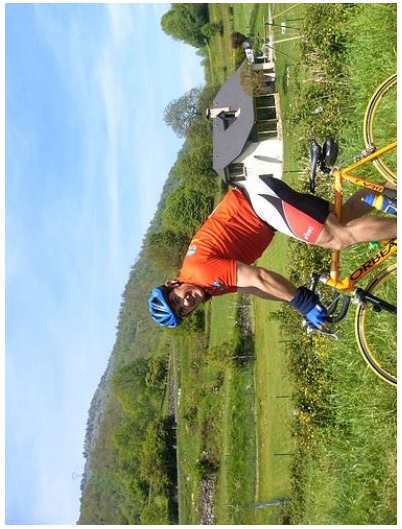
Automatic Image Orientation Detection

Mike Ghesquiere

CS9840 – Machine Learning for Computer Vision

Problem



Which orientation is correct?

Solution

Automatic Image Orientation (A. Vailaya et al)

Pipeline

Feature Extraction

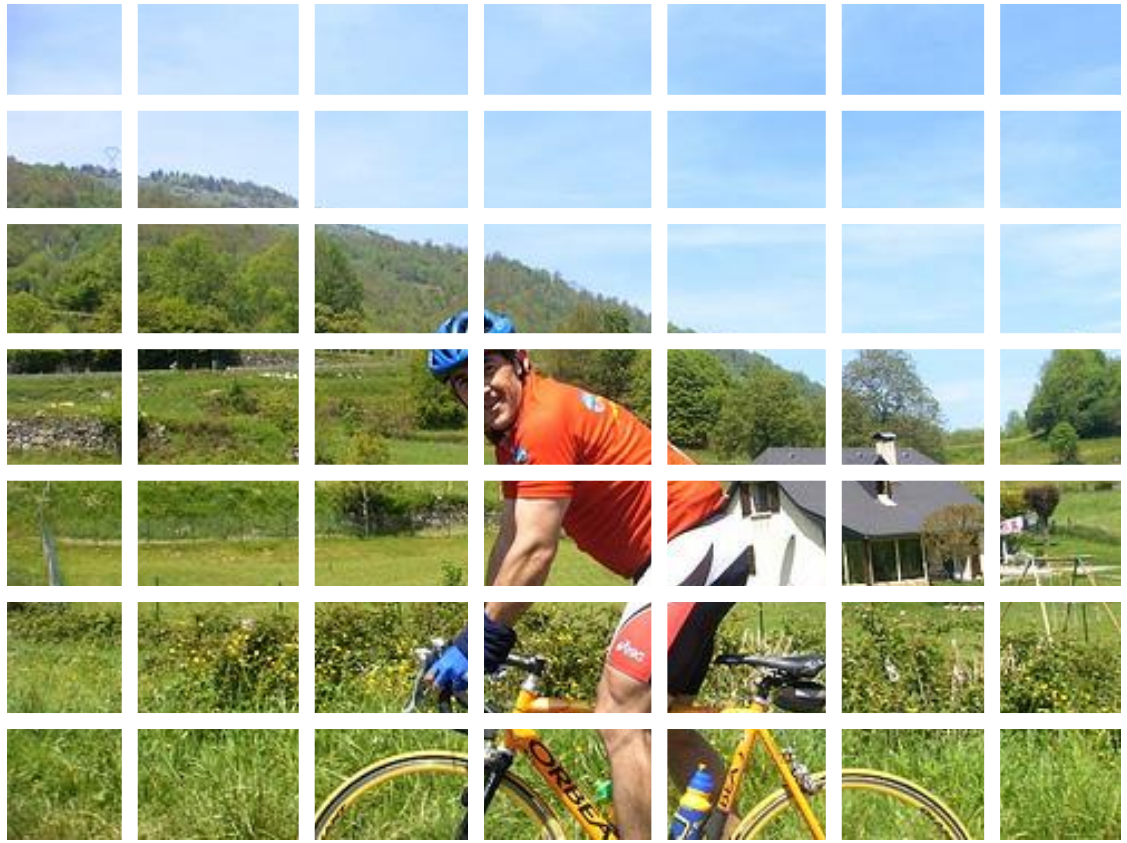
Feature Selection

Learning

Results

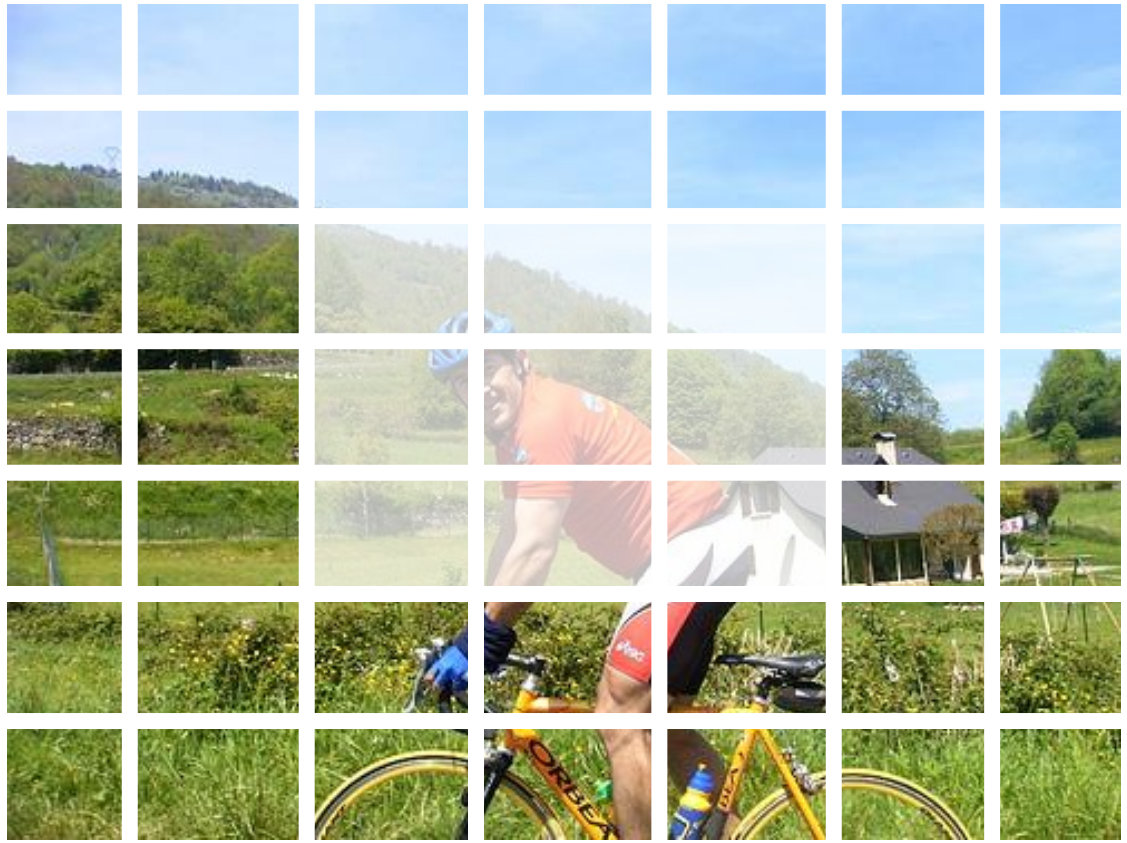
Feature Extraction

Split the image into $N \times N$ blocks.



