Introduction to Recognition

Computer Vision

CS 543 / ECE 549

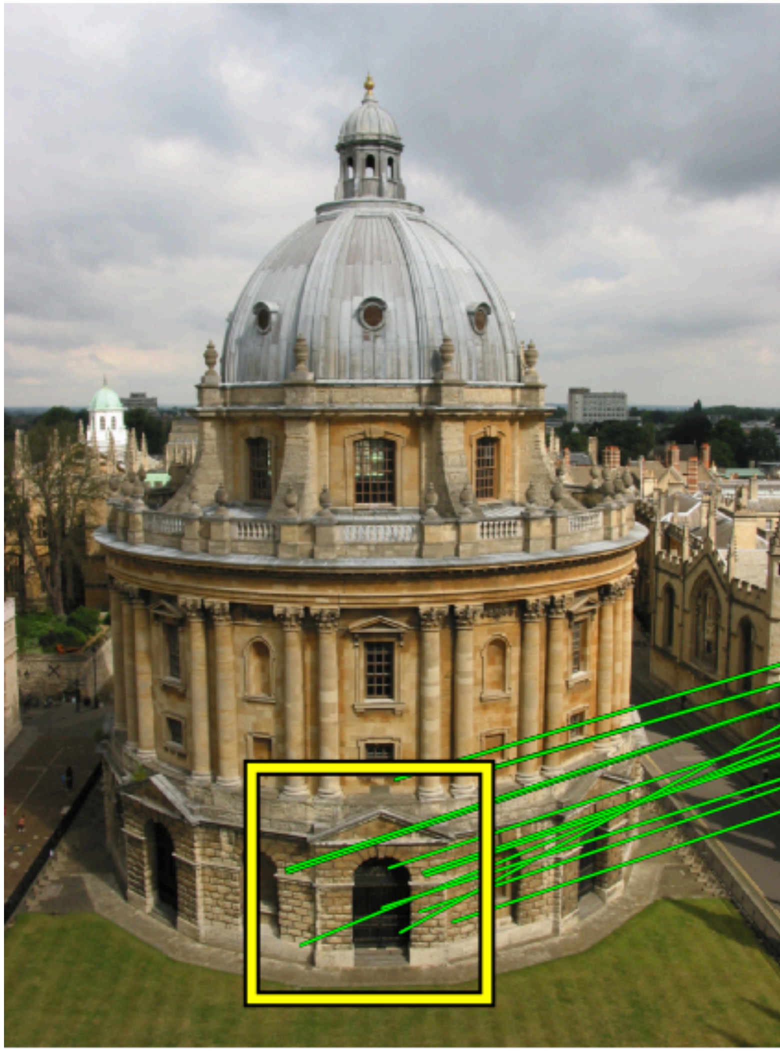
University of Illinois

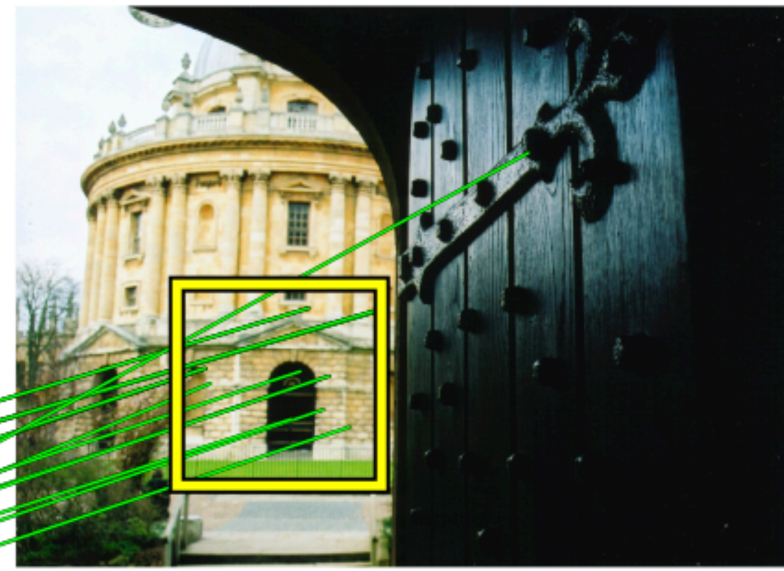
Many Slides from D. Hoiem, L. Lazebnik.

Outline

- Overview of image and region categorization
  – Task description
  – What is a category
- Example of spatial pyramids bag-of-words scene categorizer
- Key concepts: features and classification
- Deep convolutional neural networks (CNNs)

Recognition as 3D Matching





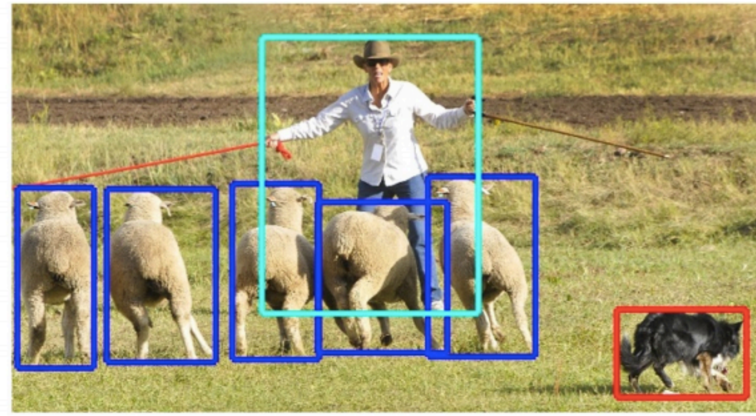
“Instance” Recognition

“Category-level” Recognition

Recognizing solid objects by alignment with an image. Huttenlocher and Ullman IJCV 1990.

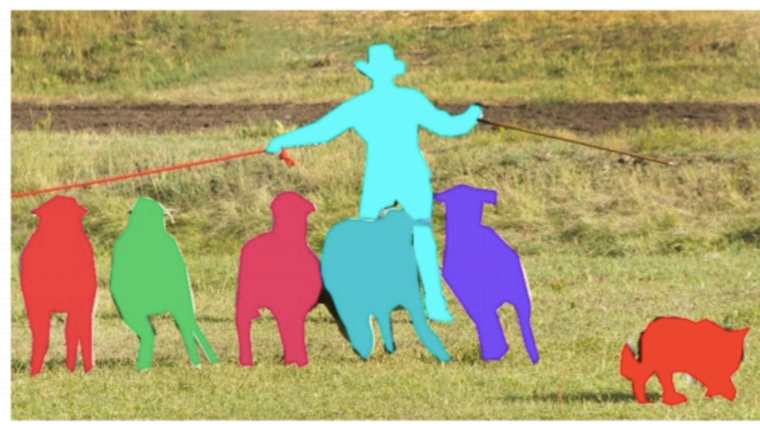
Detection, semantic segmentation, instance segmentation

