Paper Reading Seminar

Yan Wang

September 11, 2012

Combining Randomization and Discrimination for Fine-Grained Image Categorization

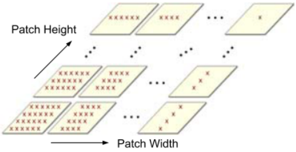
- Motivation

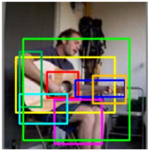
Class-generic detector • Objectness	Class-specific detector • Part model	Fine-grained detector • SVM + RF	Instance detector • Exemplar-SVM

  - Fine-grained image categorization
  - Bird species
  - Human activity classification
- Intuition
  - Dense sampling $\Rightarrow$ patches
  - Correlation among patches
  - Random forest + SVM

Approach

- Dense sampling

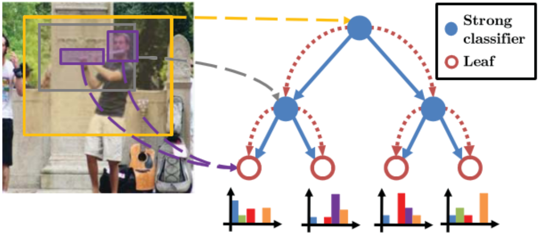




- Feature
  - Single patch: BoW
  - Patch pair: concatenation/intersection/absolute of difference of BoW histogram

Approach

- Random forest + SVM

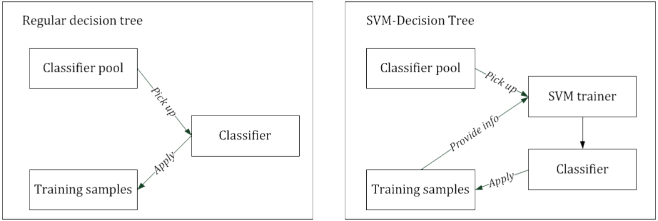


(b) Discriminative decision tree.

  - Randomly select patches (or patch pairs) + SVM
  - Train random forest with information gain
  - Use "ancestor" features
  - Q: invariance?

Discussions

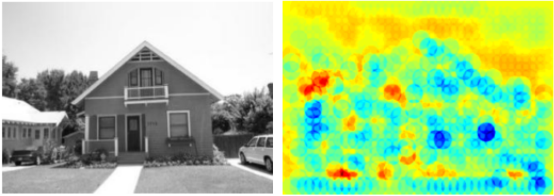
- How to train the SVM?
  - Images $\Rightarrow$ patches $\Leftrightarrow$ SVM
  - When: SVMs are trained right before evaluating the information gain



  - Train with features extracted from patches, image with labels
- How to maintain visual invariance?
  - Decision trees are able to achieve the invariance (appeared in the training data)

Application

- Heatmap



(a) Original Image (b) Heat Map Image

- Frequency a region picked up by the random forest
- Visualize from a classifier's perspective

Measuring the objectness of image windows

- Motivation

Class-generic detector	Class-specific detector	Fine-grained detector	Instance detector
• Objectness	• Part model	• SVM + RF	• Exemplar-SVM

  - Hand-crafted model ⇒ different from conventional "detectors"
- Applications
  - Preprocessing for detection
  - Visualization of classifier
  - Foreground/background separation?

Approach

- Intuition: an object should have...
  - A well defined closed boundary in space
  - A different appearance from its surroundings
  - Sometimes unique within the image (salient)
- Bayesian fusion of the cues

Approach

- Multi-scale saliency
  - Spectral residual of FFT in multiple scales
  - Measure the uniqueness of the window within the image

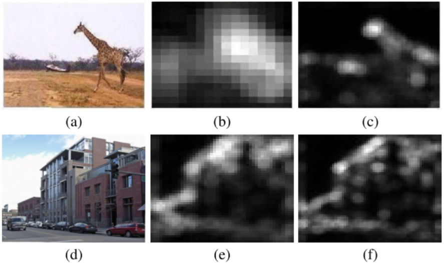
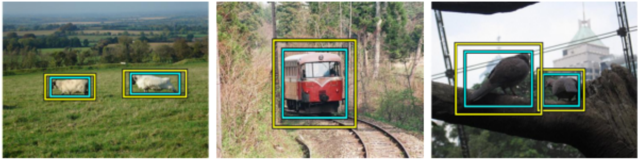


Fig. 2: **MS success and failure. Success:** *the large giraffe in the original image (a) appears as a blob in the saliency map for a high scale (b), while the tiny airplane in the map for a low scale (c). Having multi-scale saliency maps is important for finding more objects in challenging datasets. Interestingly, at the low scale the head of the giraffe is salient, rather than the whole giraffe.* **Failure:** *the numerous cars in the original image (d) are not salient at any scale. We show the saliency maps for 2 scales in (e) and (f). The contour of the building appears more salient than the cars.*

Approach

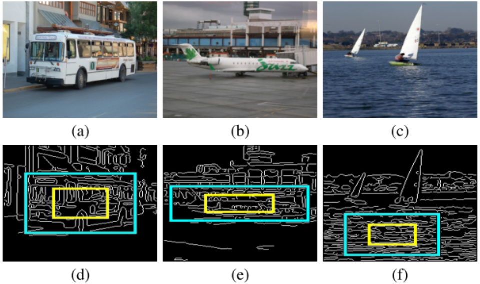
- Color contrast
  - Dissimilarity of a window to its immediate surrounding area
  - Chi-square distance between LAB histogram
  - Scores a whole window as whether it contains an entire project



(a) (b) (c)

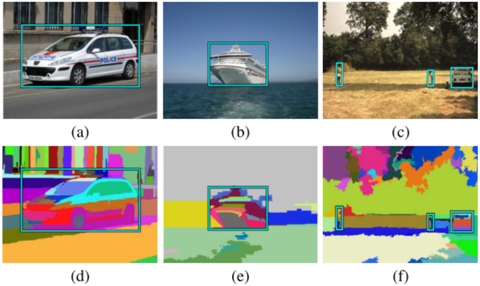
Approach

- Edge Density
  - Canny detector in the inner ring. Normalized with perimeter
  - Captures the closed boundary characteristic of objects



Approach

- Superpixel Straddling
  - Superpixels shouldn't cross the bondary
  - Rely on over-segmentation



Approach

- Spatial priori
  - Location and size
  - Kernel density estimation from training data
- Bayesian fusion

$$p(\text{obj}|C) = \frac{p(C|\text{obj})p(\text{obj})}{p(C)}$$