Classification and Feature Selection with Human Performance Data

Christina Pavlopoulou and Stella X. Yu
Computer Science Department
Boston College

Scene Classification Framework

Related Work

- Large amount of training data [Hays '07, Torralba '08]
- Learning from a few examples [Miller '00, Fei-Fei '06)]
- Semi-supervised Learning [Li '09, Fergus '09]
- Active Learning [Collins '08, Vijayanarasimhan '10]
- *Our approach: priors based on human performance*

Psychophysics-based “labelings”

- Images briefly seen by subjects may not be categorized perfectly.
- Categorization errors result from limited available features.
- Instead of “hard” labels, use categorization errors (“soft” labels).
- Benefits:
    - better generalizing classifiers
    - infer features employed by human vision

Ultra-Rapid Categorization

fixation dot
1sec

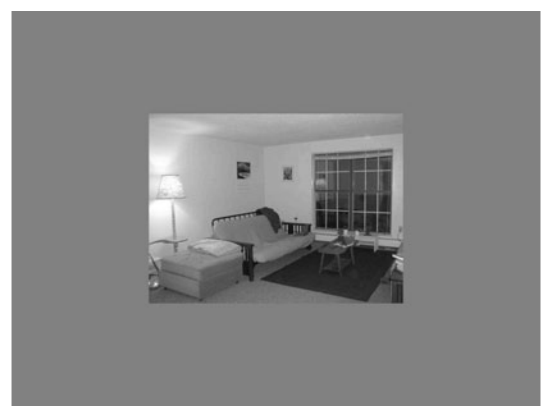
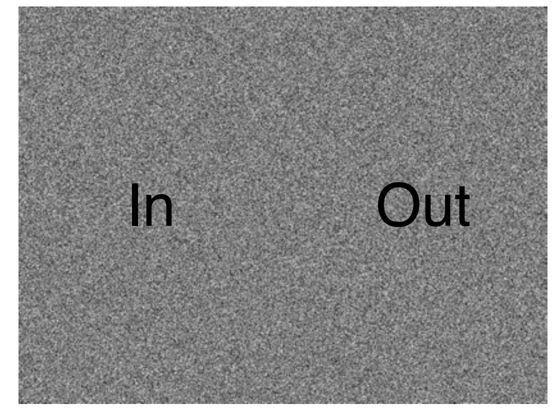
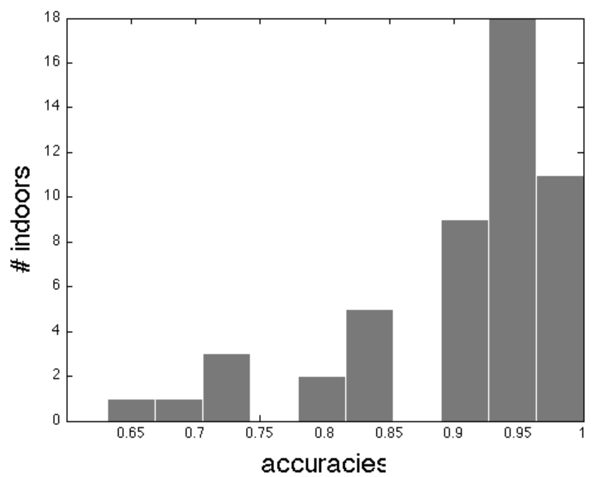