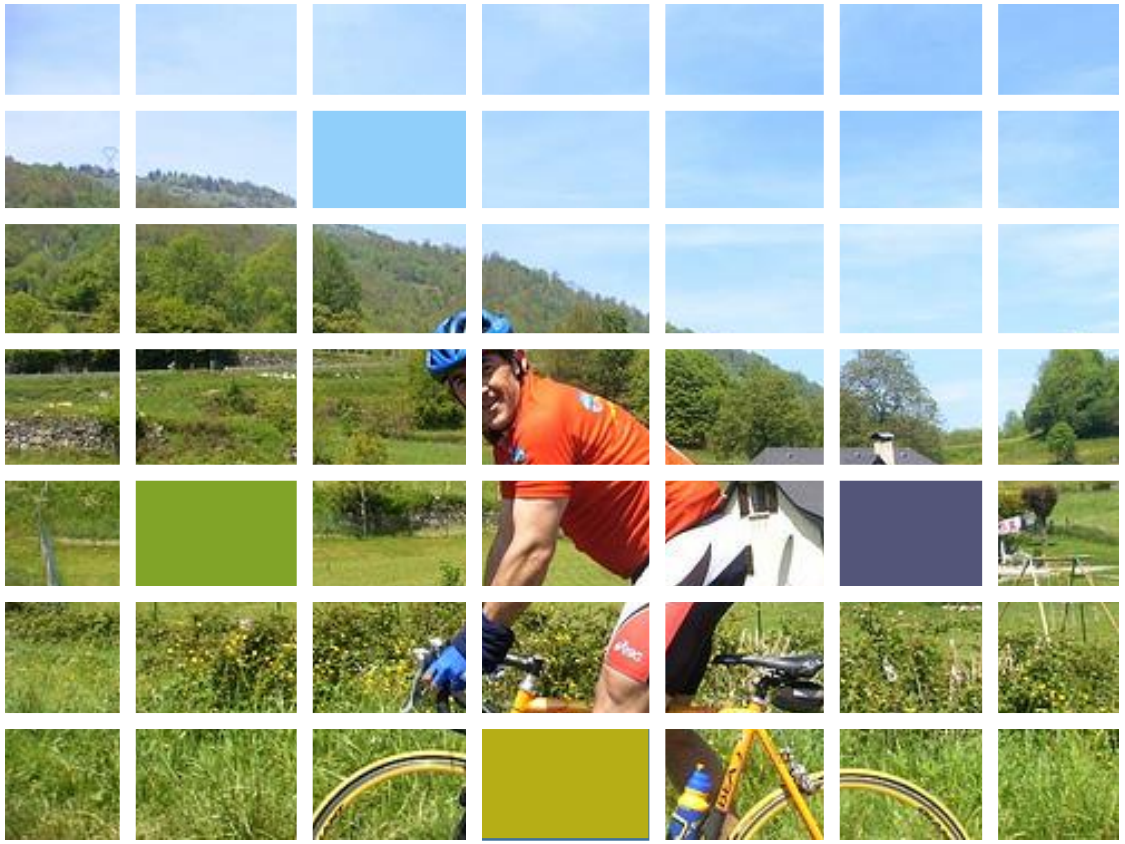
Feature Extraction

(the center can be discarded)



