Lecture 6:

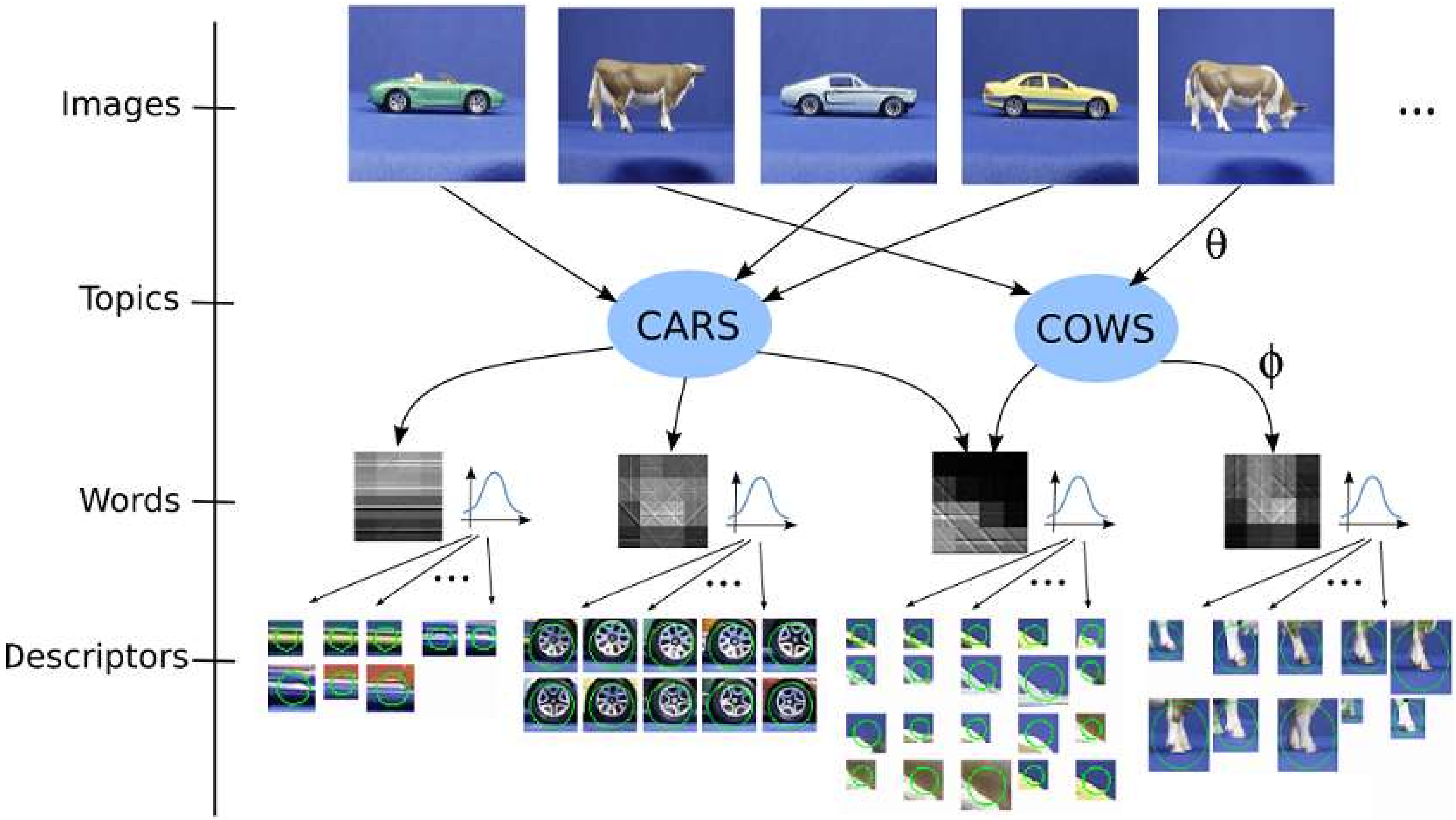
Object Categorization

Some slides were adapted/taken from various sources, including 3D Computer Vision of Prof. Hee, NUS, Air Lab Summer School, The Robotic Institute, CMU, Computer Vision of Prof. Mubarak Shah, UCF, Computer Vision of Prof. William Hoff, Colorado School of Mines and many more. We thankfully acknowledge them. Students are requested to use this material for their study only and **NOT** to distribute it.