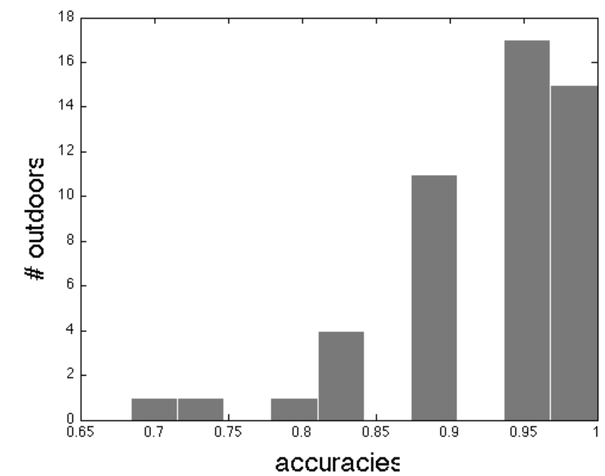
photo for 16ms



choice screen

Example Accuracies





100%

100%

100%

100%

79%

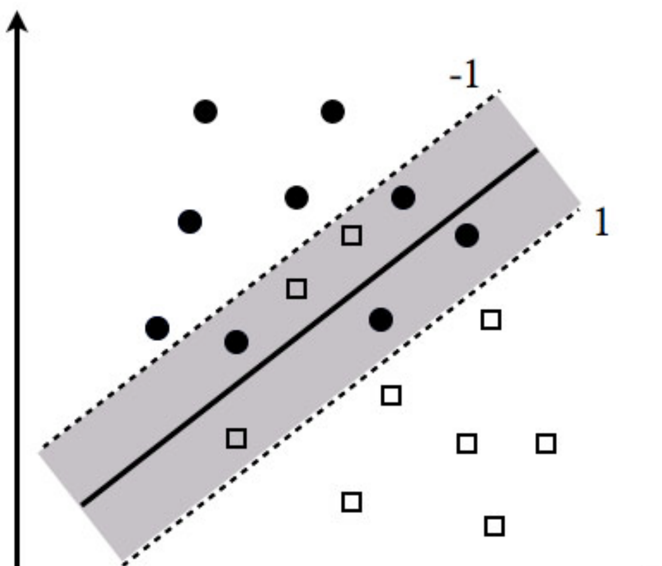
63%

74%

68%

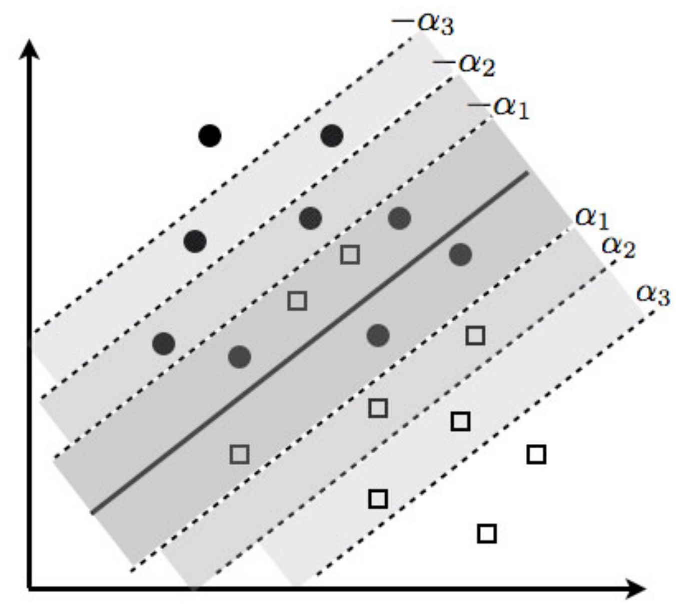
Formulation

SVM



$$\begin{aligned} \min \quad & ||\mathbf{w}||_p + C \textstyle\sum_i \xi_i \\ \text{s.t.} \quad & y_i \left(\mathbf{x_i} \cdot \mathbf{w} + b\right) && \geq 1 - \xi_i \\ & \xi_i && \geq 0, \quad i = 1, \dots n \end{aligned}$$

DSVM



$$\begin{aligned} \min \quad & ||\mathbf{w}||_1 + C \textstyle\sum_i \xi_i \\ \text{s.t.} \quad & y_i \left(\mathbf{x_i} \cdot \mathbf{w} + b\right) && \geq \alpha_i - \xi_i \\ & \xi_i && \geq 0, \quad i = 1, \dots n \end{aligned}$$

Non-linear SVM

$$\begin{array}{lll} \min & \|\mathbf{w}\|_1 + C\sum_i \xi_i & \\ \text{s.t.} & y_i\left(\sum_{j\neq i} y_j\, k(\mathbf{x_i},\mathbf{x_j})\, w_j + b\right) & \geq \alpha_i - \xi_i \\ & \xi_i & \geq 0, \quad i = 1,\ldots n \end{array}$$

