CS688/WST665: Web-Scale Image Retrieval
Intro to Object Recognition

Sung-Eui Yoon
(윤성의)

Course URL:
http://sglab.kaist.ac.kr/~sungeui/IR
What we will learn today?

- Introduction to object recognition
  - Representation
  - Learning
  - Recognition
What are the different visual recognition tasks?
Classification:
Does this image contain a building? [yes/no]
Classification:
Is this an beach?
Image Search

Organizing photo collections

Fei-Fei Li
Detection:

Does this image contain a car? [where?]
Detection:
Does this image contain a car? [where?]
Detection:
Which object does this image contain? [where?]
Detection:
Accurate localization (segmentation)
Detection: Estimating object semantic & geometric attributes

Object: Building, 45° pose, 8-10 meters away
It has bricks

Object: Person, back; 1-2 meters away

Object: Police car, side view, 4-5 m away
Applications of Object Recognitions and Image Retrieval
Categorization vs Single instance recognition

Does this image contain the Chicago Macy building’s?
Categorization vs Single instance recognition

Where is the crunchy nut?
Applications of Object Recognitions and Image Retrieval

• Recognizing landmarks in mobile platforms

Fei-Fei Li
Activity or Event recognition
What are these people doing?
Visual Recognition

• Design algorithms that are capable to
  – Classify images or videos
  – Detect and localize objects
  – Estimate semantic and geometrical attributes
  – Classify human activities and events

Why is this challenging?
How many object categories are there?

~10,000 to 30,000
Challenges: viewpoint variation

Michelangelo 1475-1564

Fei-Fei Li
Challenges: illumination

Fei-Fei Li
Challenges: scale
Challenges: deformation
Challenges: occlusion

Magritte, 1957

Fei-Fei Li
Challenges: background clutter

Kilmeny Niland. 1995

Fei-Fei Li
Challenges: intra-class variation
Basic issues

- Representation
  - How to represent an object category; which classification scheme?

- Learning
  - How to learn the classifier, given training data

- Recognition
  - How the classifier is to be used on novel data
Representation
- Building blocks: Sampling strategies

Interest operators

Dense, uniformly

Multiple interest operators

Randomly

Fei-Fei Li
Representation

- Building blocks: Choice of descriptors
  [SIFT, HOG, codewords....]
Representation

- Appearance only or location and appearance
Representation

- Invariances
  - View point
  - Illumination
  - Occlusion
  - Scale
  - Deformation
  - Clutter
  - etc.
Representation

- To handle intra-class variability, it is convenient to describe an object categories using probabilistic models
- Object models: Generative vs Discriminative vs hybrid
Object categorization: the statistical viewpoint

\[ p(\text{zebra} \mid \text{image}) \quad \text{vs.} \quad p(\text{no zebra} \mid \text{image}) \]

- Bayes rule:
  \[
P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}
  \]
  \[
  \frac{p(\text{zebra} \mid \text{image})}{p(\text{no zebra} \mid \text{image})}
  \]
Object categorization: the statistical viewpoint

$p(\text{zebra} \mid \text{image})$

vs.

$p(\text{no zebra} \mid \text{image})$

- Bayes rule:

\[
\frac{p(\text{zebra} \mid \text{image})}{p(\text{no zebra} \mid \text{image})} = \frac{p(\text{image} \mid \text{zebra})}{p(\text{image} \mid \text{no zebra})} \cdot \frac{p(\text{zebra})}{p(\text{no zebra})}
\]

posterior ratio

likelihood ratio

prior ratio
Object categorization: the statistical viewpoint

- Discriminative methods model posterior
- Generative methods model likelihood and prior

Bayes rule:

\[
\frac{p(\text{zebra} \mid \text{image})}{p(\text{no zebra} \mid \text{image})} = \frac{p(\text{image} \mid \text{zebra})}{p(\text{image} \mid \text{no zebra})} \cdot \frac{p(\text{zebra})}{p(\text{no zebra})}
\]

- Posterior ratio
- Likelihood ratio
- Prior ratio
Discriminative models

- Modeling the posterior ratio:
  \[
  \frac{p(\text{zebra} \mid \text{image})}{p(\text{no zebra} \mid \text{image})}
  \]
Discriminative models

Nearest neighbor
10^6 examples
Shakhnarovich, Viola, Darrell 2003
Berg, Berg, Malik 2005...

