Object Recognition

Lennart Waltke

University of Hamburg

17th May 2018

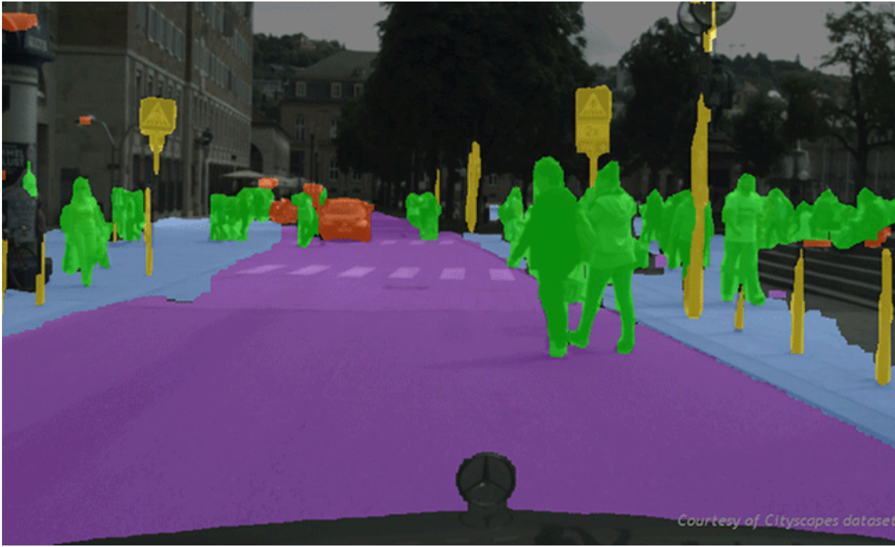
Outline

1. Introduction
2. Histogram Comparison
3. Feature Based Recognition
4. Segmentation
5. CNN

Motivation

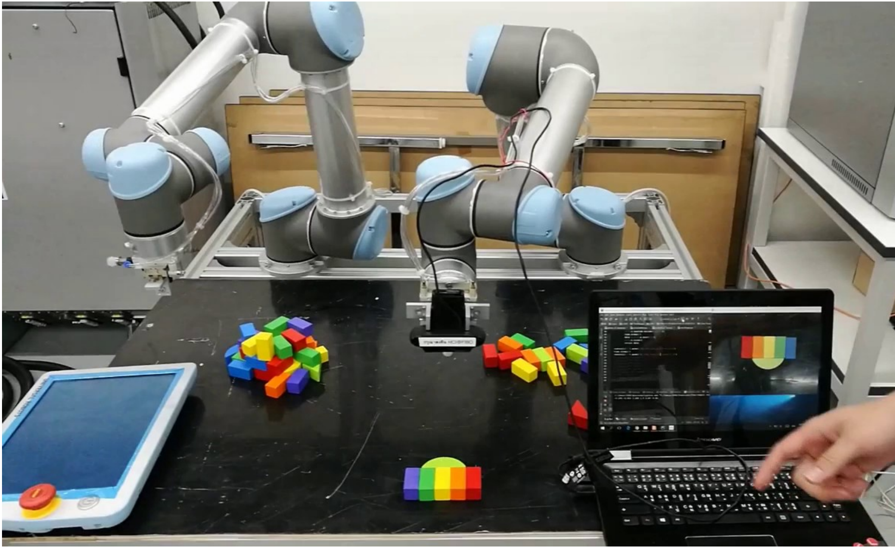
Why use Object Recognition?

Motivation



Motivation

Motivation



Histogram

- graphical representation of distribution of values
- intervals of histogram called bins
- Examples: gray value histogram, color histogram

Gray Value Histogram

- intensity values for each bin

Color Histogram

- histogram for each color channel of the picture

Histogram Comparison

How similar are histograms?

L_x-Distance

- All cells are weighted equally
- Not very robust to outliers
- Popular: Manhattan-Distance L_1 and Euclidian-Distance L_2

Appearance Based Recognition

- Objects can be represented as set of images
- For recognition it is sufficient to compute 2D-appearances
- No 3D-model needed

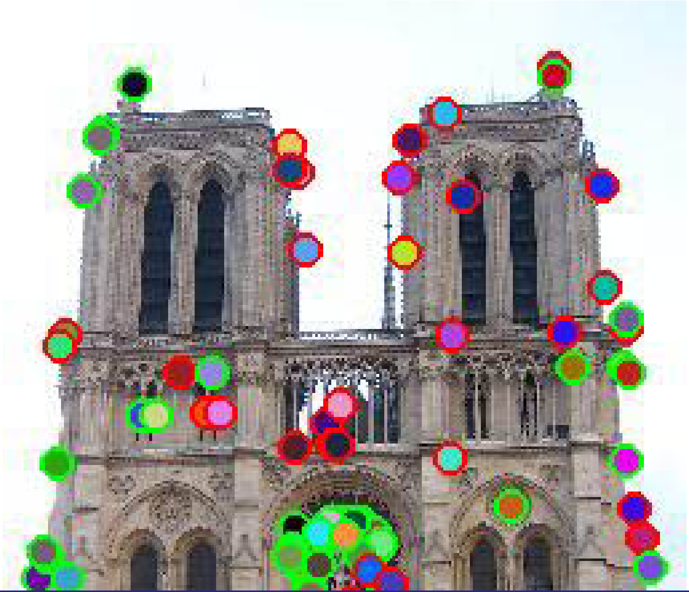
Idea

Represent each object (view) by global descriptor and match descriptors for recognition.

