly stored in our genes and we do not need to learn it (unlike walking on two legs)
- In brain there is anatomic structure responsible particularly for face recognition
- Its malfunction causes no other incapability if it happens to adult, but has serious effects if it is congenital, including no empathy to other people



# Machine Learning

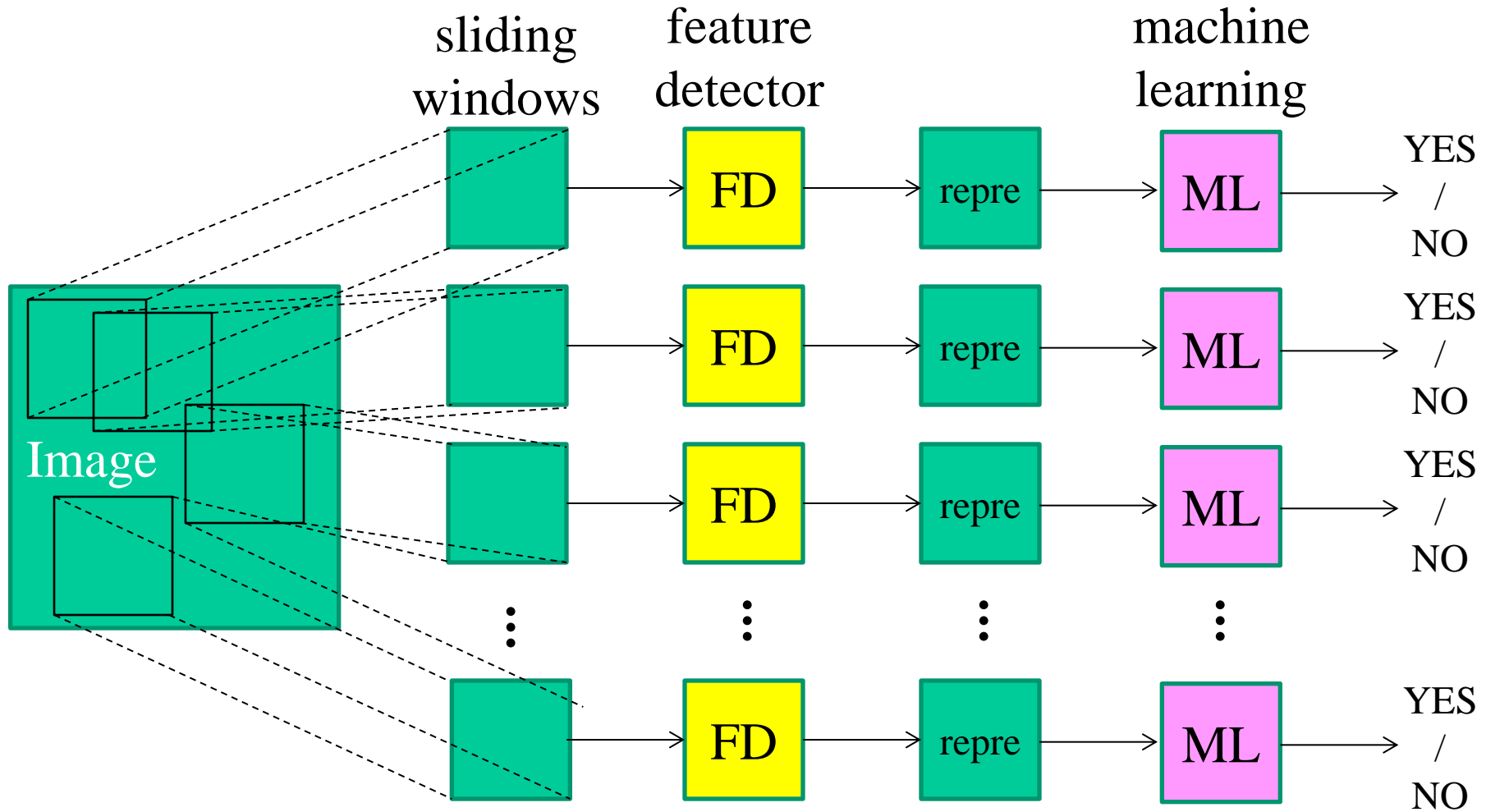
- In robotics, we have to put into robot a **model** (instead of the human genes)
- The model is created from **dataset** by a method of **supervised machine learning**
- Data in the dataset have to be **annotated**
- The typical model is a **classifier**
- Classifier tells category (face: yes or no, animal: cat, dog, elephant, monkey, other)

# Dataset



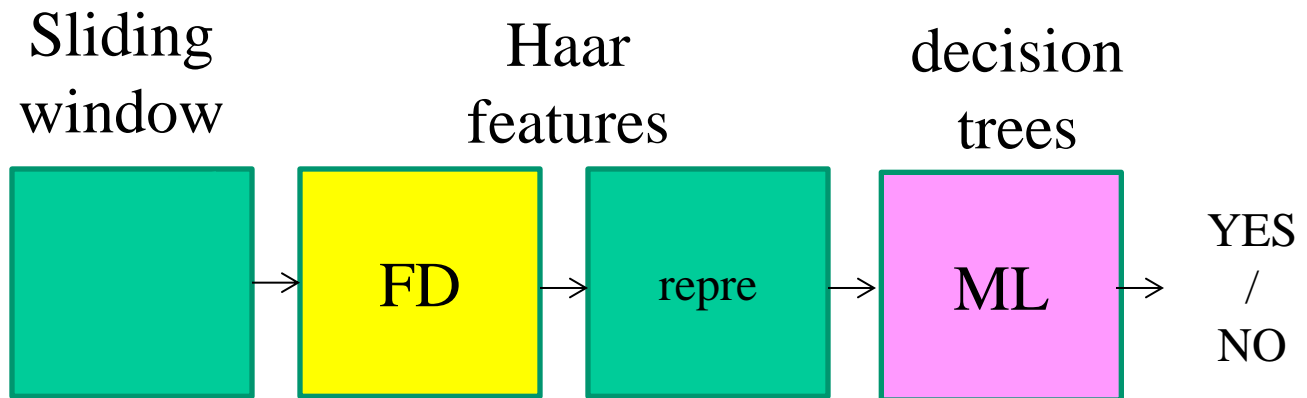
It contains positive and also negative samples

# General schema of classifier based detector



# Viola Jones Algorithm

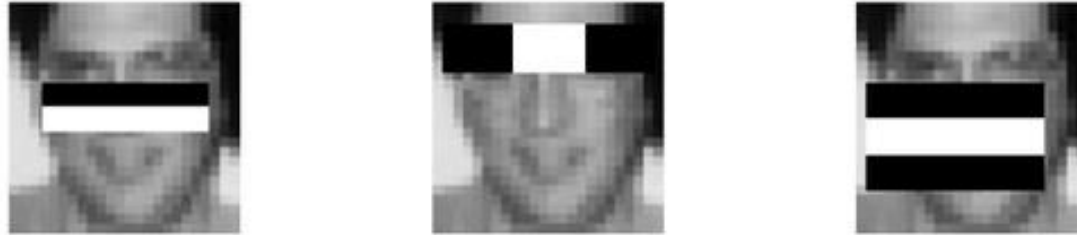
The first operational face detector [2001]