image classification



object detection

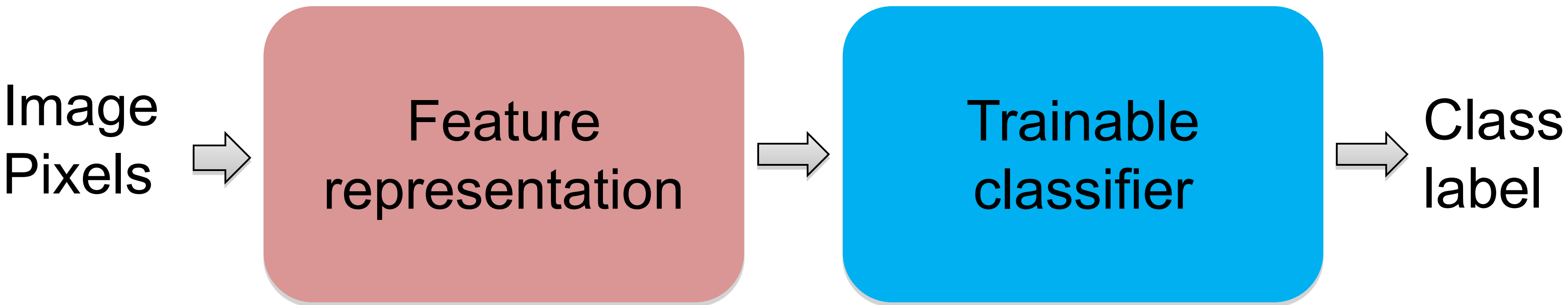
semantic segmentation



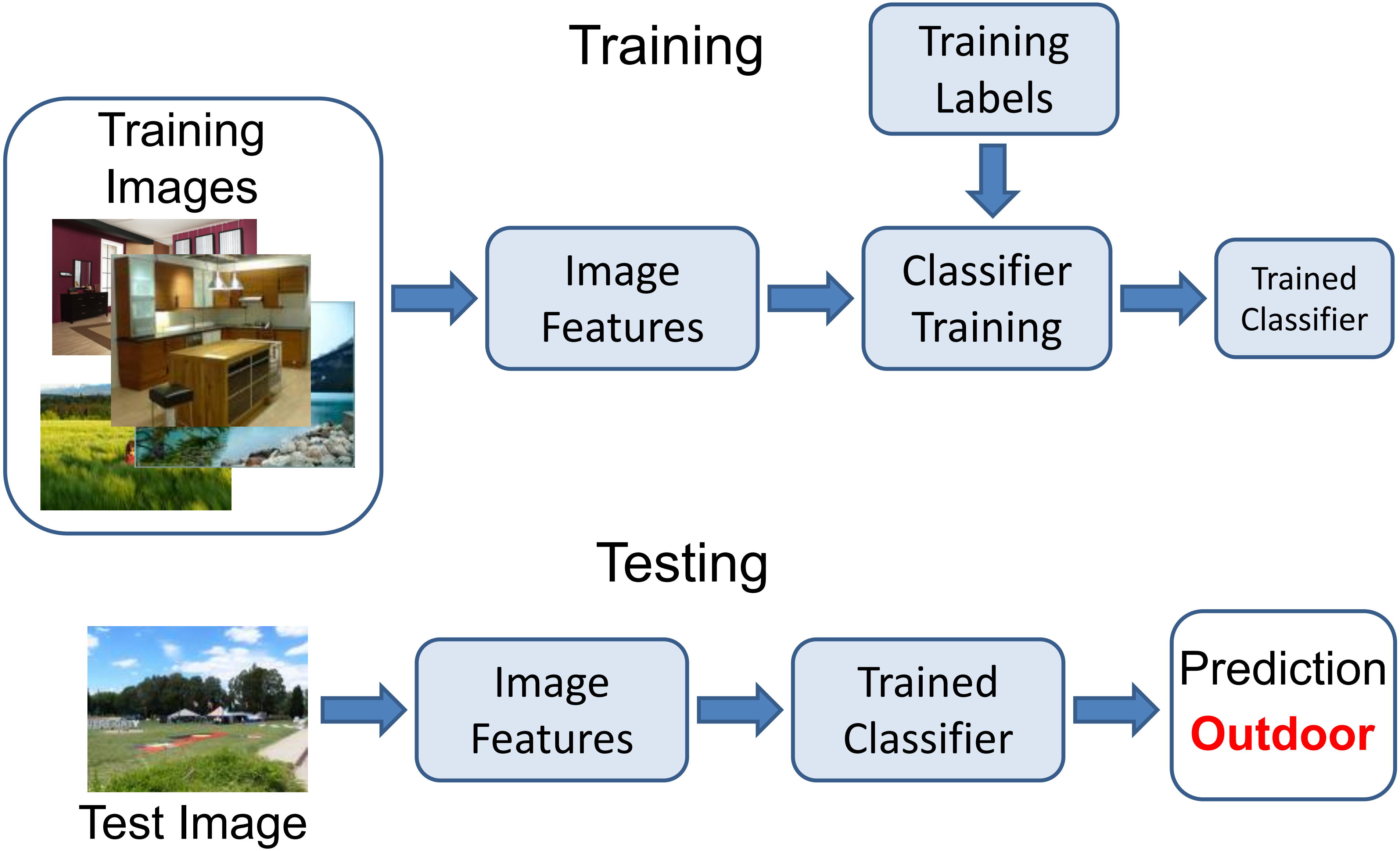
instance segmentation

Image source

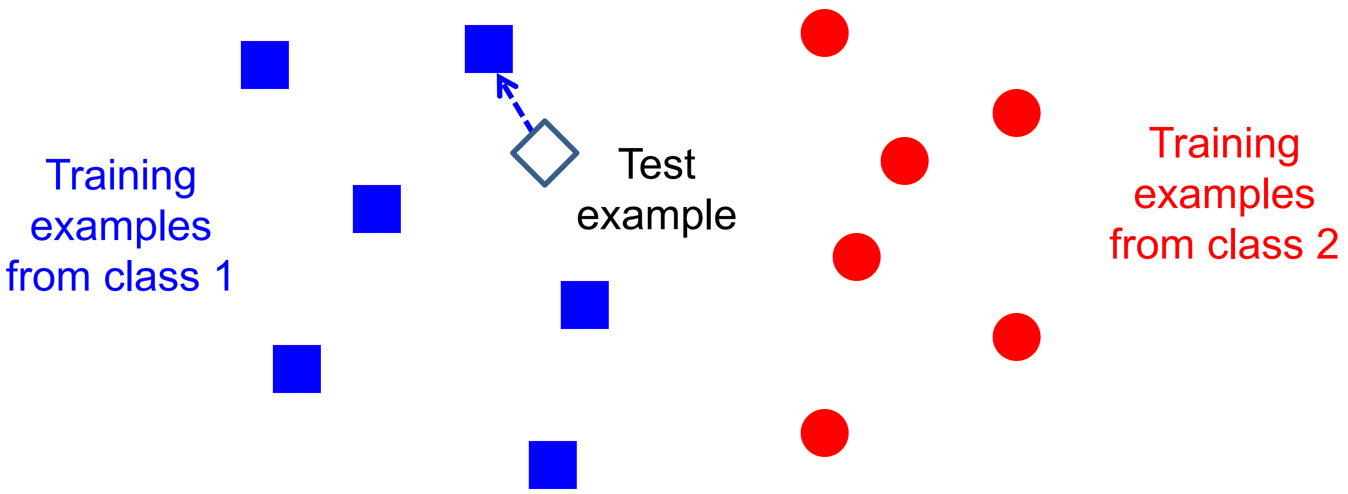
“Classic” recognition pipeline



Overview


Training
Training Labels
Training Images
Image Features
Classifier Training
Trained Classifier
Testing
Image Features
Trained Classifier
Prediction
Outdoor
Test Image

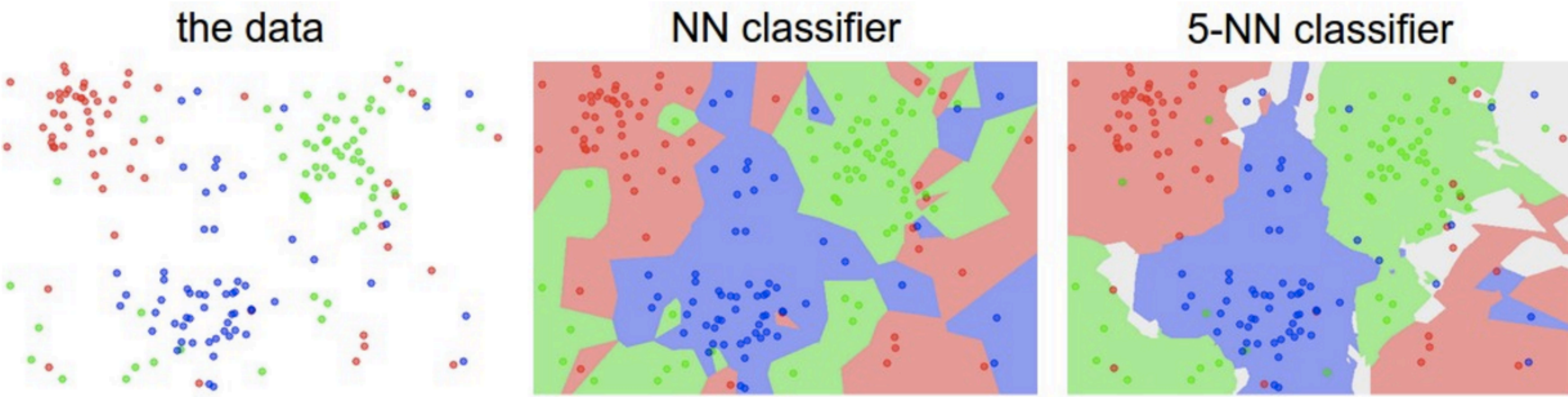
Classifiers: Nearest neighbor



f($\mathbf{x}$) = label of the training example nearest to $\mathbf{x}$