Features

- Properties of image/image region
- Divided into
    - global features
    - local features
- Feature information stored in descriptors
- Feature detection types
    - edge detection
    - corner detection
    - blob detection

Features



SIFT

- Scale Invariant Feature Transform
- Combines detector and descriptor
- Very robust to viewpoint and illumination changes
- Can run in real time

SIFT - Phases

1. Keypoint detection
2. Remove unstable keypoints
3. Assign orientation
4. Determine descriptor

SIFT - Keypoint Detection

- Use a image pyramid with Difference-of-Gaussian filters
- Several images per octave, each smoothed with different kernel size
- Find stable keypoints
  - compare each pixel with it's 26 neighbors
  - minima and maxima are keypoints

SIFT - Remove Unstable Keypoints

1. Remove keypoints with low contrast
2. Remove keypoints on edges

SIFT - Assign Orientation

- Compute magnitude m and Orientation Θ for k sample points
- Weight by magnitude and Gaussian Window
- Create orientation histogram
  - strongest bin is orientation of keypoint

SIFT - Determine Descriptor

- Divide patch into 4-by-4 subpatches
- Weight gradients by gaussian centered at keypoint
- Compute histogram of gradient orientation with 8 bins for subpatches
- Descriptor: $4 * 4 * 8 = 128 Dimensions$

Segmentation

- Divides image into regions
    - easier and faster to analyze
- Problems
    - Undersegmentation
    - Oversegmentation

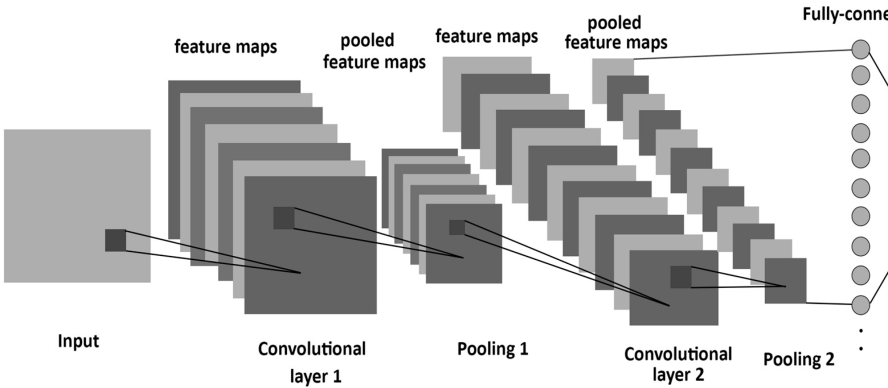
CNN

- Neural network for images as Input
- Properties:
    - operate directly on image data
    - are locally connected
    - input are not vectors but 2D-matrices or 3D-volumes

CNN - Layer Types

1. Convolutional Layer
2. Activation Function Layer
3. Pooling Layer
4. Fully Connected Layer to compute class scores

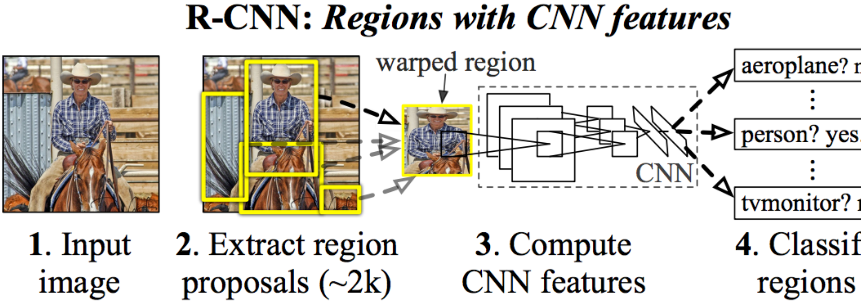
CNN


feature maps
pooled feature maps
feature maps
pooled feature maps
Fully-conne
Input
Convolutional layer 1
Pooling 1
Convolutional layer 2
Pooling 2

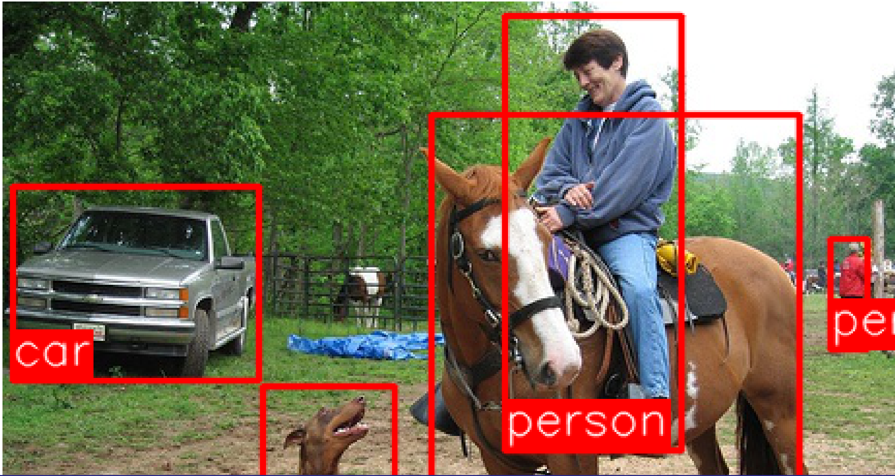
R-CNN

- Regions with CNNs
- Output: Image with bounding boxes and classifications
- Adds region proposal
- Modification towards Real-Time: Faster R-CNN
- Modification of Faster R-CNN: Mask R-CNN

R-CNN


R-CNN: *Regions with CNN features*
warped region
aeroplane? n
⋮
person? yes.
⋮
tvmonitor? n
CNN
1. Input image
2. Extract region proposals (~2k)
3. Compute CNN features
4. Classif regions

R-CNN Output



Qeustions