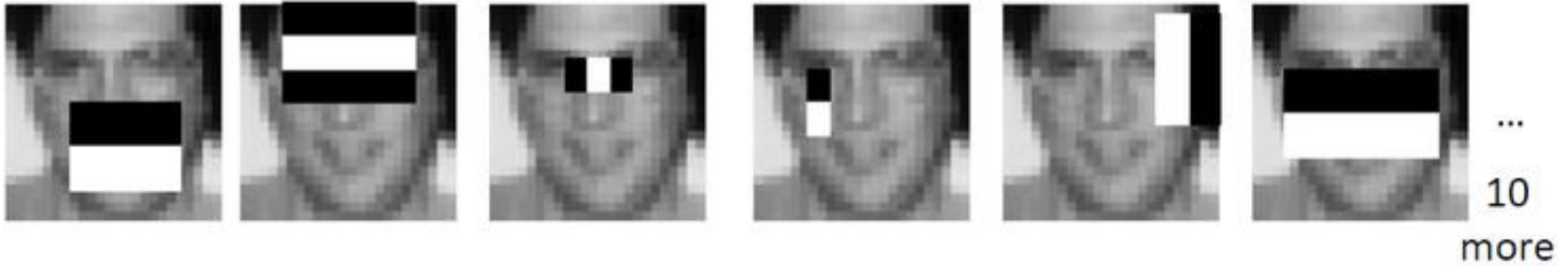


# Haar Features

Stage 0

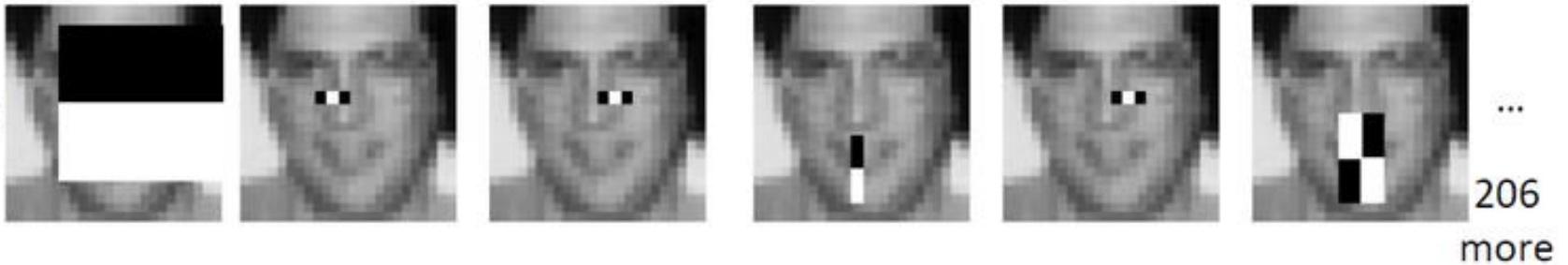


Stage 1

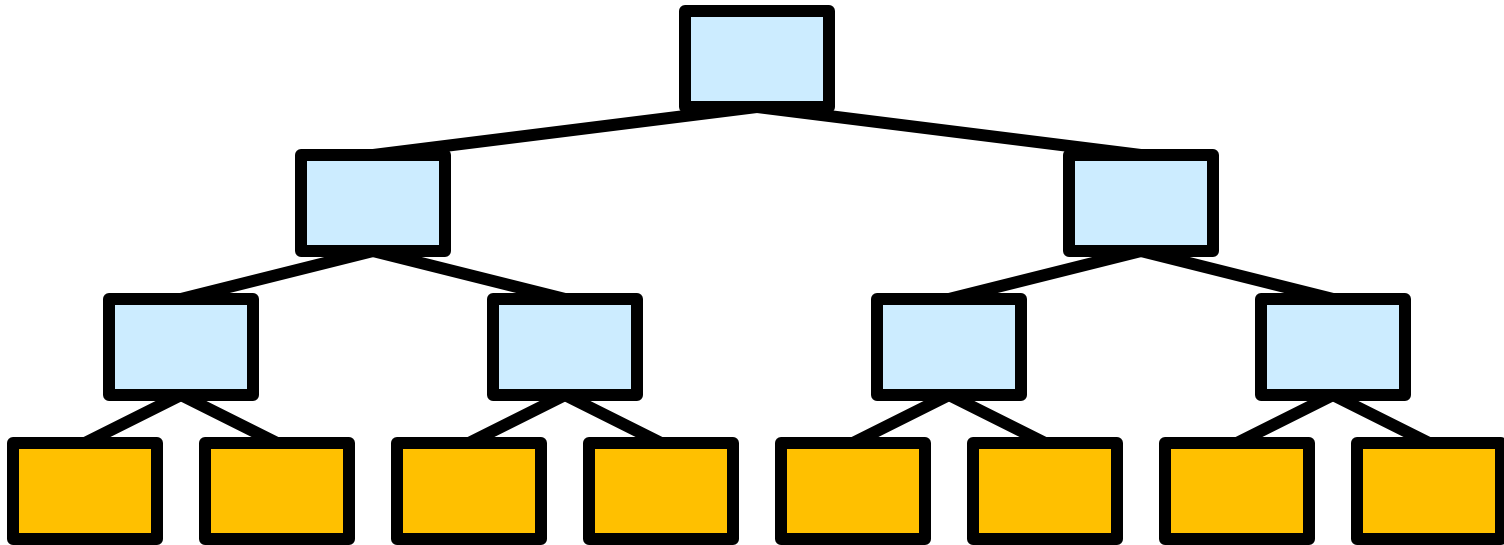


⋮

Stage 21



# Weak Classifier

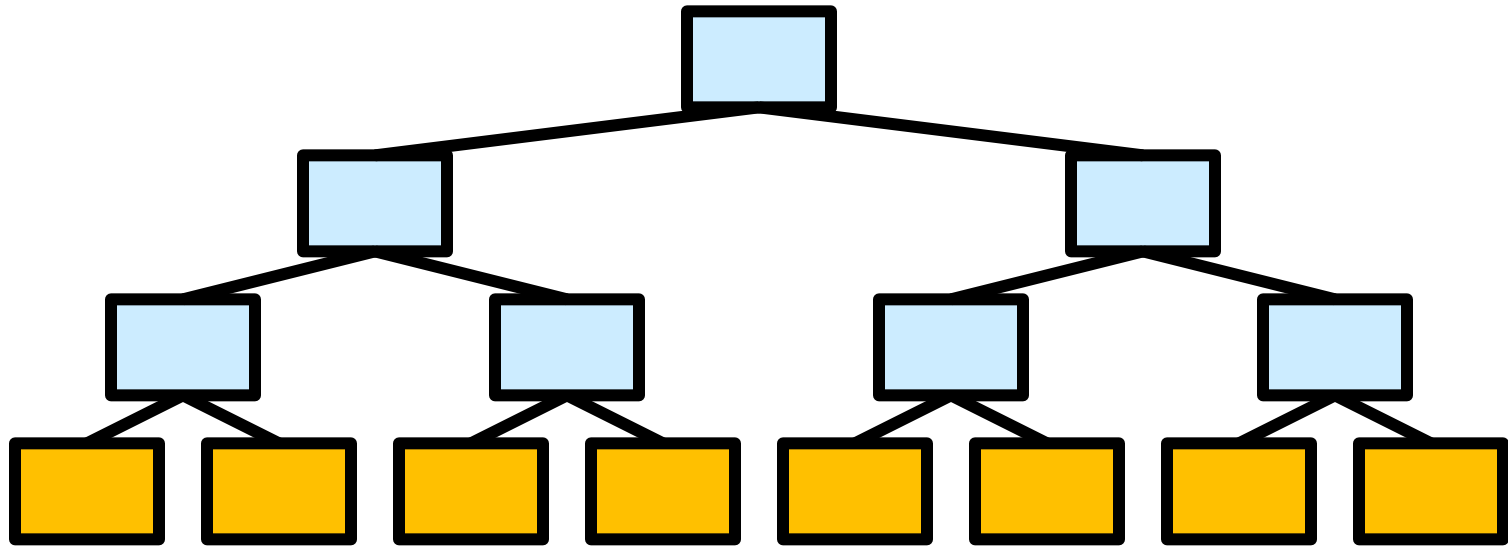


test on data



category or value

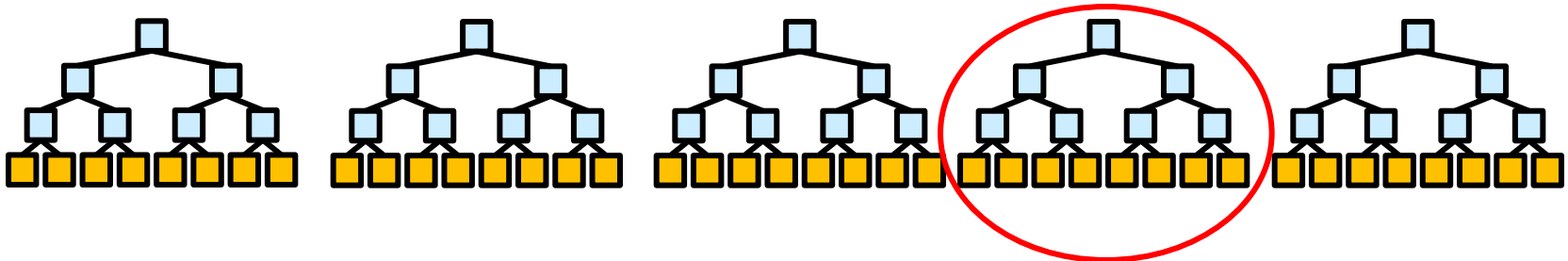
# How to get a classification tree?