- All we need is a distance or similarity function for our inputs
- No training required!

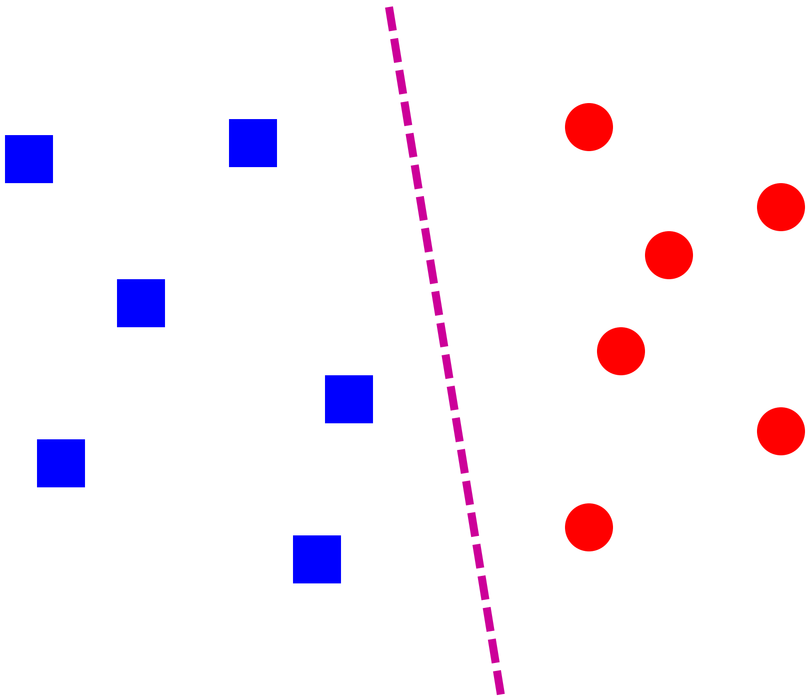
K-nearest neighbor classifier



- Which classifier is more robust to *outliers*?

Credit: Andrej Karpathy, http://cs231n.github.io/classification/

Linear classifiers

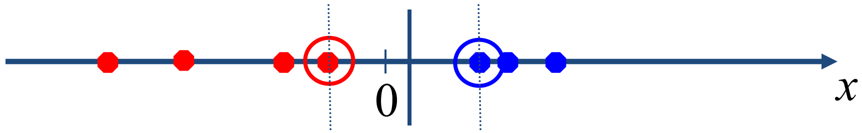


- Find a *linear function* to separate the classes:

$$f(\mathbf{x}) = \text{sgn}(\mathbf{w} \cdot \mathbf{x} + b)$$

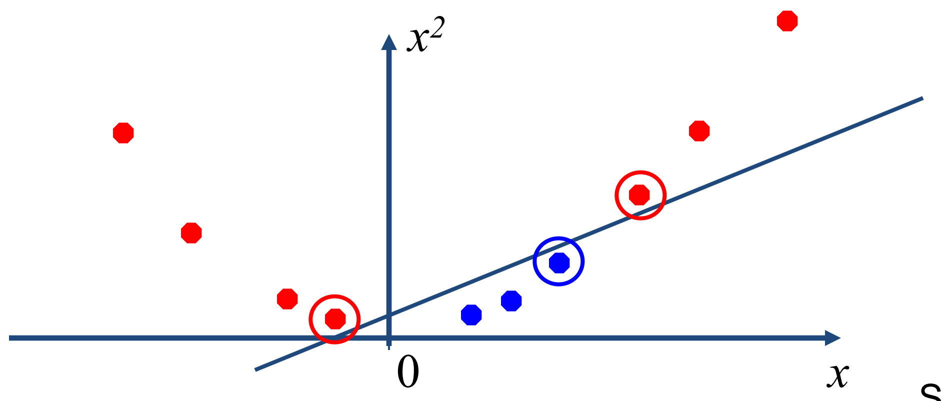
Nonlinear SVMs

- Linearly separable dataset in 1D:



- Non-separable dataset in 1D:

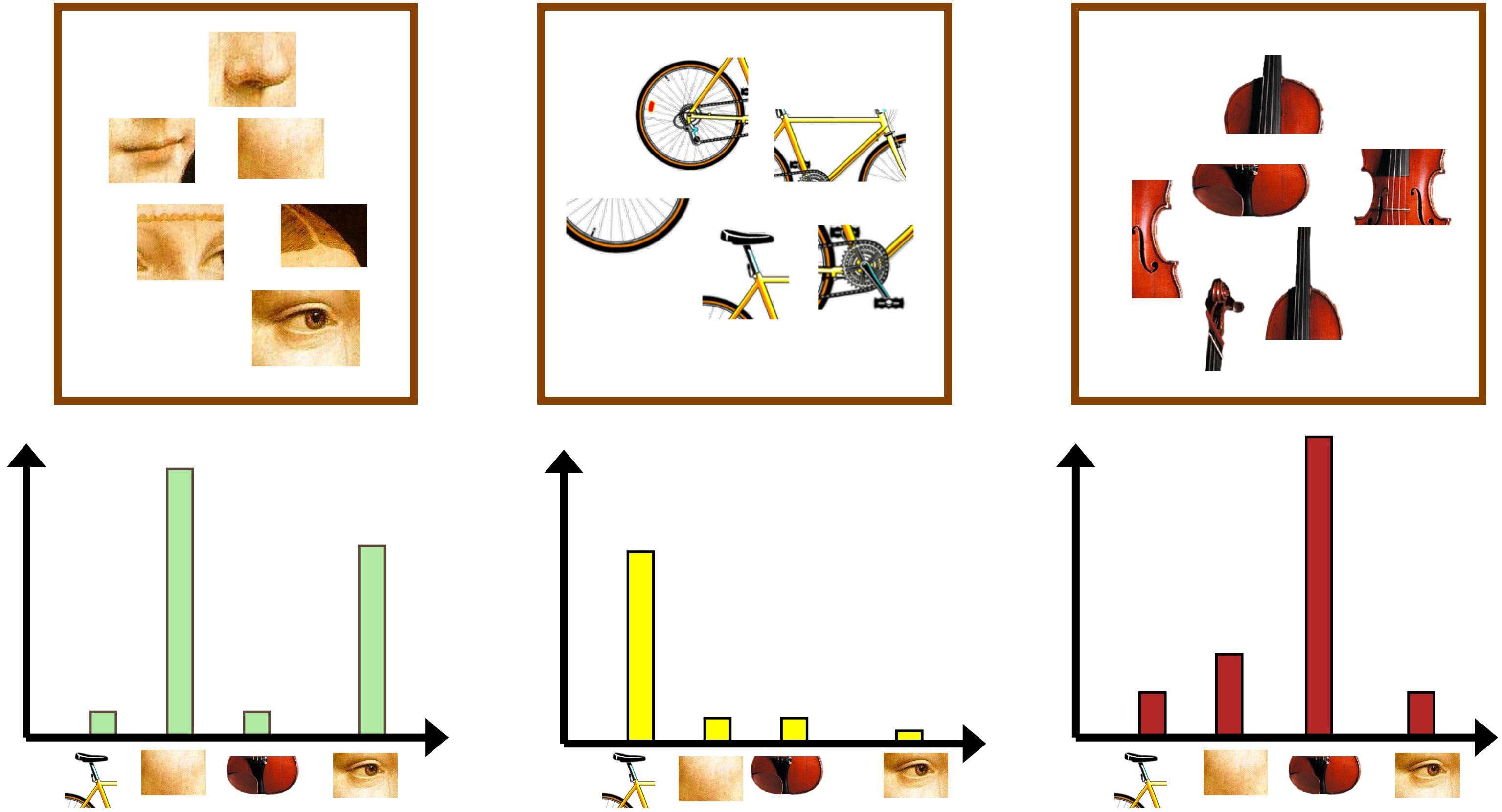
- We can map the data to a *higher-dimensional space*:



Slide credit: Andrew Moore

Bag of features

1. Extract local features
2. Learn "visual vocabulary"
3. Quantize local features using visual vocabulary
4. Represent images by frequencies of "visual words"