Outlines

- Image Representation and Filtering
- Edge detection
- Object categorization
- Interest point detection
- Depth Estimation
- Optical Flow for video
- Motion model
- Mean shift tracking etc.

Object Categorization



Challenges

- Viewpoint variation.
- Scale variation.
- Deformation
- Occlusion.
- Illumination conditions.
- Background clutter.
- Intra-class variation.

Challenges: Viewpoint variation

- Viewpoint variation: A single instance of an object can be oriented in many ways with respect to the camera

Challenges: Scale variation

- Scale variation: Visual classes often exhibit variation in their size (size in the real world, not only in terms of their extent in the image)

Challenges: Deformation

- Deformation. Many objects of interest are not rigid bodies and can be deformed in extreme ways

Challenges: Occlusion

- Occlusion: The objects of interest can be occluded. Sometimes only a small portion of an object (as little as few pixels) could be visible.

Challenges: Illumination conditions

- Illumination conditions: The effects of illumination are drastic on the pixel level.

Challenges: Background clutter

- Background clutter. The objects of interest may blend into their environment, making them hard to identify

Challenges: Intra-class variation

- Intra-class variation. The classes of interest can often be relatively broad, such as chair. There are many different types of these objects, each with their own appearance.

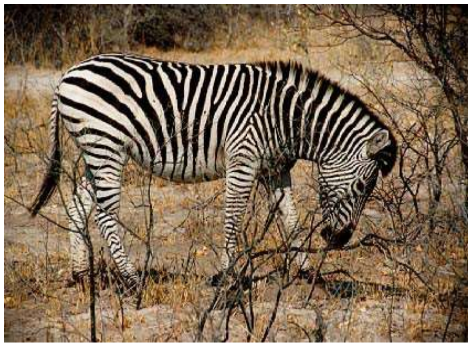




Outlines

- Introduction to object categorization
- Brief overview
  - Generative
  - Discriminative
- Generative models
- Discriminative models

Object Categorization: Statistical Approach



p(zebra | image)

vs

p(no zebra | image)

Bayes rule:

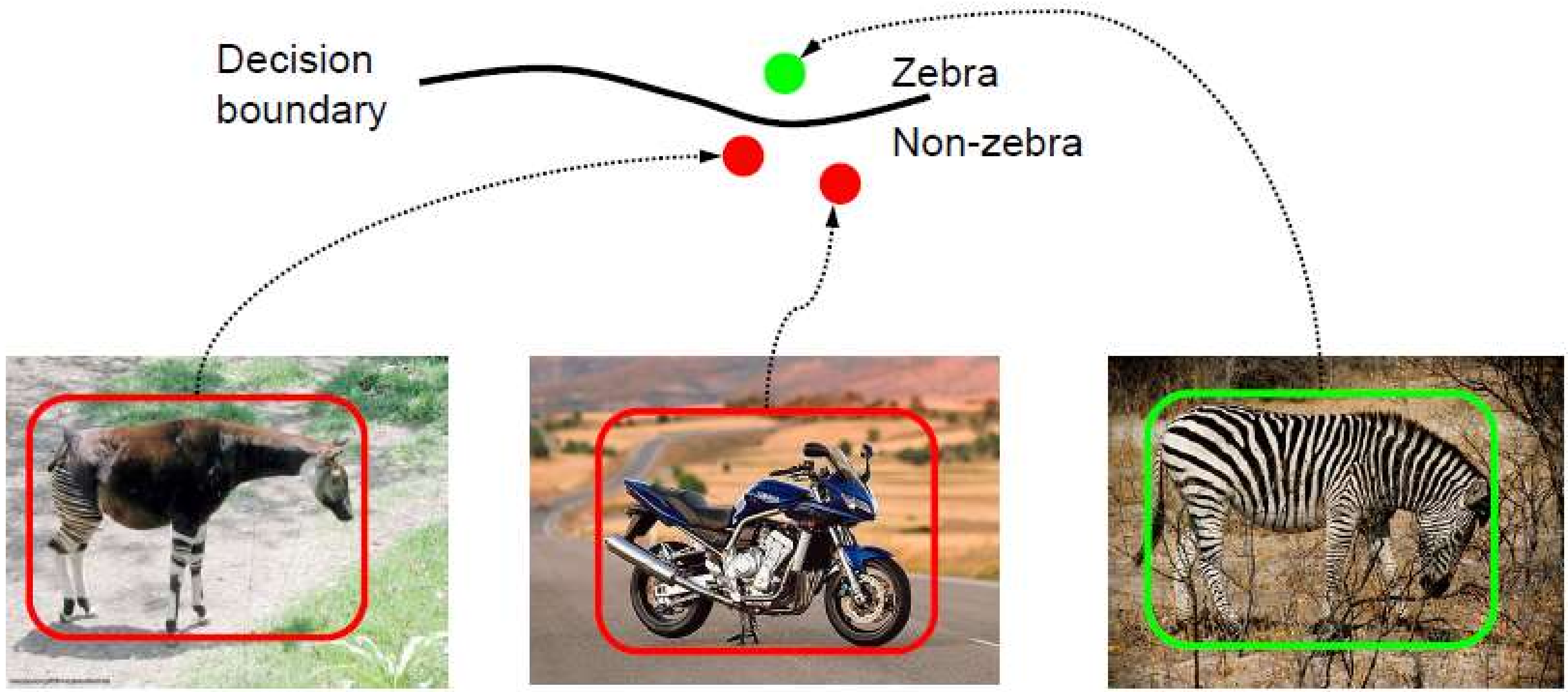
$$\frac{p(zebra|image)}{p(no\ zebra\ |\ image)} = \frac{p(image|zebra)}{p(imag|no\ zebra)} \cdot \frac{p(zebra)}{p(no\ zebra)}$$

Posterior ratio   Likelihood ratio   Prior ratio

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

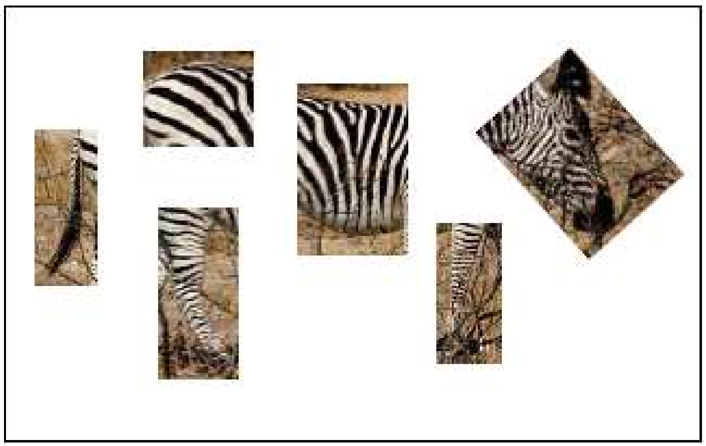
Discriminative Modeling

- Modelling of $\dfrac{p(zebra|image)}{p(no\ zebra\ |\ image)}$



Generative Modeling

- Modelling of $\frac{p(image|zebra)}{p(imag|no\ zebra)}$ and $\frac{p(zebra)}{p(no\ zebra)}$



	$p(image \mid zebra)$	$p(image \mid no\ zebra)$
	Low	Middle
	High	Middle→Low

Main steps of object categorization

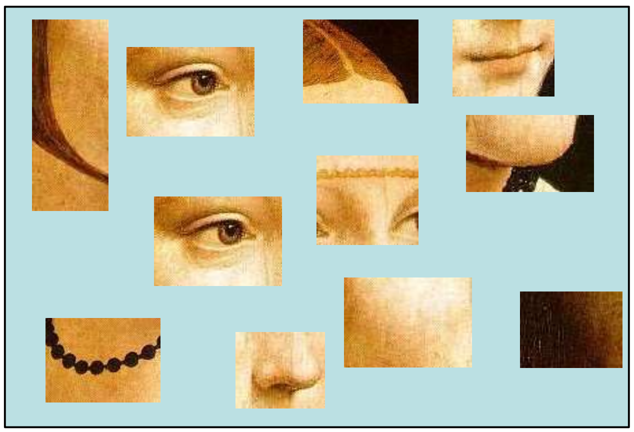
- Representation (Feature Selection)
  - How to represent an object category
- Learning (Training the classifier)
  - How to form the classifier, given training data
- Recognition (Testing Performance)
  - How the classifier is to be used on novel data