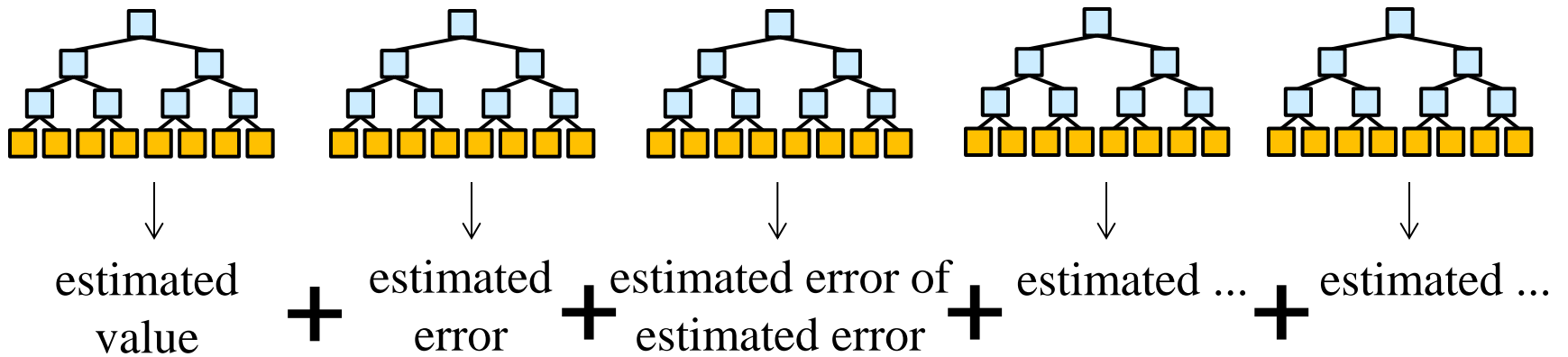
- We specify depth, e.g. 3
- In the inner nodes we test a feature (successful test – left, failed test – right)
- In the list we report: prevailing category from samples in dataset which falls in that list

# How to find a good tree?

1. Generating splitting tests we generate randomly more features and select such a one which splits the samples to sets with significant size and lowest variation in each set
2. We generate more trees and select the lowest error of classification

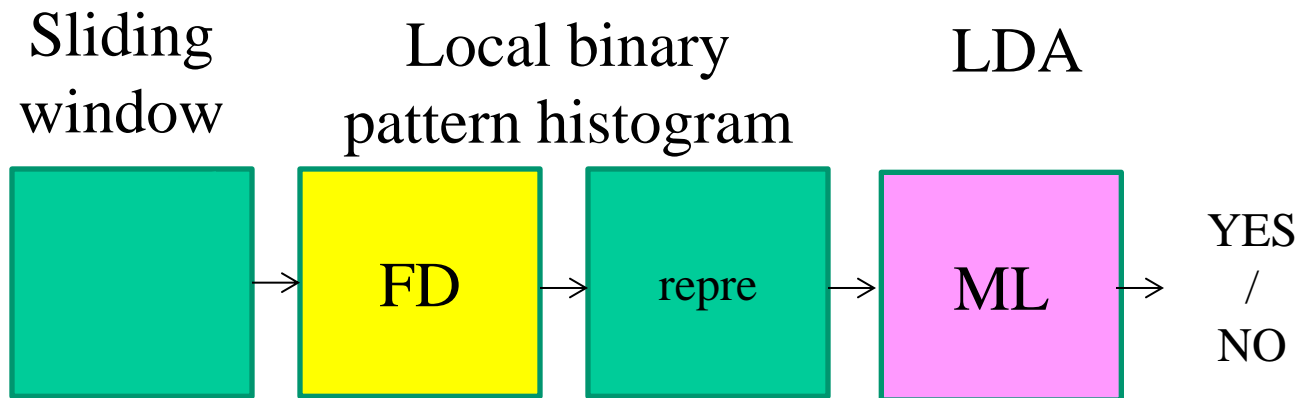


# Cascade classifier



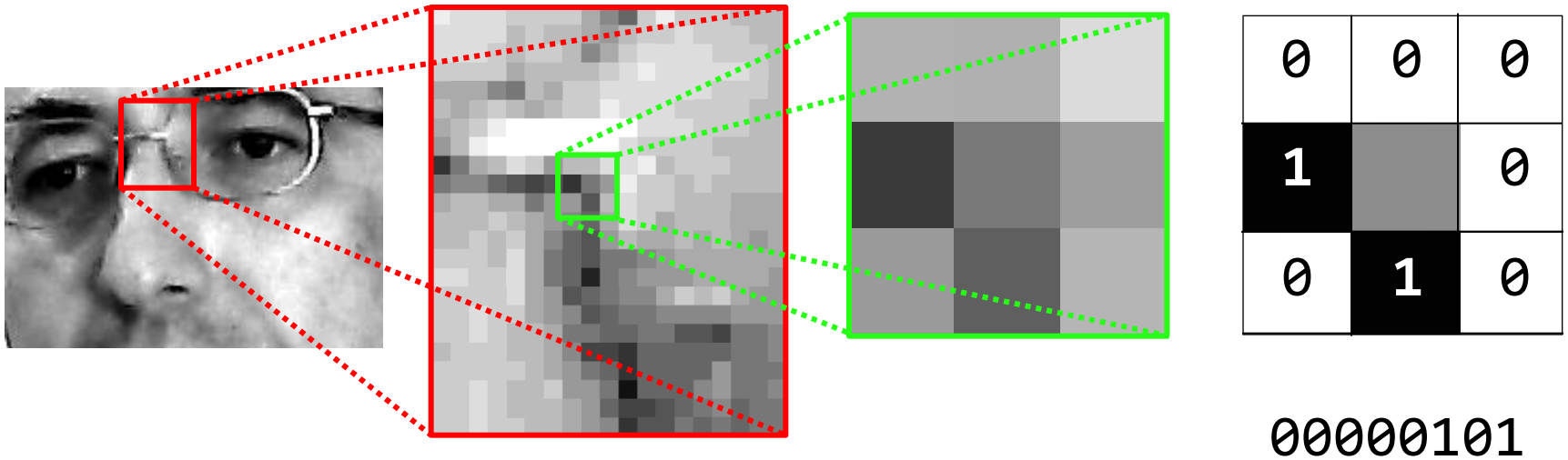
We can join more weak classifier by gradient boosting method

# LBPH approach

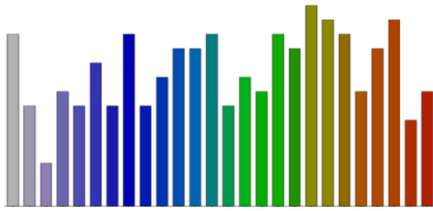


[T. Ojala, M. Pietikäinen, and D. Harwood 1994]