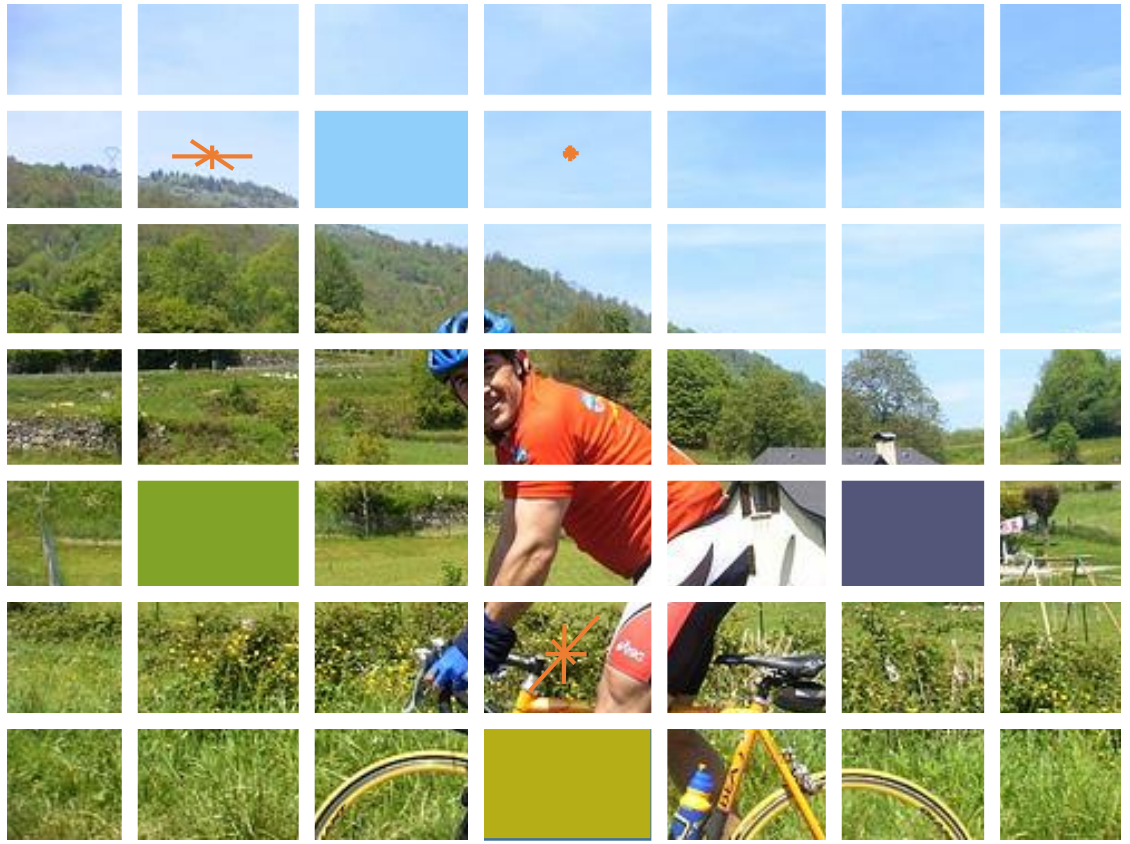
Feature Extraction

Extract features from each block.



Feature Extraction

Extract features from each block.



Feature Extraction

Color Moments (CM) in L* u* v*

Edge Detection Histogram (EDH)

Feature Extraction

$$N_{CM} \times N_{CM} \times \{mean, variance\} \times \{L, U, V\}$$

$$N_{EDH} \times N_{EDH} \times (histogram\ buckets)$$

Feature Extraction

$$10 \times 10 \times 2 \times 3$$

$$0 \times 0$$

600 features

Feature Extraction

	N_{CM}	N_{EDH}	EDH bins	Length of feature vector
Vailaya et al. [1]	10	0	-	600
Luo et al. [2]	7	5	16+1	719
Takahashi et al. [3]	7	5	16+1	719
Cingovska et al. [4]	(8)[1]	(8)[1]	(8+1)[2]	720
Le Borgne et al. [5]	4	4	(4+1)[2]	176
Liu et al. [6]	5	5	12+1	475