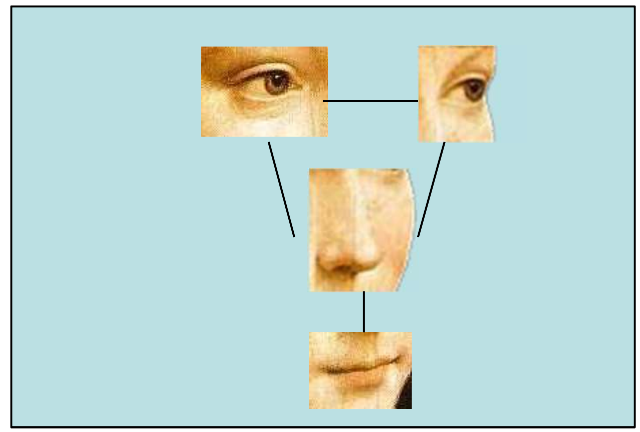
Representation

- Generative / discriminative / hybrid

Representation

- Generative / discriminative / hybrid
- Appearance only or location and appearance





Representation

- Generative / discriminative / hybrid
- Appearance only or location and appearance
- Invariance
  - View point
  - Illumination
  - Occlusion
  - Scale
  - Deformation
  - Clutter etc.

Representation

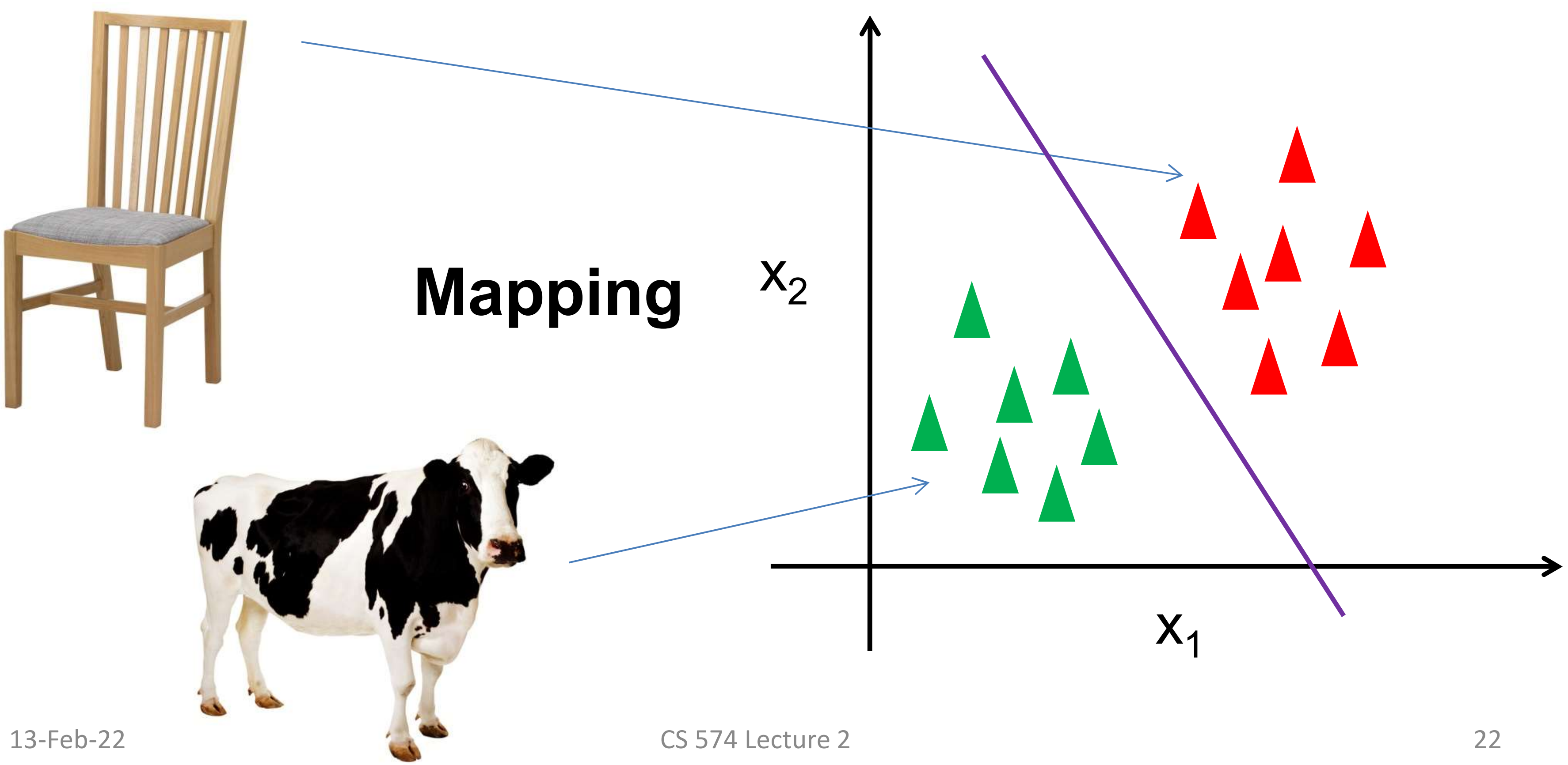
- Generative / discriminative / hybrid
- Appearance only or location and appearance
- Invariance
  - View point
  - Illumination
  - Occlusion
  - Scale
  - Deformation
  - Clutter etc.
- Local or global characteristics / statistics