# LBPH Features



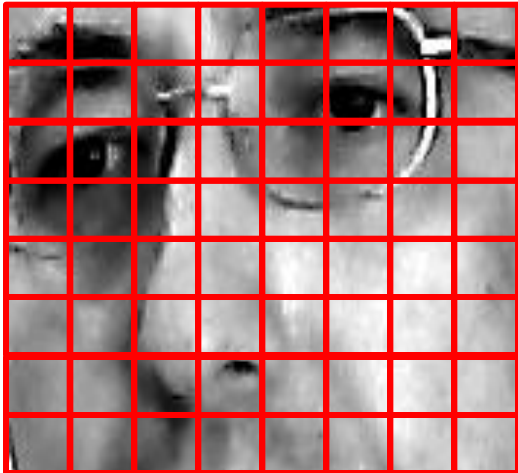
- Each pixel has its LBP  
– 0 .. 255



# LBPH Features

3F 2E 01 ... FF  
256 bytes

- Each region can be associated with histogram of LBP

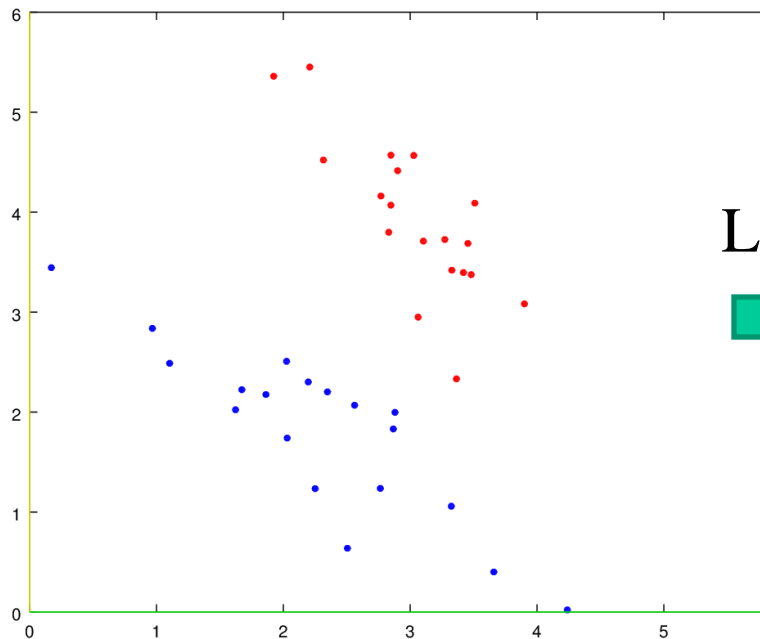


- Object is represented by set of LPBH

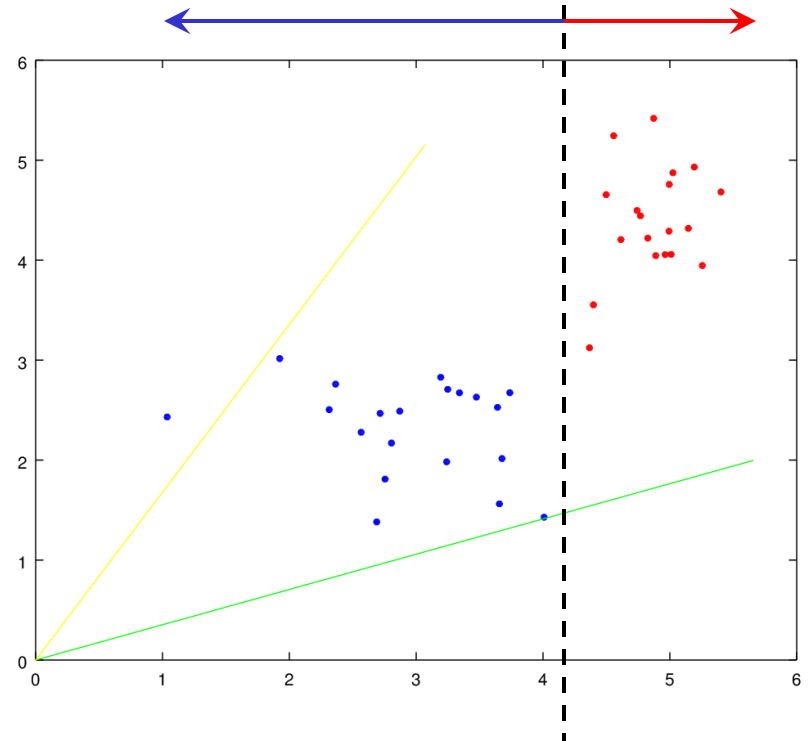
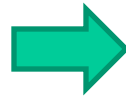


# Linear Discriminant Analysis

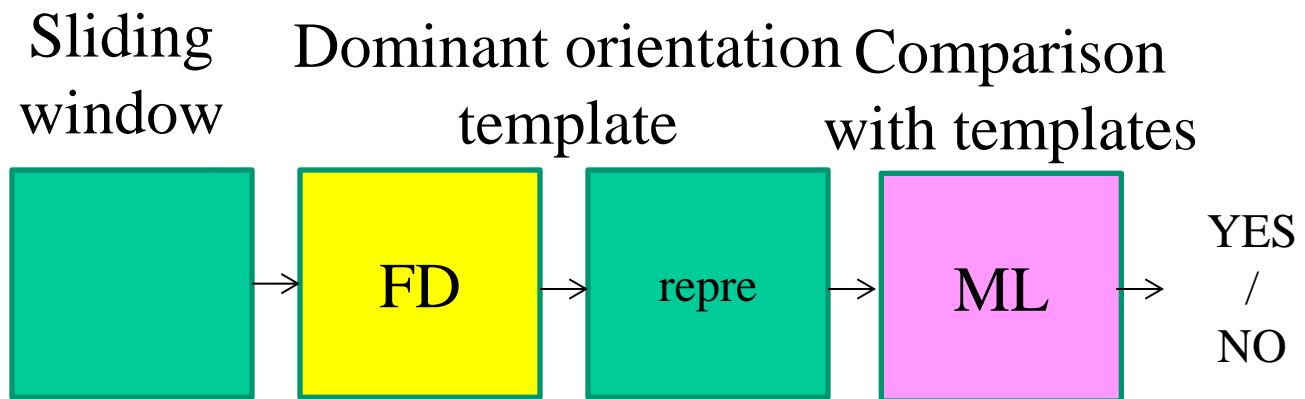
- Data are represented as point in multidimensional space (fixed number of dimensions)
- The space is reduced and transformed to easy distinguish e.g. data categories



LDA



# Dominant Orientation Templates



[Hinterstoisser, 2010]

# DOT Features



obraz



dx



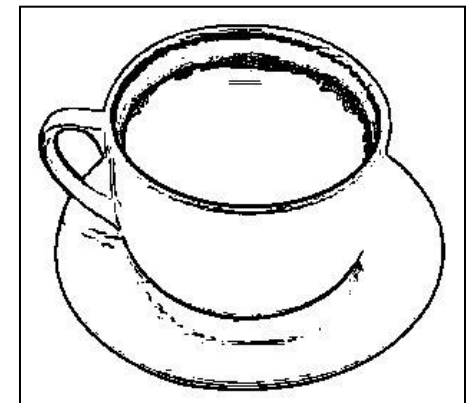
dy



|gradient|

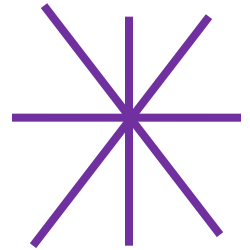
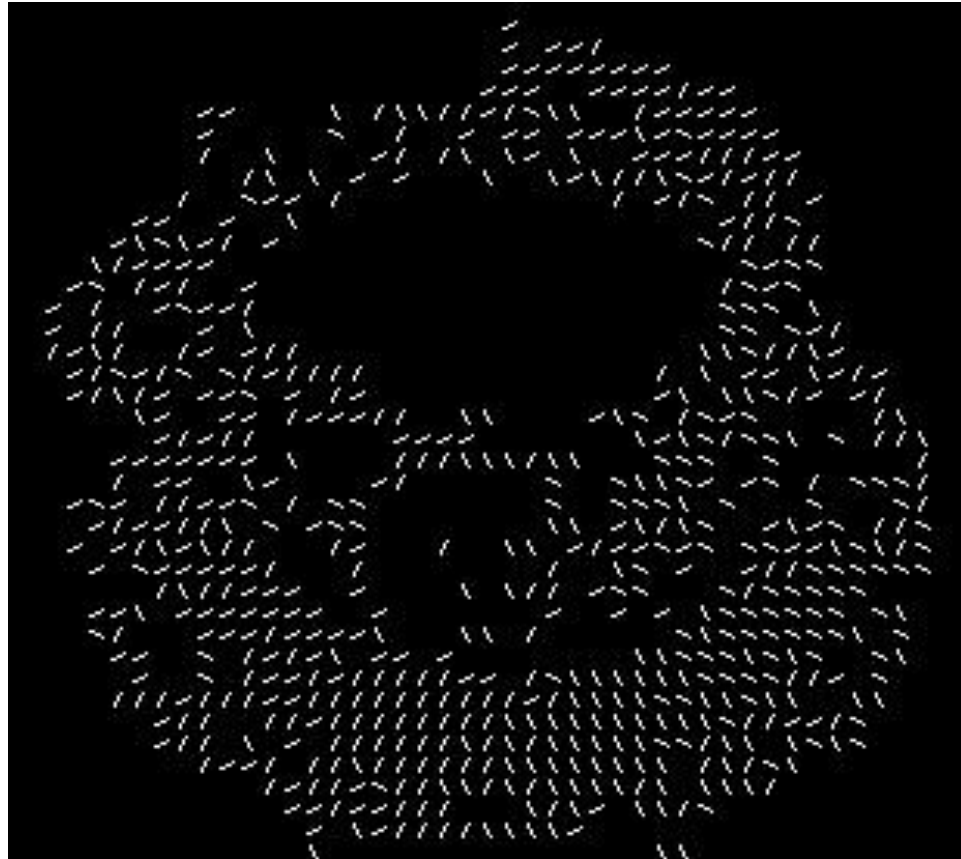
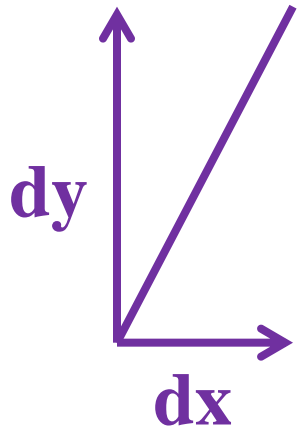