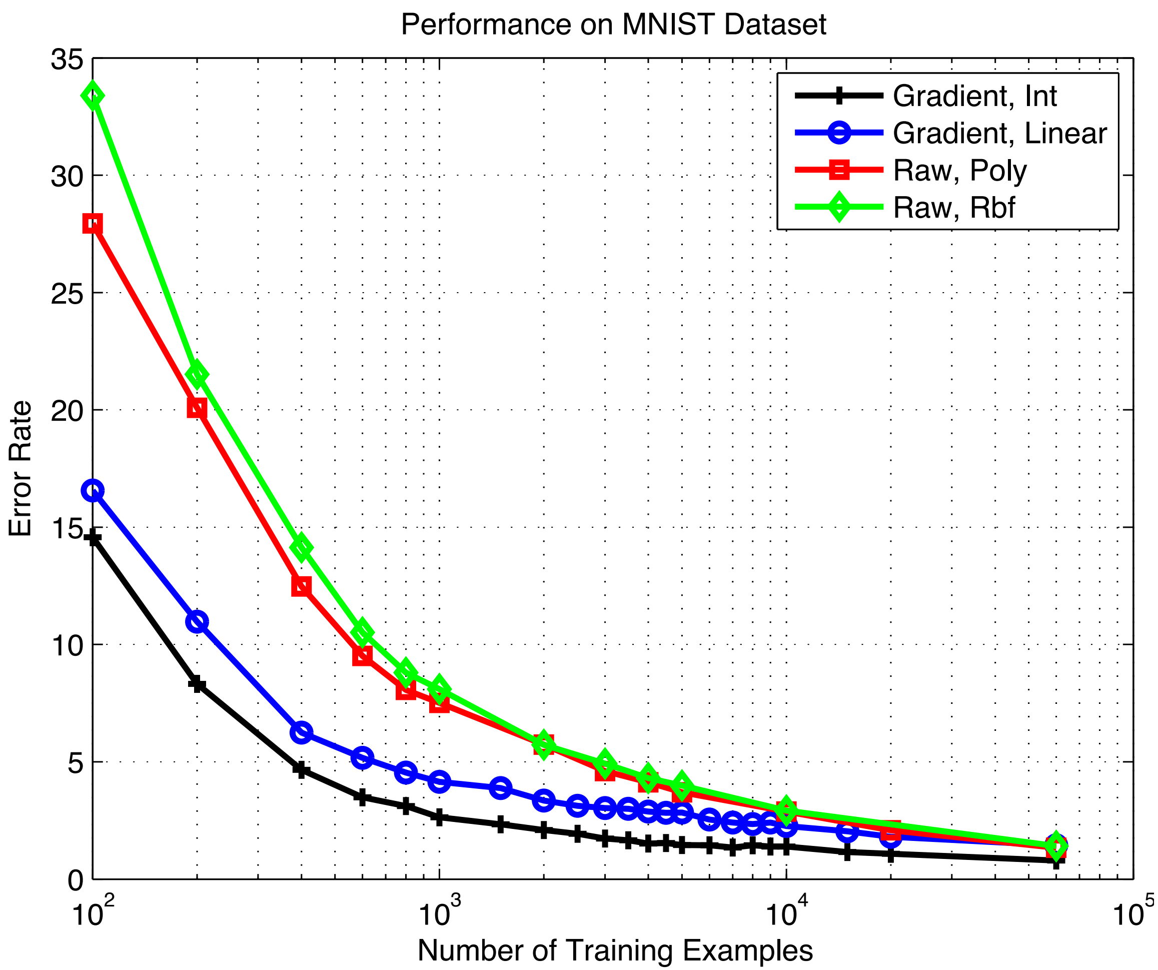


Digit Classification Case Study

The MNIST DATABASE of handwritten digits
Yann LeCun & Corinna Cortes

- Has a training set of 60 K examples (6K examples for each digit), and a test set of 10K examples.
- Each digit is a 28 x 28 pixel grey level image. The digit itself occupies the central 20 x 20 pixels, and the center of mass lies at the center of the box.

Bias-Variance Trade-off



Bias and Variance

Bias-Variance Trade-off

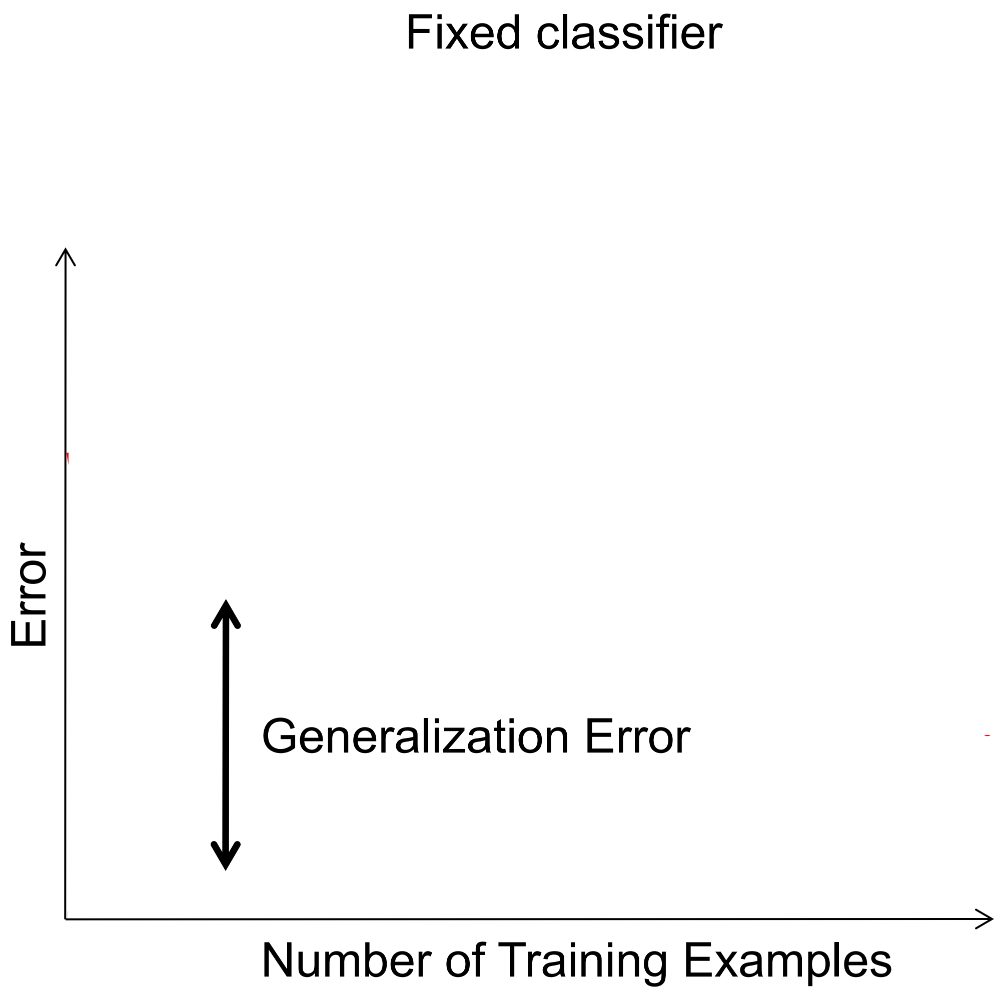
Performance as a function of model complexity (SVM)