Pipeline

Feature Extraction

Feature Selection

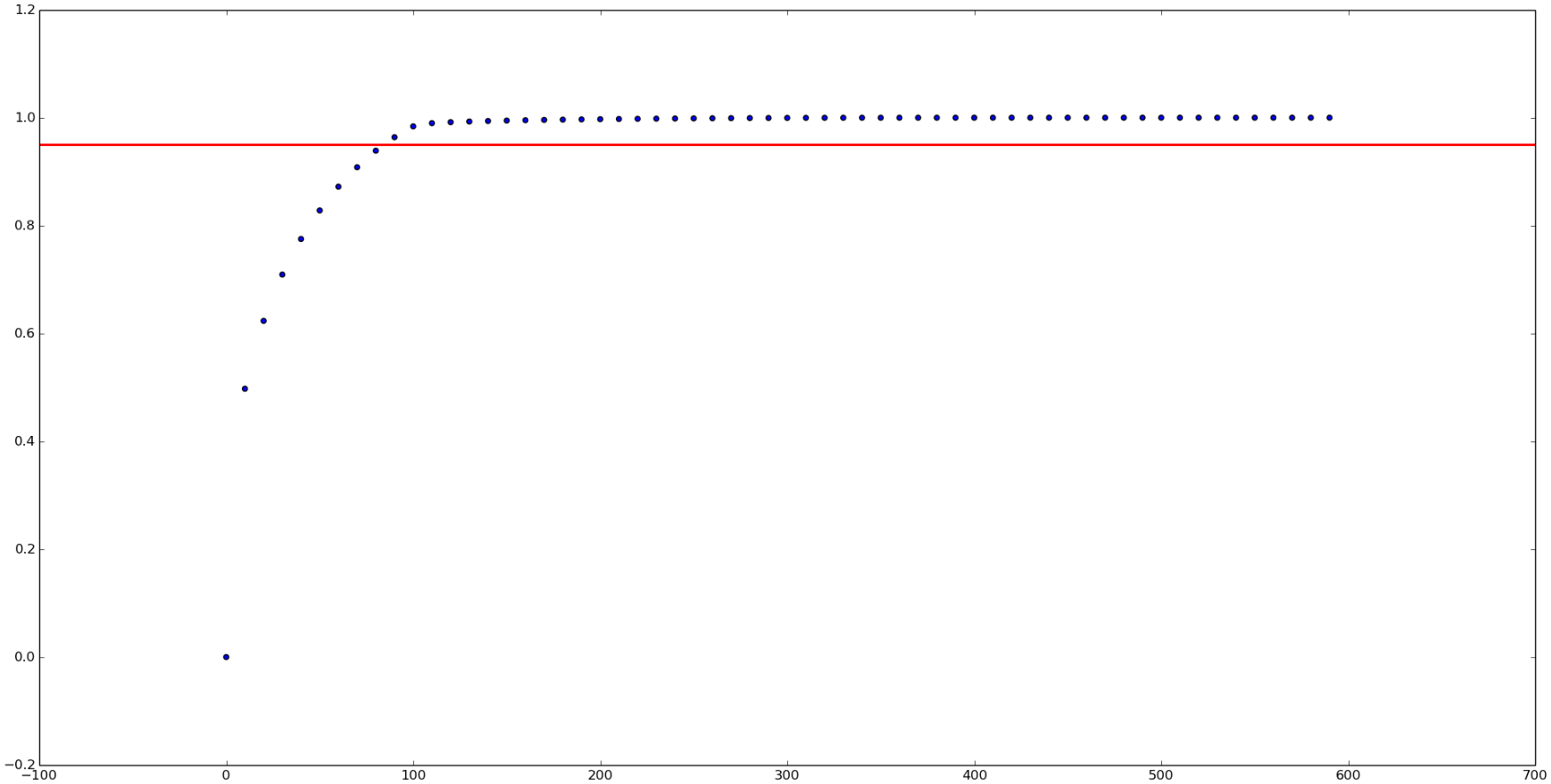
Learning

Results

Feature Selection

PCA or LDA?

Principal Component Analysis (PCA)