slope



edges

# Edge orientations (=slopes)



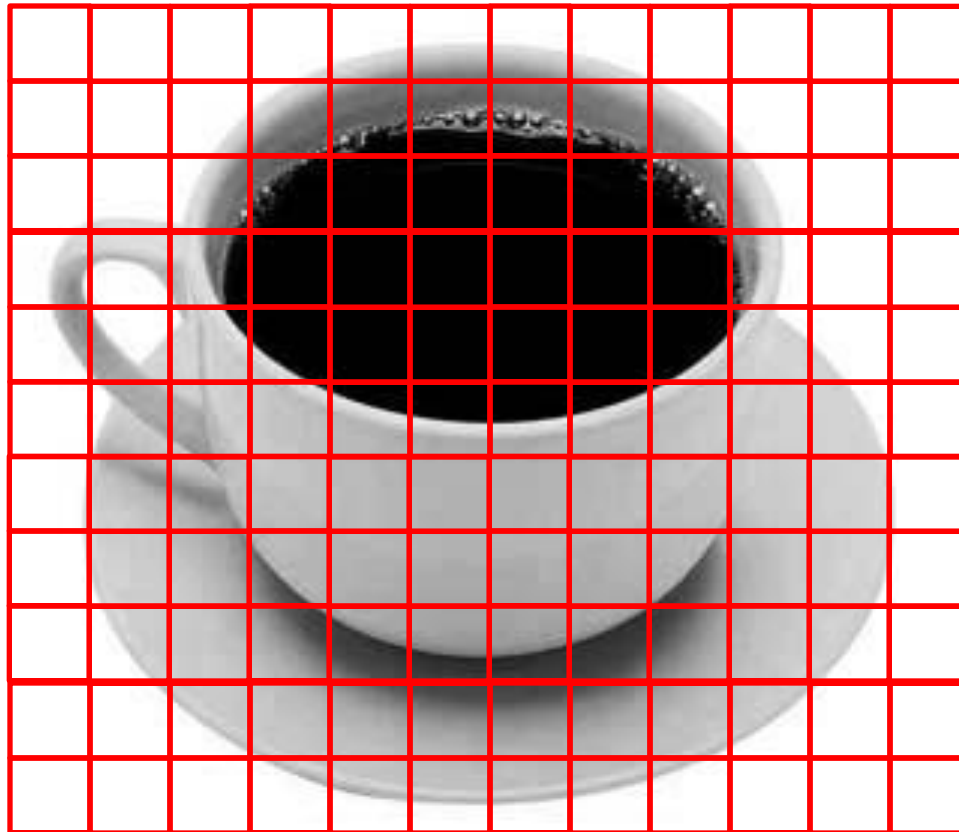
# DOT Templates

- Just clustered dominant orientations will represent the object



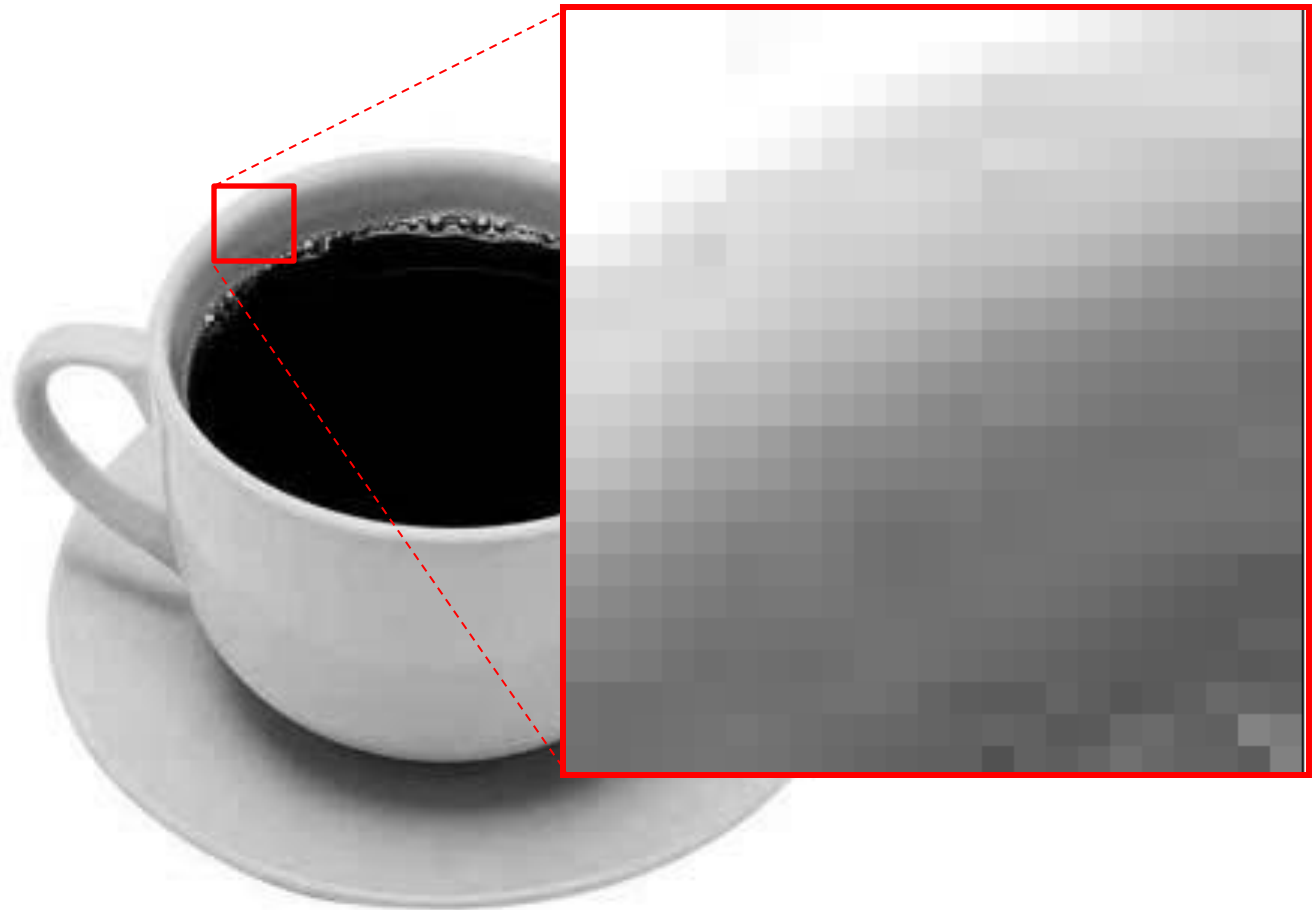
# DOT Templates

- We cover object with set of non-overlapping regions