Model Selection

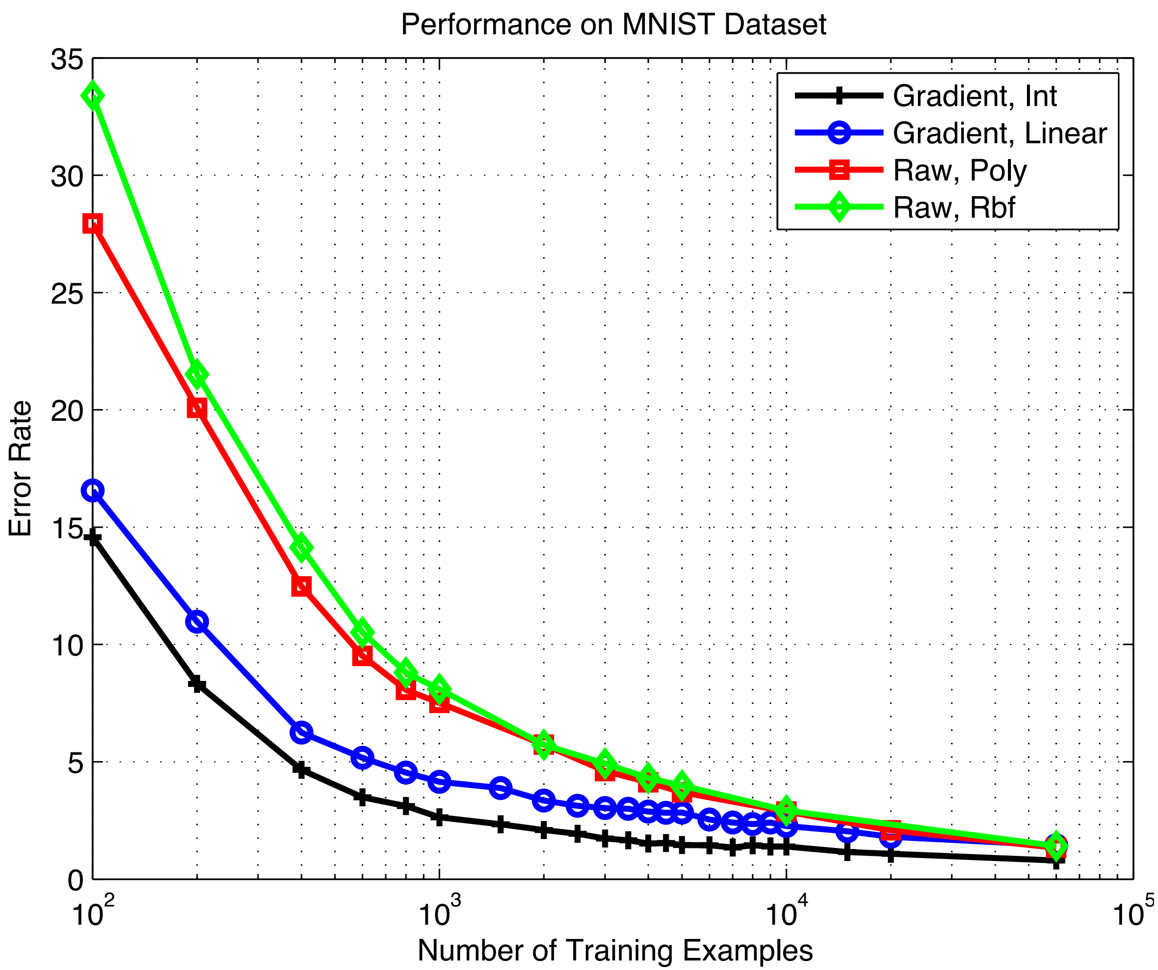
Bias-Variance Trade-off

As a function of dataset size

Generalization Error



Features vs Classifiers



What are the right features?

Depend on what you want to know!

- Object: shape
  - Local shape info, shading, shadows, texture
- Scene : geometric layout
  - linear perspective, gradients, line segments
- Material properties: albedo, feel, hardness
  - Color, texture
- Action: motion
  - Optical flow, tracked points