Neural networks
LeCun, Bottou, Bengio, Haffner 1998
Rowley, Baluja, Kanade 1998
...

Support Vector Machines
Guyon, Vapnik, Heisele, Serre, Poggio...

Latent SVM
Structural SVM
Felzenszwalb 00
Ramanan 03...

Boosting
Viola, Jones 2001,
Torralba et al. 2004,
Opelt et al. 2006, ...

Source: Vittorio Ferrari, Kristen Grauman, Antonio Torralba
Fei-Fei Li
Generative models

- Modeling the likelihood ratio:

\[
p(\text{image} \mid \text{zebra}) \frac{1}{p(\text{image} \mid \text{no zebra})}
\]
## Generative models

<table>
<thead>
<tr>
<th>$p(\text{image} \mid \text{zebra})$</th>
<th>$p(\text{image} \mid \text{no zebra})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

![Diagram](image)

Lecture 14
Generative models

• Naïve Bayes classifier
  - Csurka Bray, Dance & Fan, 2004

• Hierarchical Bayesian topic models (e.g. pLSA and LDA)
  - Natural scene categorization: Fei-Fei et al. 2005

• 2D Part based models
  - Constellation models: Weber et al 2000; Fergus et al 2000
  - Star models: ISM (Leibe et al 05)

• 3D part based models:
Basic issues

• Representation
  – How to represent an object category; which classification scheme?

• Learning
  – How to learn the classifier, given training data

• Recognition
  – How the classifier is to be used on novel data
Learning

- Learning parameters: What are you maximizing? Likelihood (Gen.) or performances on train/validation set (Disc.)
Learning

- Learning parameters: What are you maximizing? Likelihood (Gen.) or performances on train/validation set (Disc.)
- Level of supervision
  - Manual segmentation; bounding box; image labels; noisy labels
- Batch/incremental
- Priors
Learning

- Learning parameters: What are you maximizing? Likelihood (Gen.) or performances on train/validation set (Disc.)

- Level of supervision
  - Manual segmentation; bounding box; image labels; noisy labels

- Batch/incremental

- Priors

- Training images:
  - Issue of overfitting
  - Negative images for discriminative methods
Basic issues

• Representation
  – How to represent an object category; which classification scheme?

• Learning
  – How to learn the classifier, given training data

• Recognition
  – How the classifier is to be used on novel data
Recognition

- Recognition task: classification, detection, etc..
Recognition

- Recognition task

- Search strategy: Sliding Windows
  - Simple
  - Computational complexity (x, y, S, \( \theta \), N of classes)

Viola, Jones 2001,

- BSW by Lampert et al 08
- Also, Alexe, et al 10
Recognition

- Recognition task

- Search strategy: Sliding Windows
  - Simple
  - Computational complexity (x, y, S, θ, N of classes)
    - BSW by Lampert et al 08
    - Also, Alexe, et al 10
  - Localization
    - Objects are not boxes

Viola, Jones 2001,
Recognition

- Recognition task
- Search strategy: Sliding Windows
  - Simple
  - Computational complexity \((x, y, S, \theta, N\) of classes\)
    - BSW by Lampert et al 08
    - Also, Alexe, et al 10
  - Localization
    - Objects are not boxes
    - Prone to false positive

Non max suppression:
Canny ’86
......
Desai et al, 2009
Recognition

- Recognition task
- Search strategy
- Attributes

- It has metal
- it is glossy
- has wheels

Category: car
Azimuth = 225°
Zenith = 30°

- Savarese, 2007
- Sun et al 2009
- Liebelt et al., ’08, 10
- Farhadi et al 09

- Farhadi et al 09
- Lampert et al 09
- Wang & Forsyth 09

Fei-Fei Li
Recognition

- Recognition task
- Search strategy
- Attributes
- Context

Semantic:
- Torralba et al 03
- Rabinovich et al 07
- Gupta & Davis 08
- Heitz & Koller 08
- L-J Li et al 08
- Yao & Fei-Fei 10

Geometric
- Hoiem, et al 06
- Gould et al 09
- Bao, Sun, Savarese 10
Basic issues

• Representation
  – How to represent an object category; which classification scheme?

• Learning
  – How to learn the classifier, given training data

• Recognition
  – How the classifier is to be used on novel data
What have we learned today?

• Introduction to object recognition
  – Representation
  – Learning
  – Recognition
Next Time…

- Bag of visual words approach