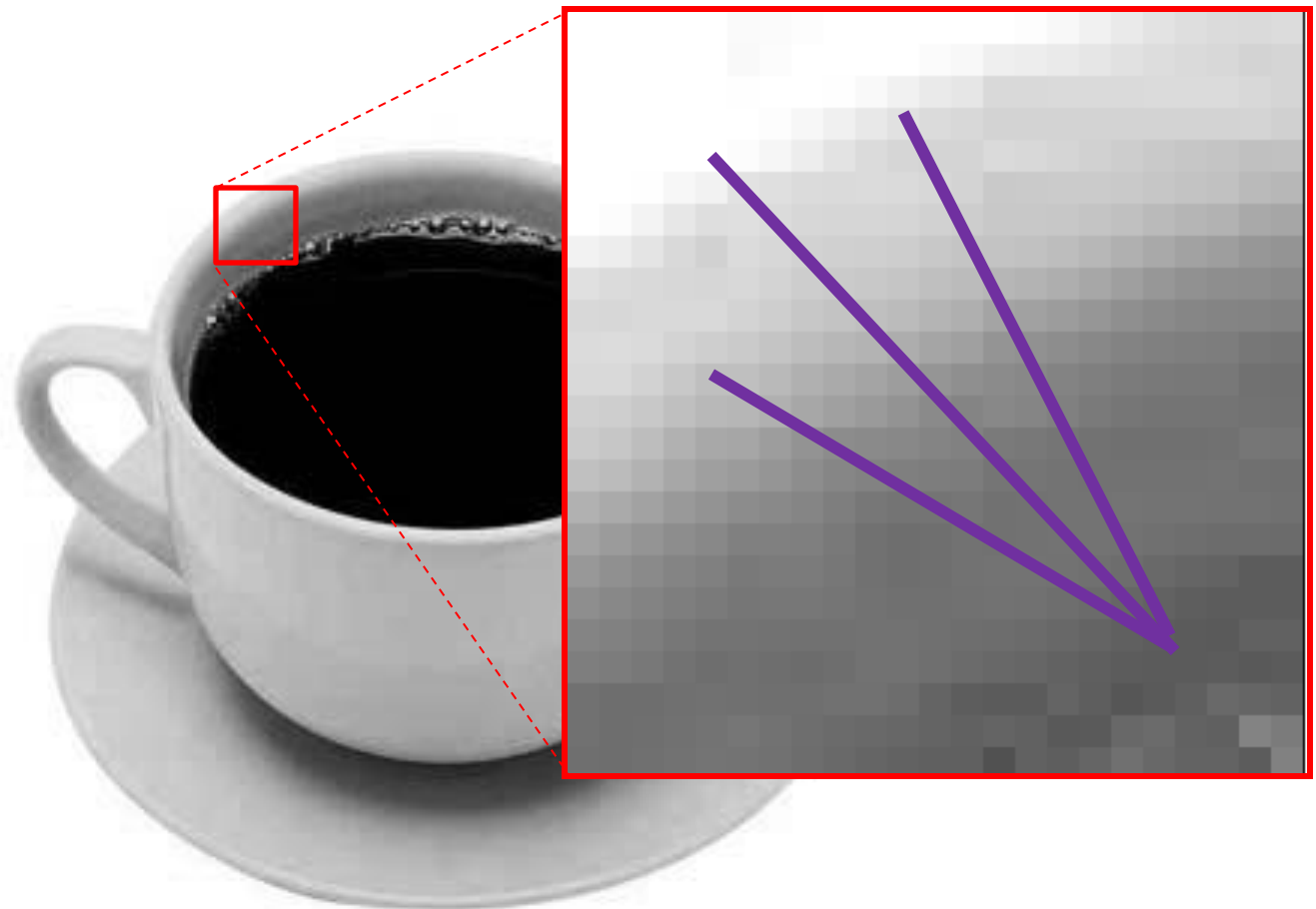
# DOT Templates

- For each pixel we calculate edge slope



# DOT Templates

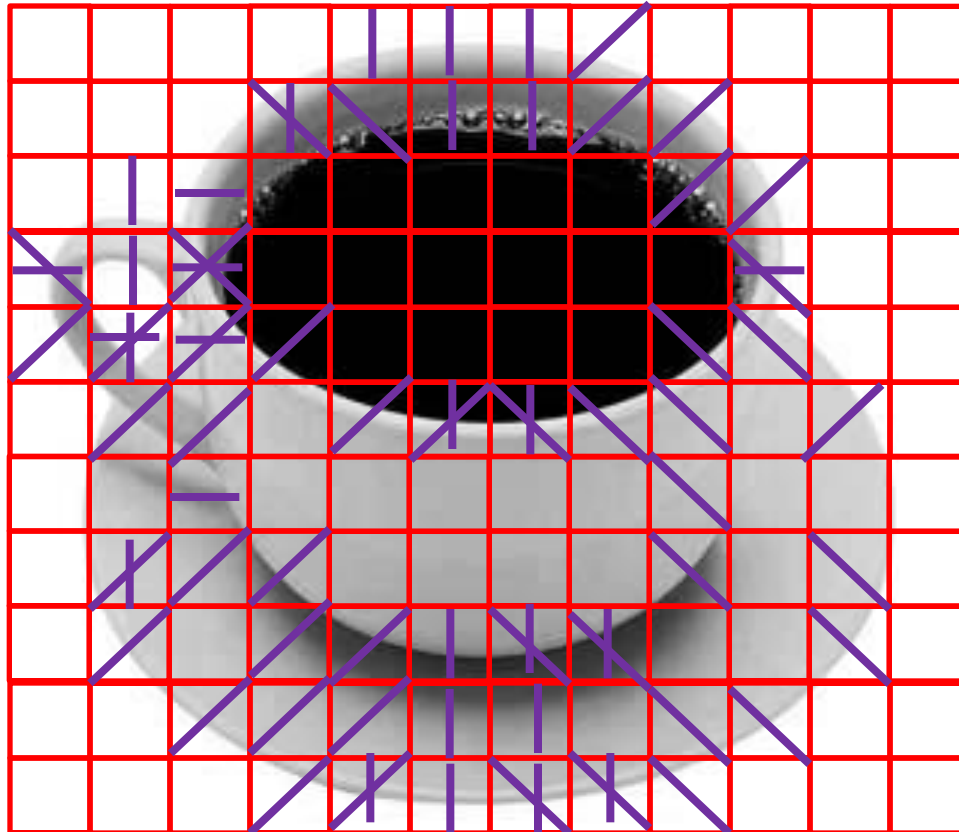
- We select prevailing slopes





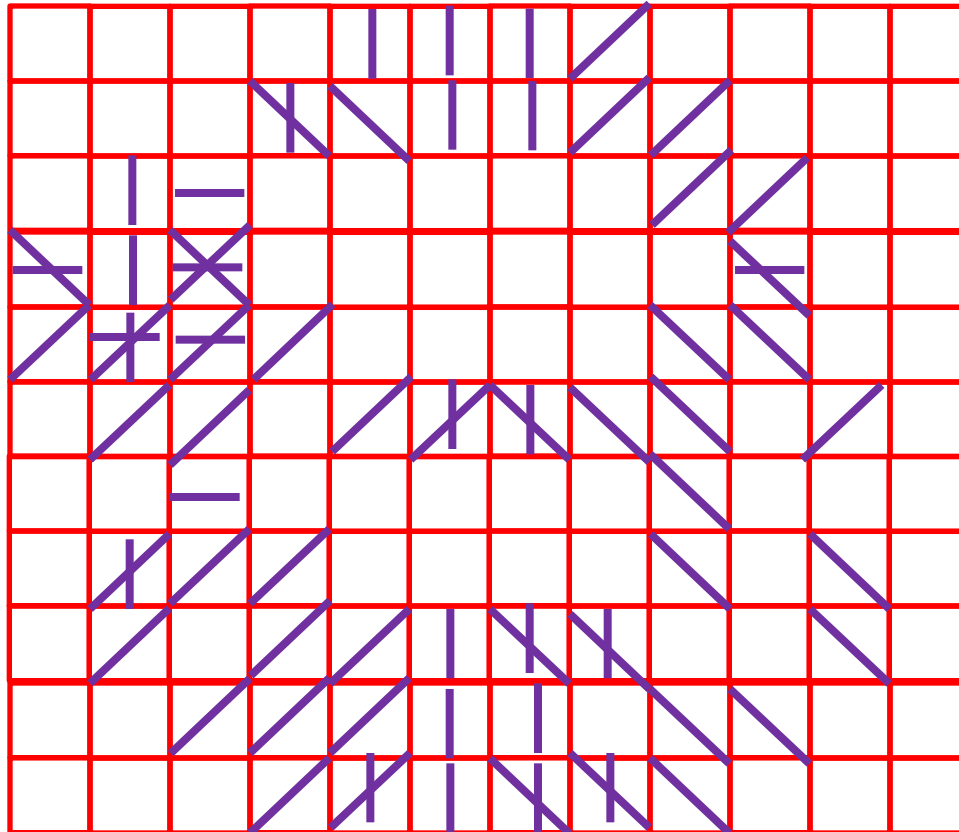
# DOT Templates

- And this our template ...