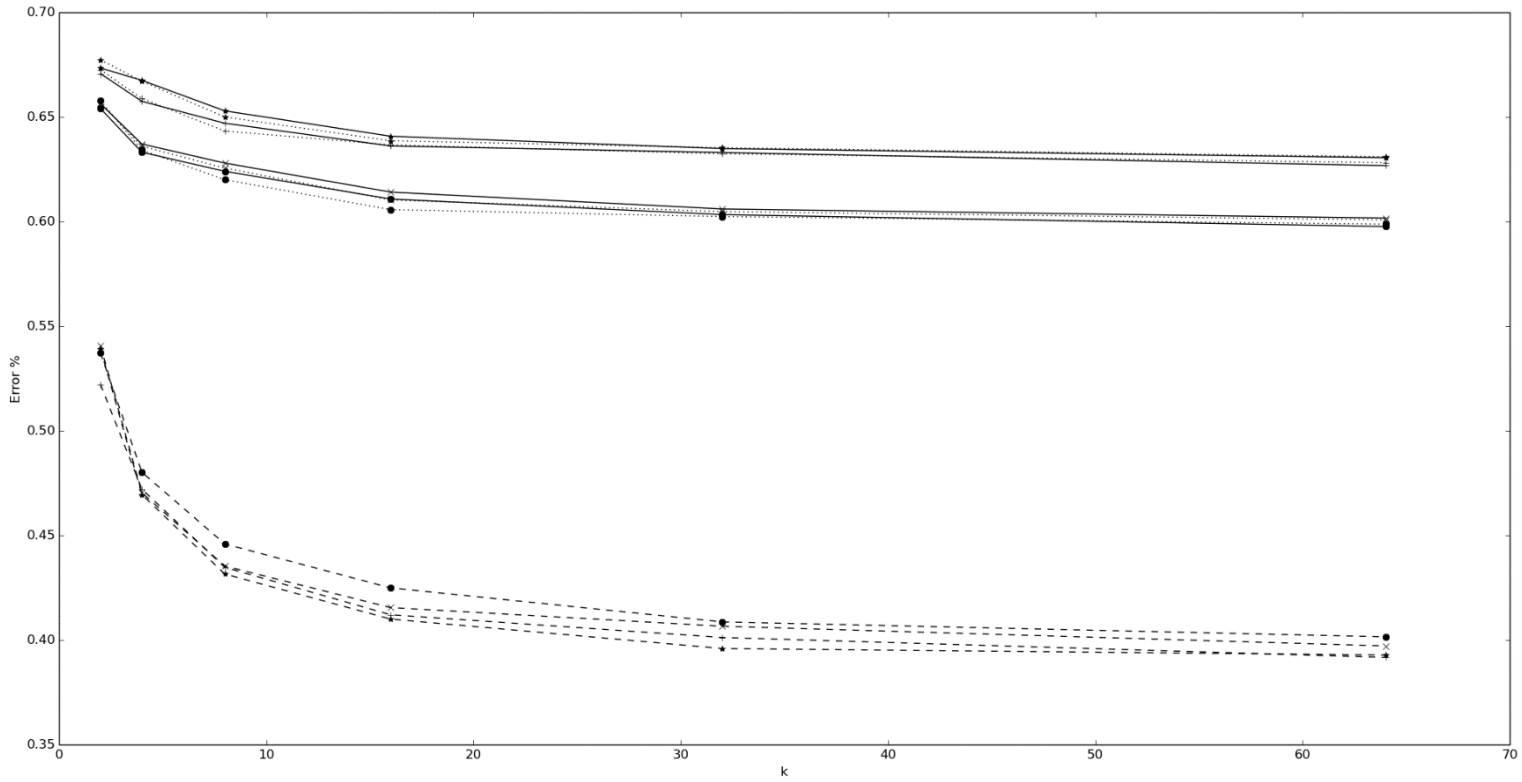


Linear Discriminant Analysis (LDA)

$$d = 3$$

(that's it)

Error rate with k-NN



Pipeline

Feature Extraction

Feature Selection

Learning

Results

Learning

Learning Vector Quantization (LVQ)

k-NN

SVM

Mixture of Gaussians

Hierarchical Discriminating Regression (HDR) tree

Learning

Learning Vector Quantization (LVQ)

k-NN

SVM

Mixture of Gaussians

~~Hierarchical Discriminating Regression (HDR) tree~~

AdaBoost

Learning

Learning Vector Quantization (LVQ)

