

What we covered!

- We started with challenges in **Visual Recognition**
- Some intuition about how to go about **designing features** (SIFT, HOG)
- Local Features to **Global Features** (Bag-of-Words, Spatial Pyramids)
- **Discriminative Models** (SVM, concepts of margin)
 - Feature Maps: Kernels
 - Speeding SVM using additive kernels.
- **Generative Models** (Naïve Bayes, pLSA, LDA)
 - Generative vs Discriminative Models
 - Graphical Models
 - Intuition of topic models.
- **Object Detections**
 - Dalal and Trigs Object Detection.
 - Poselets
 - PASCAL VOC
- **Were not able to cover**
 - Detailed object detection
 - Image Retrieval
 - Semantic Spaces
 - Cross-modal Retrieval

Generative

Discriminative



Object Segmentation

{Which pixels are from the object}

LARGE Scale Vision

{Billions of images}

{Thousands of tags}

Natural Image

Statistics

{How is the data truly distributed}

Multi-Label

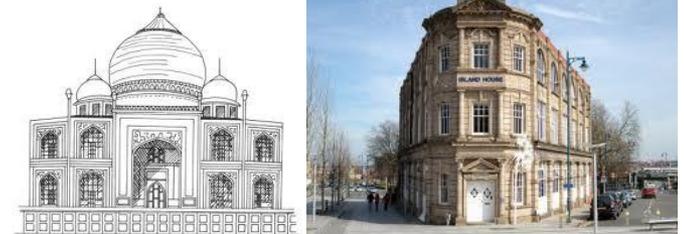
Classification

{Correlation between Class Labels}

{Hierarchy of Labels}



High Intra Class Variation



A picture is worth a thousand words.

Which words?

Fine-Grained

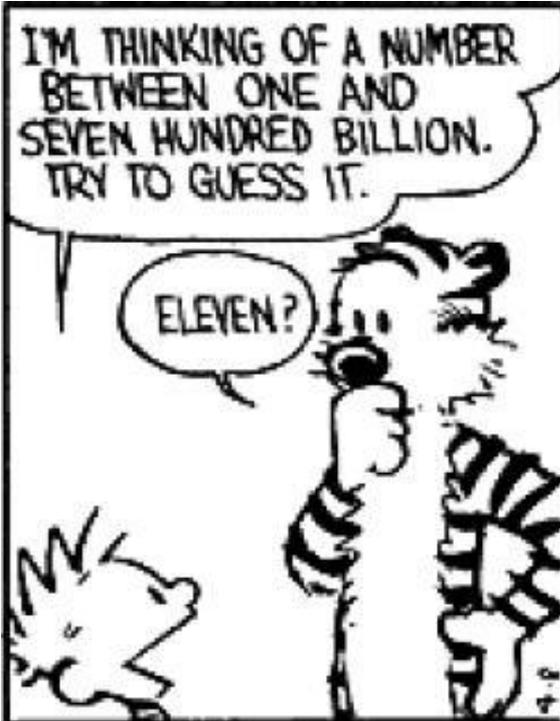
Recognition

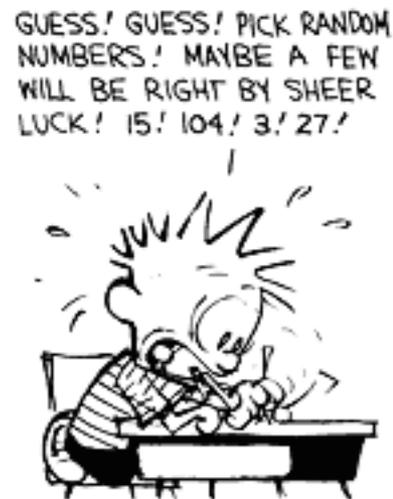
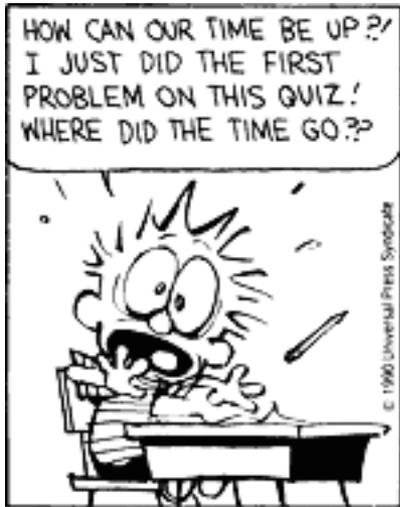
{Lotus or Water Lily}

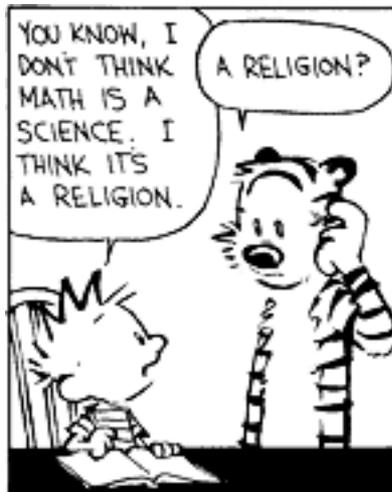
{What bird is it?}

Sparse Training

{Learn Class models from one/few image}







© 1987 Anderson. Distributed by Universal Uclick, Inc.

YEAH. ALL THESE EQUATIONS ARE LIKE MIRACLES. YOU TAKE TWO NUMBERS AND WHEN YOU ADD THEM, THEY MAGICALLY BECOME ONE *NEW* NUMBER! NO ONE CAN SAY HOW IT HAPPENS. YOU EITHER BELIEVE IT OR YOU DON'T.