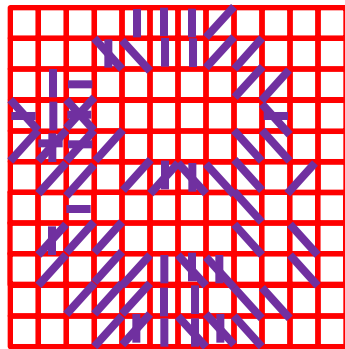
# DOT Templates

- ... which represent the object

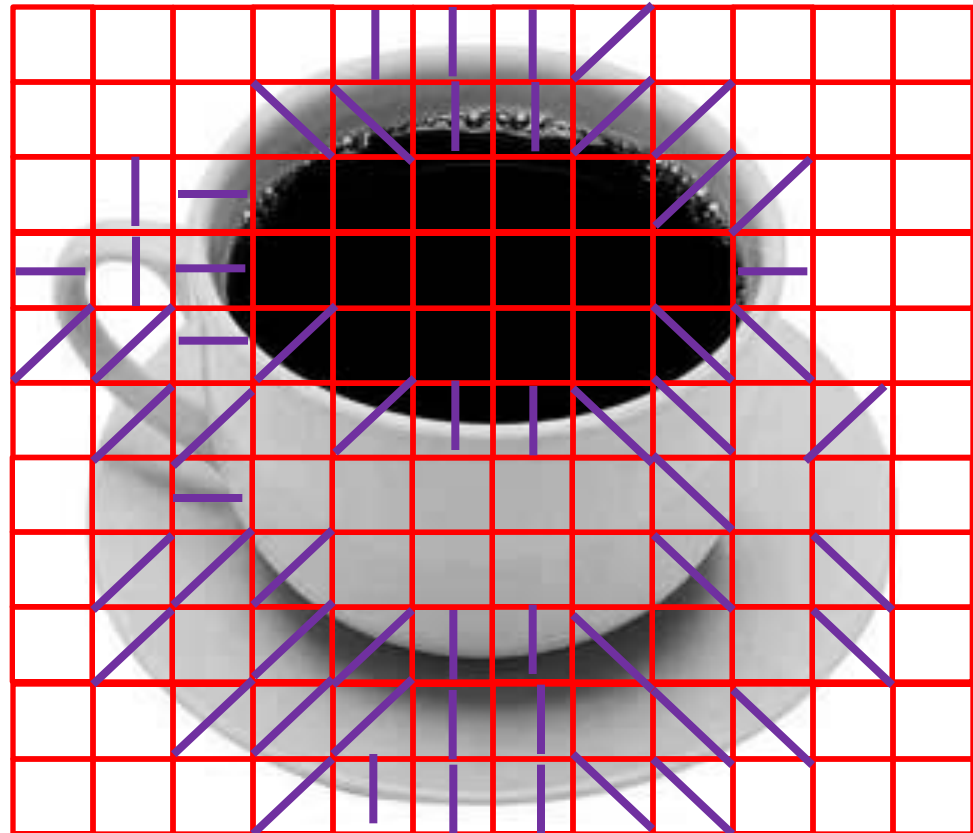


# Searching for object

- For each place we calculate one dominant orientation for each region



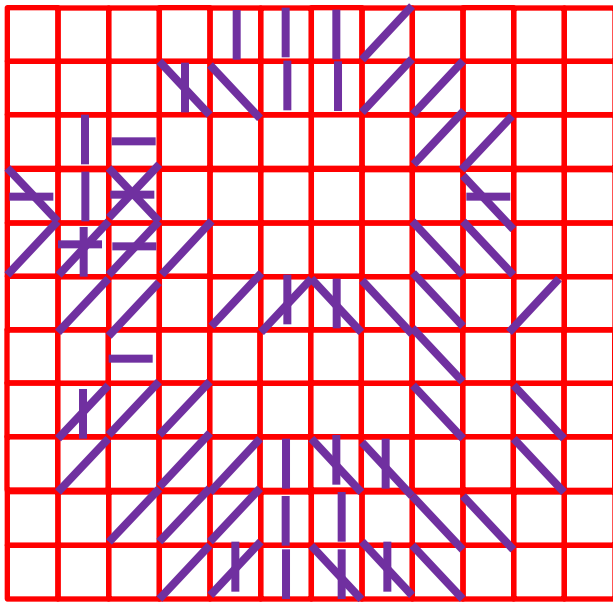
template



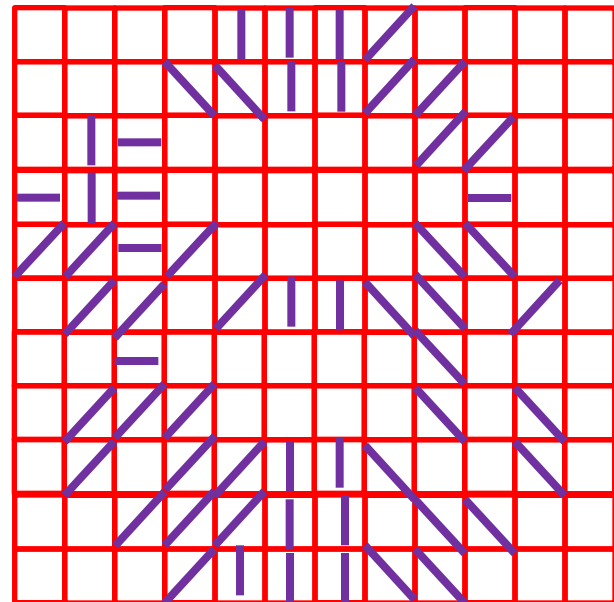
image

# Comparison with templates

- If the image matches template, we have found the object



template

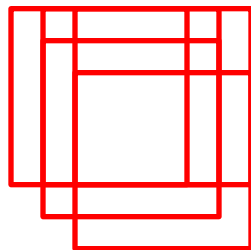
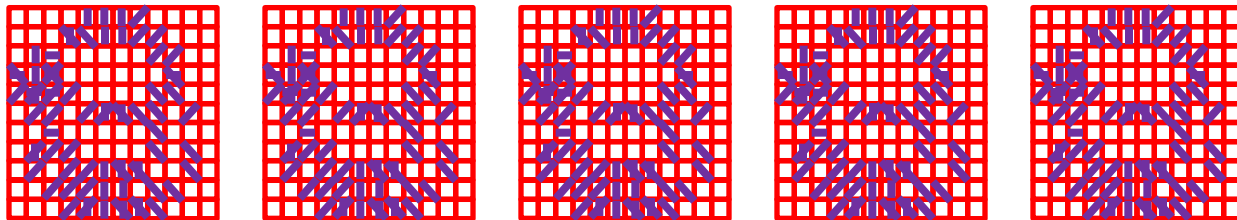


image

# What about translation of regions?

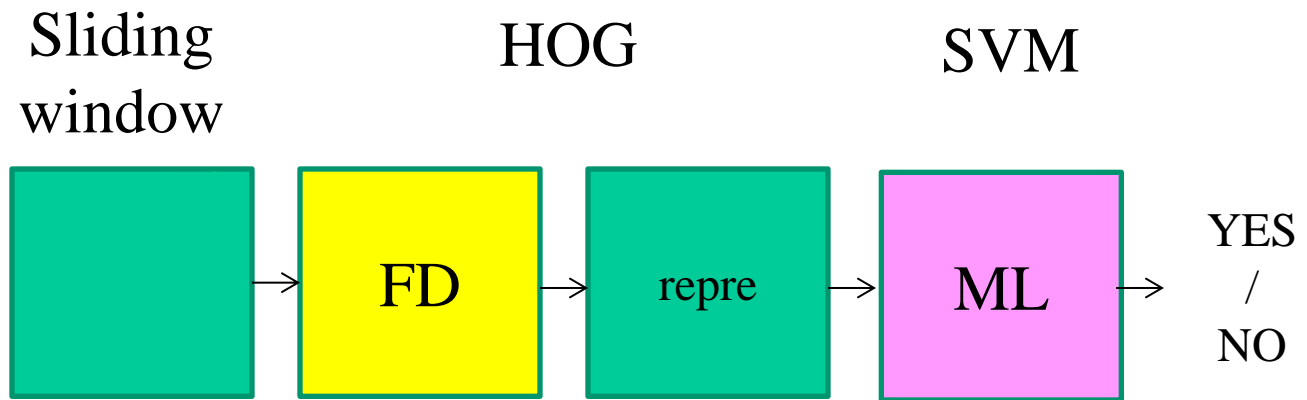
- No problem we prepare more templates for various translations, even various viewpoints, but same size