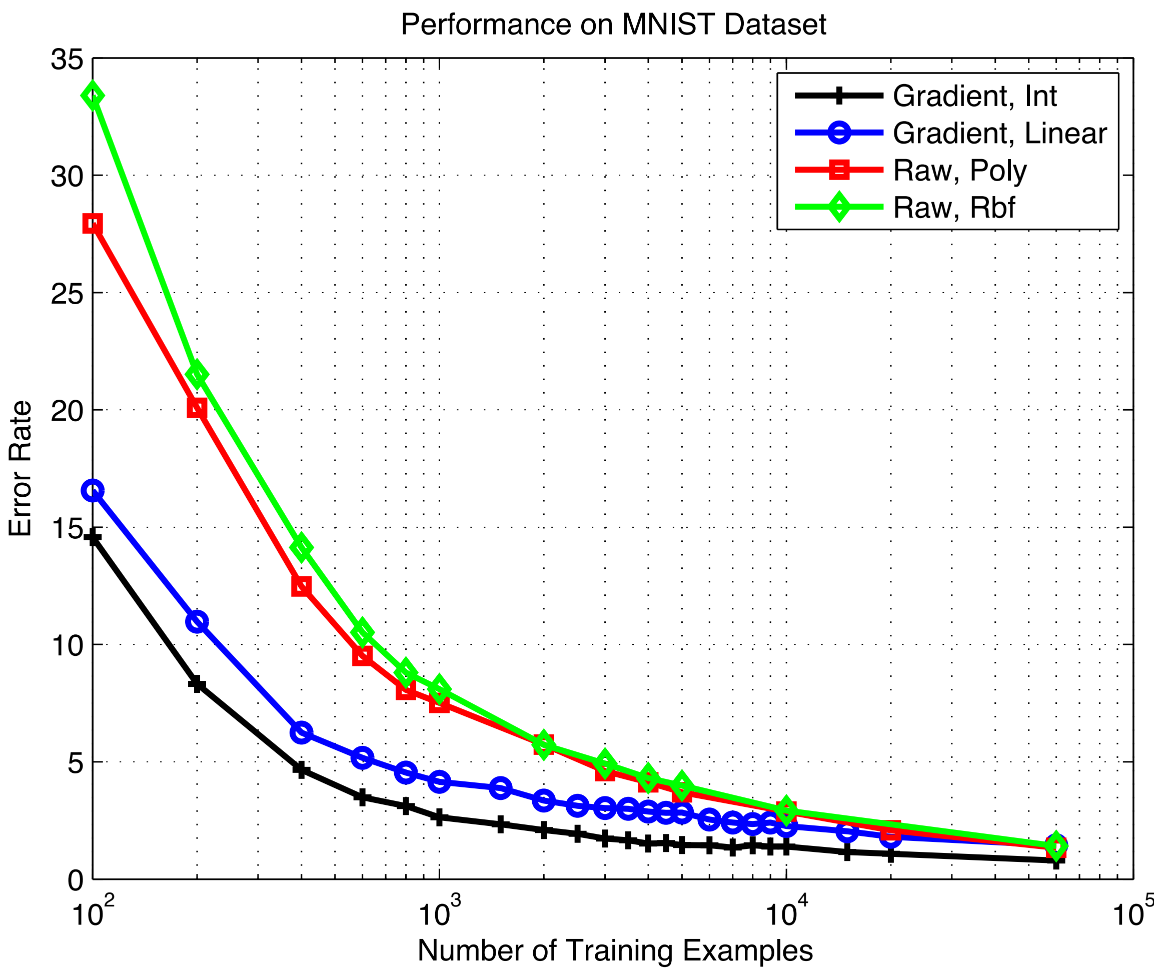
Stuff vs Objects

- recognizing cloth fabric vs recognizing cups

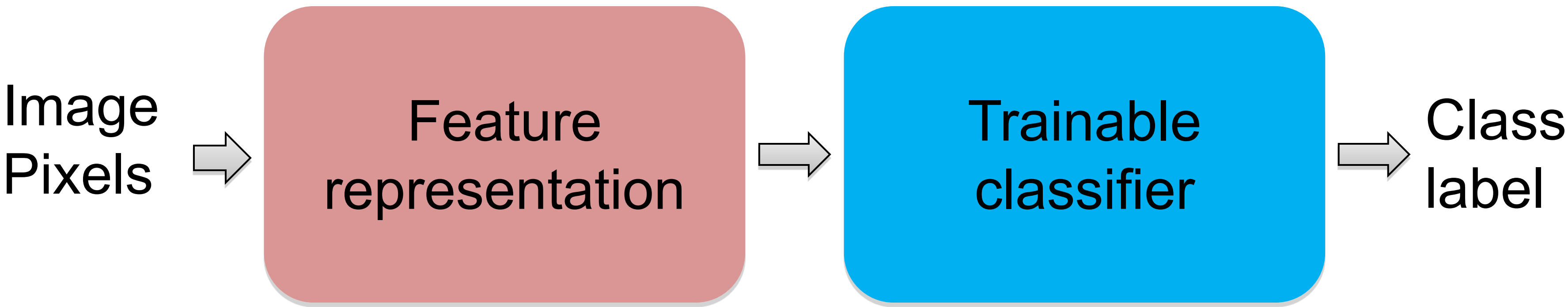
Feature Design Process

1. Start with a model
2. Look at errors on development set
3. Think of features that can improve performance
4. Develop new model, test whether new features help.
5. If not happy, go to step 1.
6. “Ablations”: Simplify system, prune out features that don’t help anymore in presence of other features.