Define kernel parameters with least squares fitting:

$$\min \sum_{i,j} [k(\mathbf{x_i},\mathbf{x_j}) - r(\mathbf{x_i},\mathbf{x_j})]^2$$

Kernel: $$k(\mathbf{x_i},\mathbf{x_j}) = \sum_{l=0}^{m} \beta_l h(\mathbf{x_i},\mathbf{x_j})^l$$

Perceptual correlation: $$r(\mathbf{x_i},\mathbf{x_j}) = \begin{cases} 1 - |A(i) - A(j)|, & \text{if } y_i = y_j \\ |A(i) - A(j)|, & \text{if } y_i \neq y_j \end{cases}$$

Experimental Evaluation

training:	50 (indoors)	50 (outdoors)
testing:	1,000 (indoors)	1,000 (outdoors)

		SVM (%)		DSVM (%)	
	# dim	indoors	outdoors	indoors	outdoors
gist	512	56.5	78.7	+5.6	+0.9
tiny images	3072	56.0	67.1	+4.2	-0.6
sparse SIFT	2000	74.9	62.8	+0.5	+0.3
textons	10752	69.3	86.2	-1.0	+0.9

Indoors correctly classified by SVM only

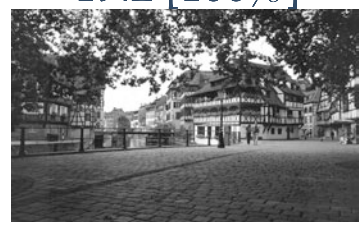
Indoors correctly classified by DSVM only

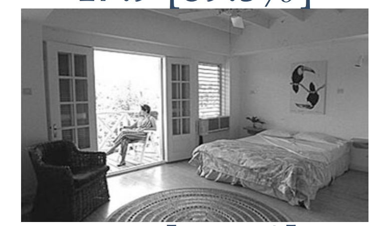
Support Vectors

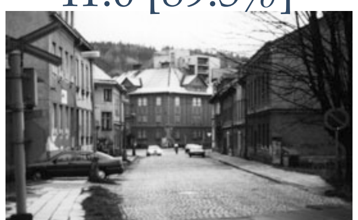
SVM

33.2 [89.5%]

29.3 [100%]

19.2 [100%]

18.5 [100%]

17.9 [89.5%]

15.4 [100%]


11.0 [89.5%]

10.9 [68.4%]

10.8 [94.7%]

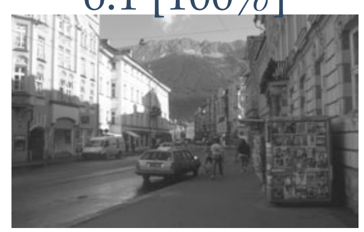
9.8 [100%]

7.7 [100%]

7.7 [100%]


7.3 [100%]

7.0 [94.7%]

6.1 [100%]

4.9 [94.7%]

4.7 [94.7%]

4.4 [94.7%]

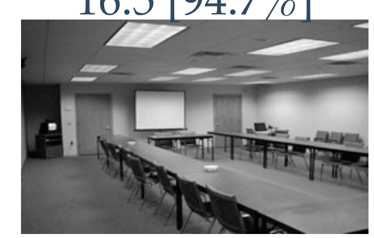

DSVM

24.6 [9.5%]

21.8 [100%]

20.9 [100%]

17.1 [89.5%]

16.5 [94.7%]

16.1 [100%]


11.6 [94.7%]

10.8 [94.7%]

10.1 [73.7%]

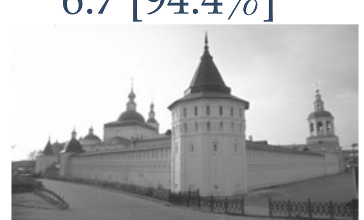
9.2 [94.7%]

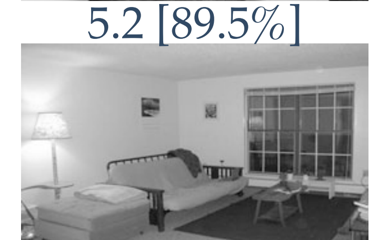
8.5 [100%]

7.9 [94.7%]