templates

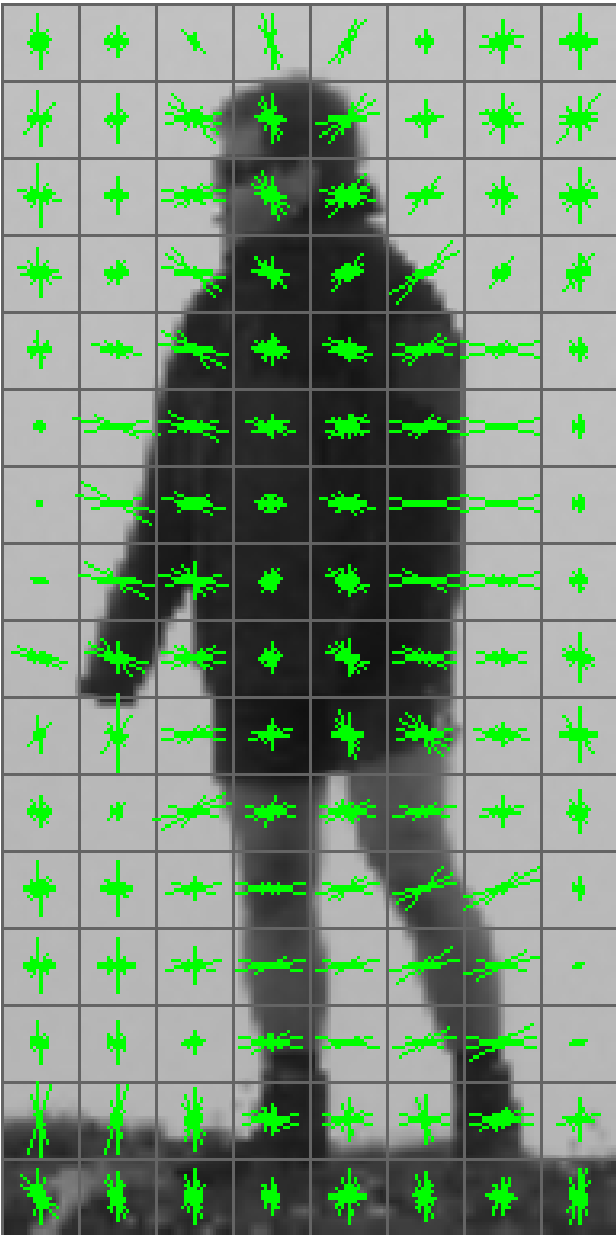


- Then we try all parts of image with variable size

# Histogram of Gradients



[Dalal, Triggs, 2005]

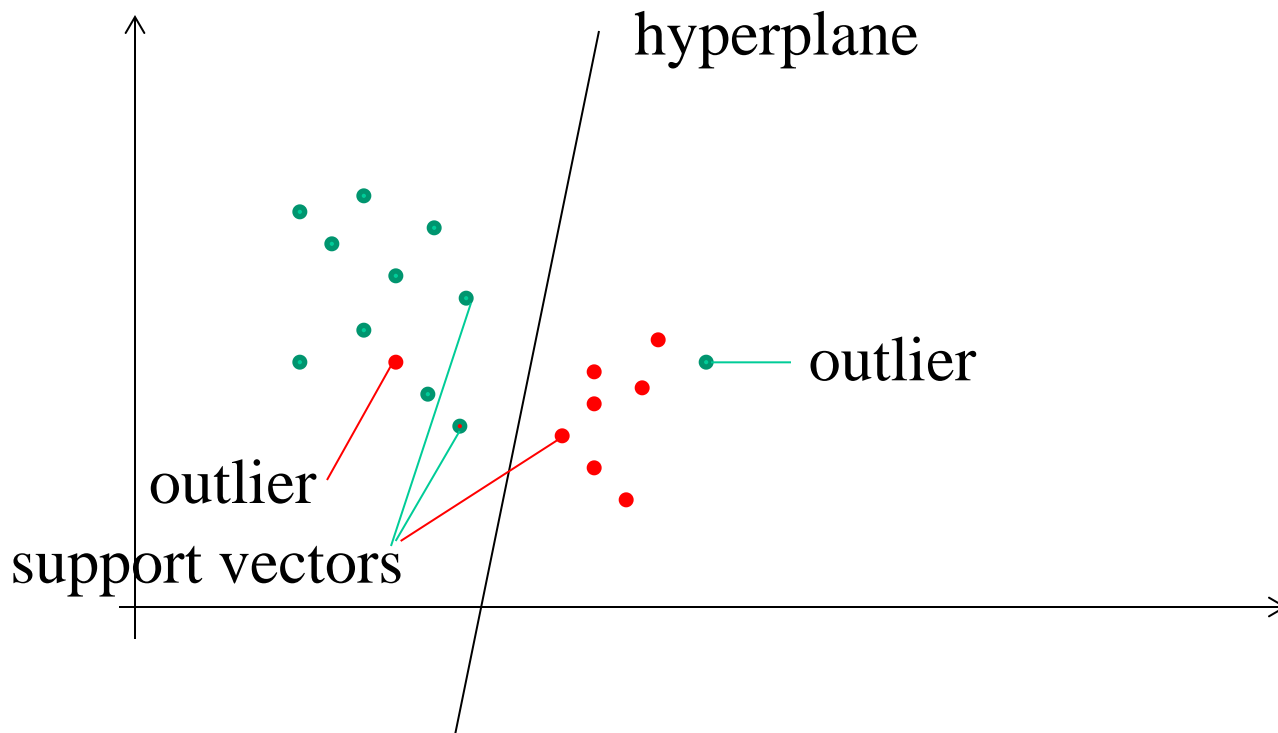


# HOG features

- Instead of few dominant slopes we take their histogram

# (Linear) Support Vector Machine

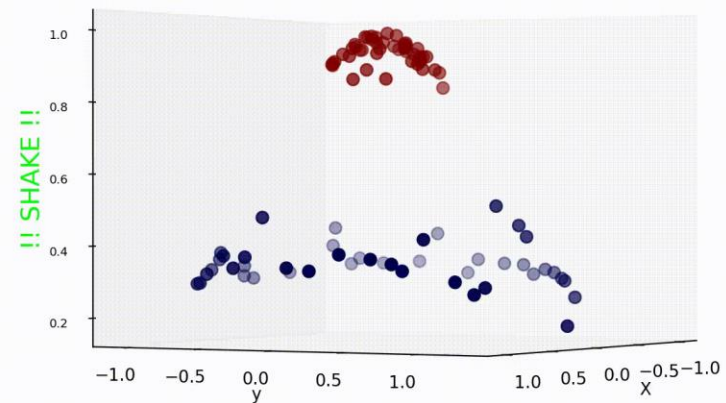
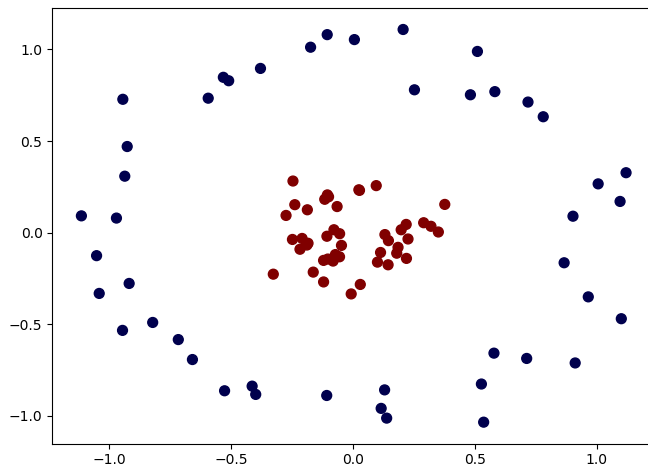
- Fast and good method which can handle outliers by maximalization of so called soft margin





# SVM Kernel trick

- SVM expand dimension of data by application of a kernel to enable separation by hyperplane



[\[more details\]](#)

- kernels are selected so that distance in the expanded space can be calculated directly in the original space

# Regression

- Analogically to classification we can have model of regression, which does not provide category but value.
- E.g. Kazemi facial landmark detector employs cascade regressor based on decision trees (and feature detector is pure selection of pixel on position relative to average face landmarks). So we put average face features on the image and regressor tells us how to move them to the right position