Representation

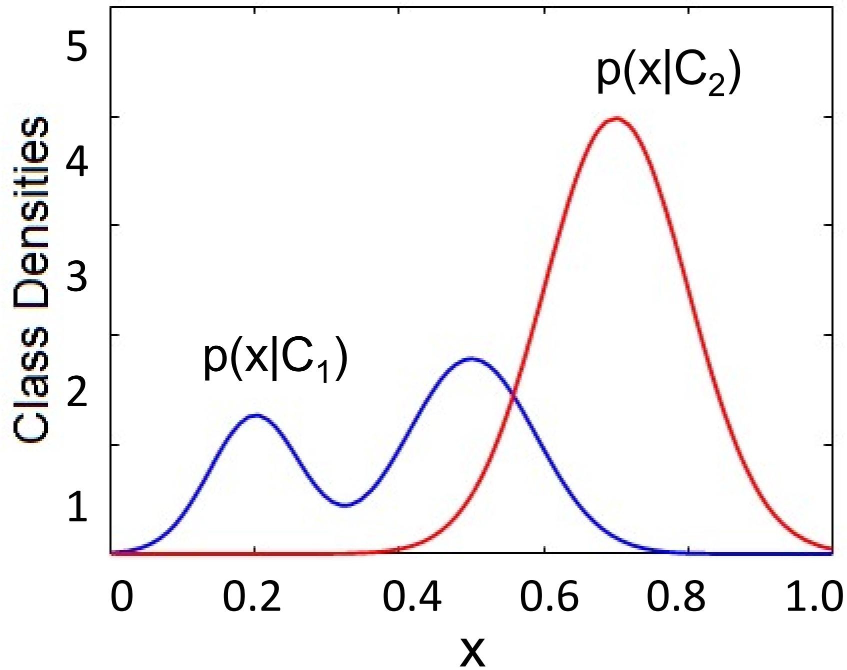
- Generative / discriminative / hybrid
- Appearance only or location and appearance
- Invariance
  - View point
  - Illumination
  - Occlusion
  - Scale
  - Deformation
  - Clutter etc.
- Local or global characteristics / statistics
- Use set of features or each pixel in image