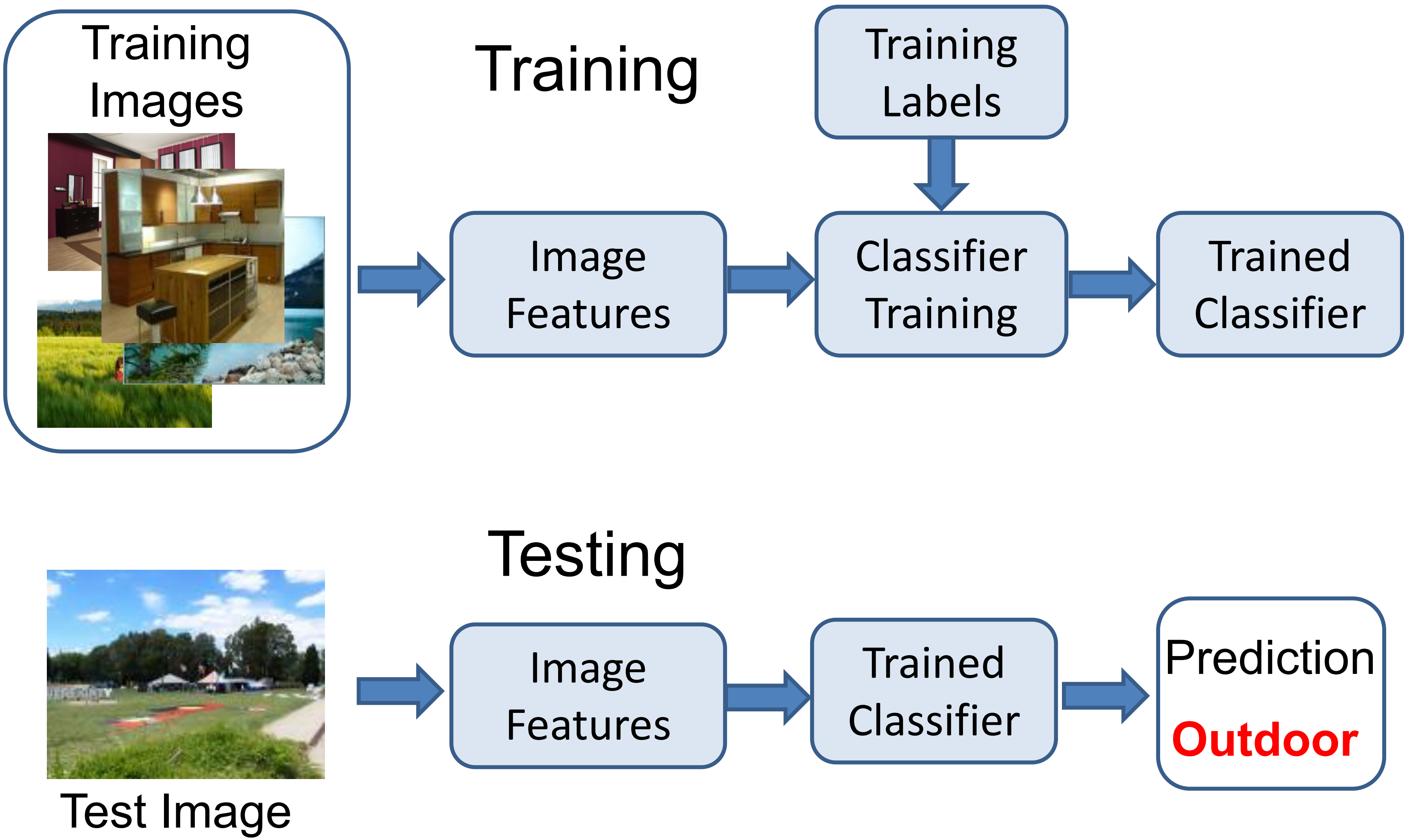
Features vs Classifiers



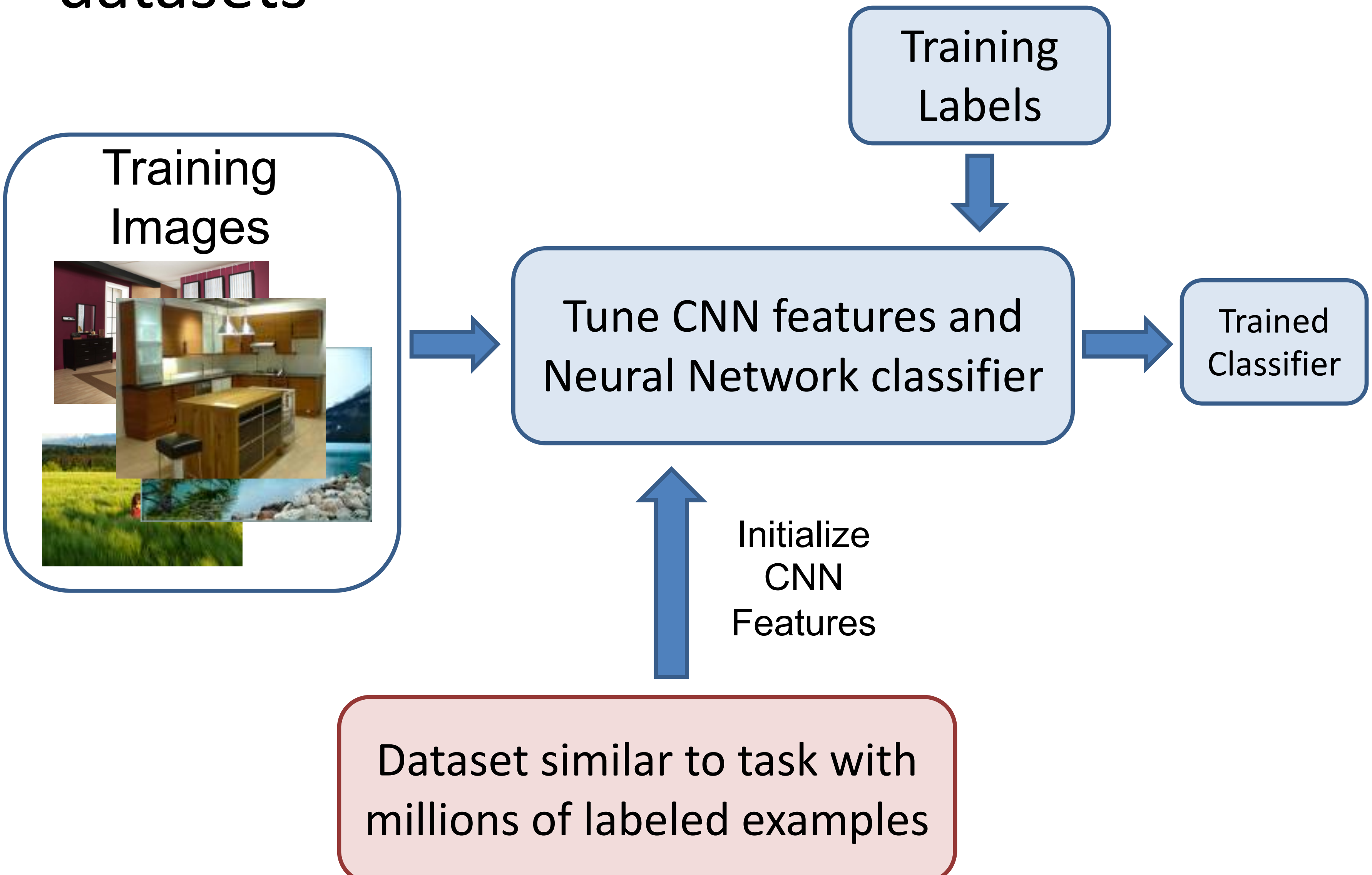
“Classic” recognition pipeline



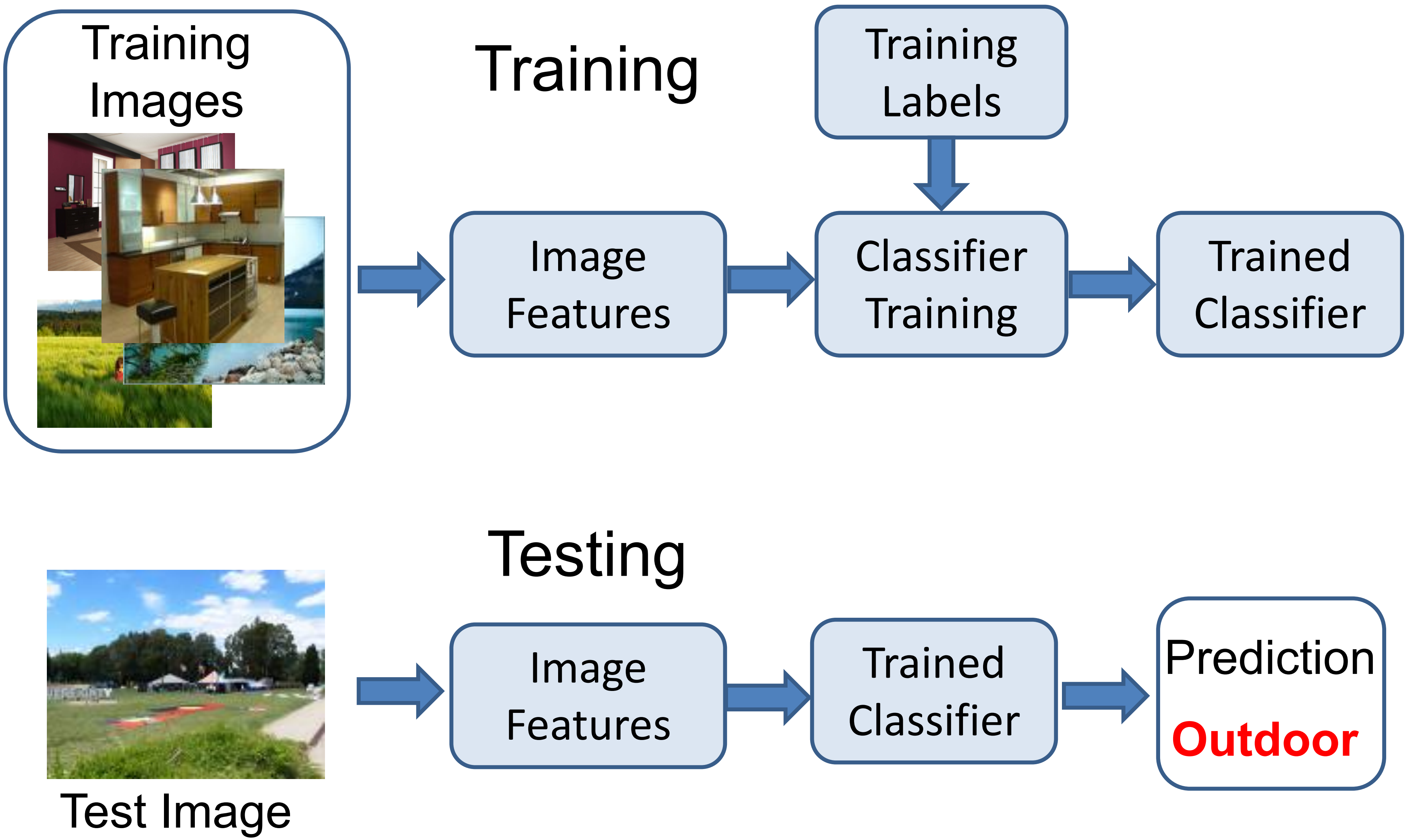
Categorization involves **features** and a classifier



New training setup with moderate sized datasets



Categorization involves **features** and a classifier



New training setup with moderate sized datasets