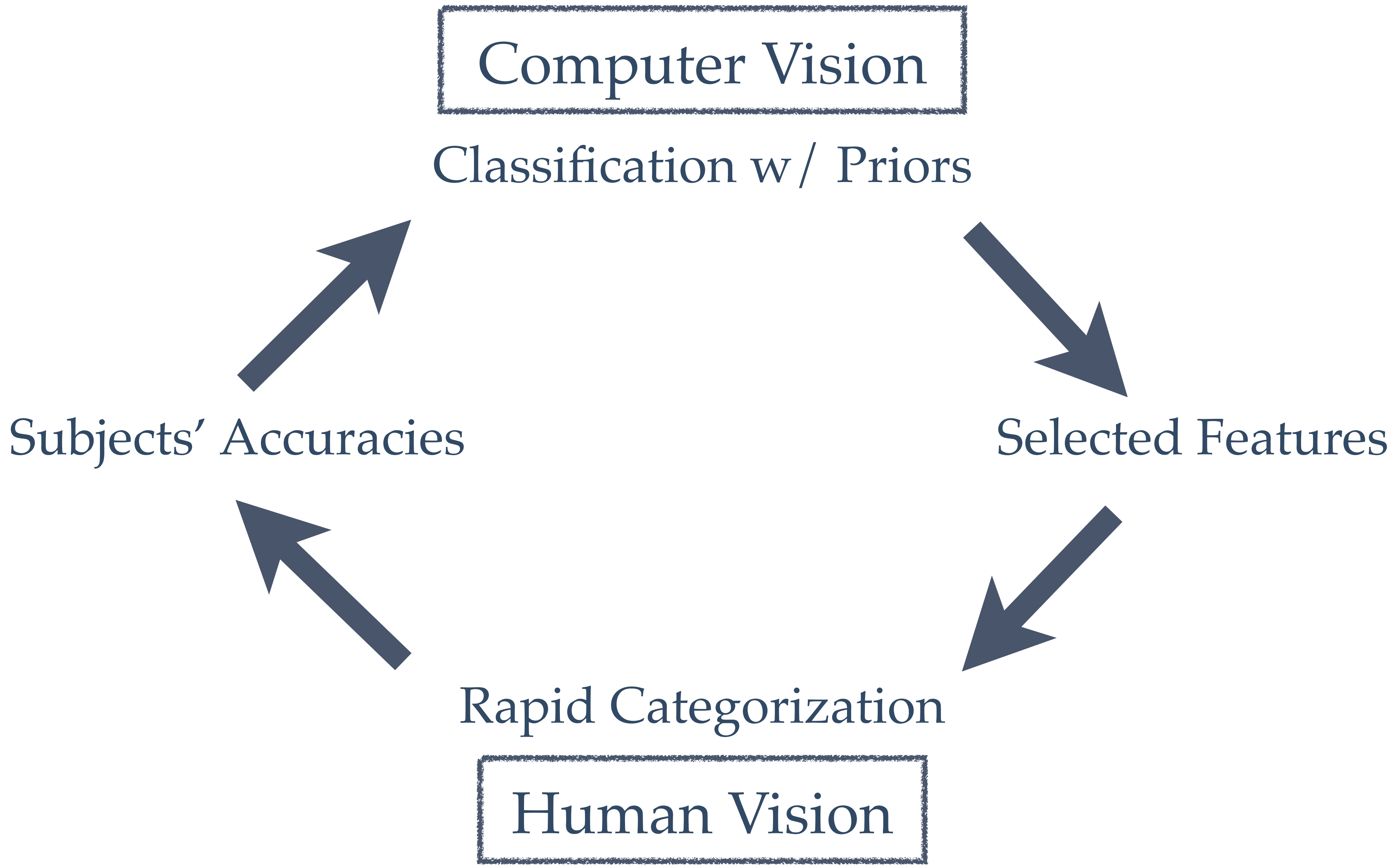

7.3 [94.7%]

6.9 [89.5%]

6.7 [94.4%]

5.3 [84.2%]

5.2 [89.5%]

3.0 [94.7%]


Summary


Computer Vision
Classification w/ Priors
Selected Features
Rapid Categorization
Human Vision
Subjects' Accuracies