Object Categorizations: Learning

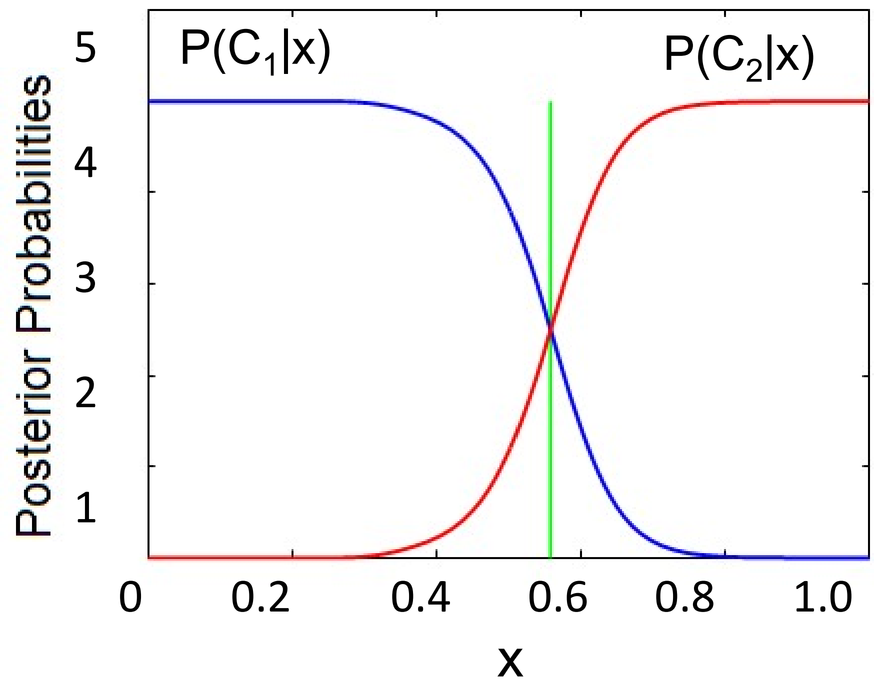
- Learn what distinguishes them rather than manually specify the difference



Methods of training: Generative vs. Discriminative



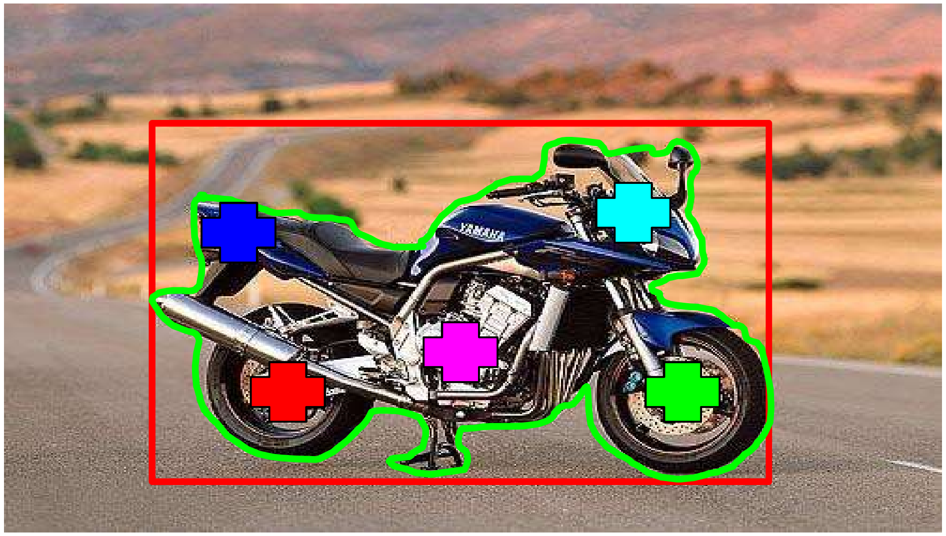
Generative: maximizing Likelihood



Discriminative: maximizing performances on train/validation set

Level of supervision

- Manual segmentation; bounding box; image labels; noisy labels



Recognition

- Performance
  - Scale / orientation range to search over
  - Speed

… to continue