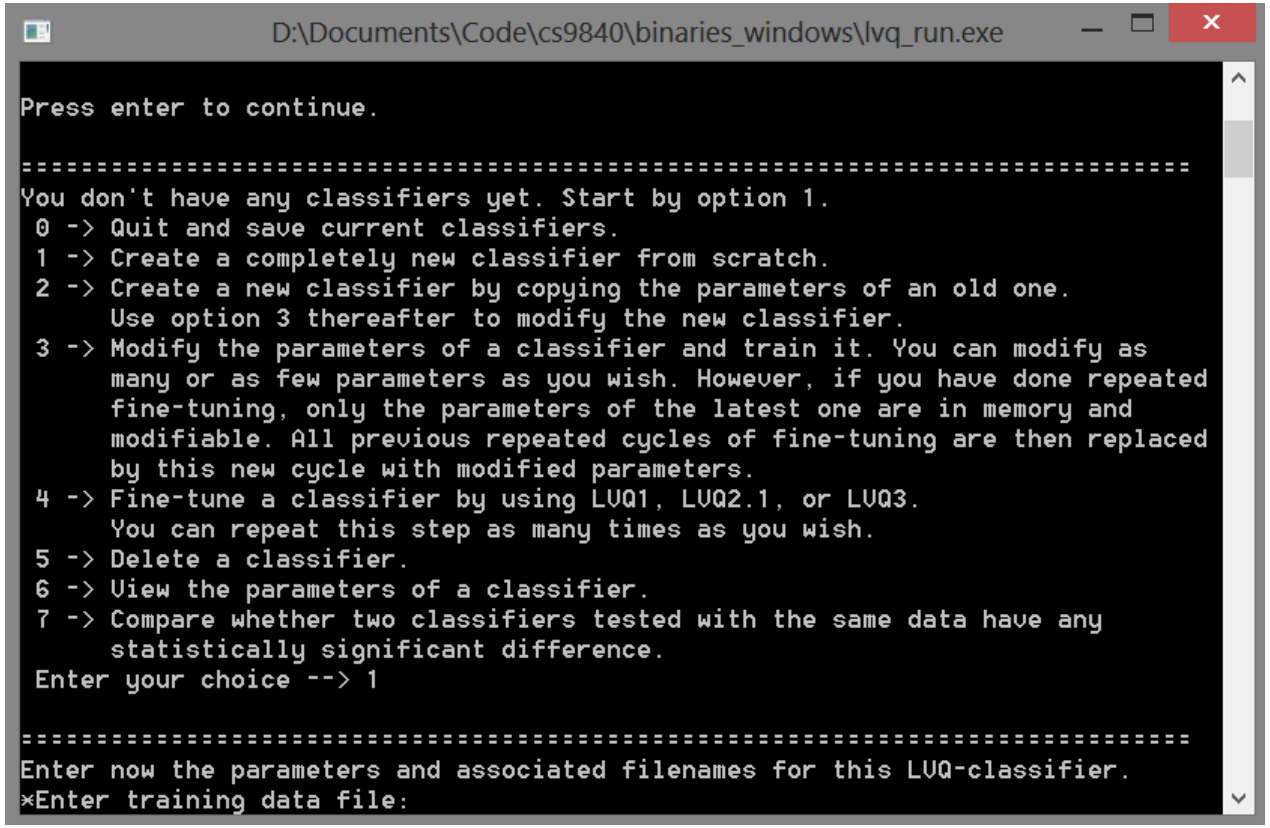
k-NN

SVM

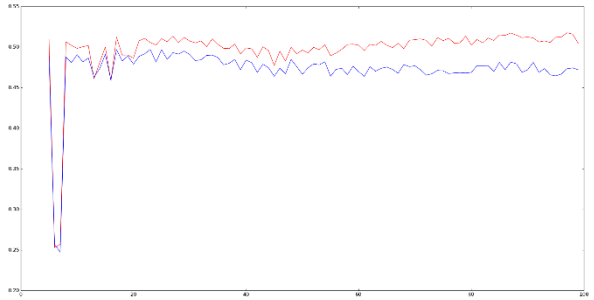
Mixture of Gaussians

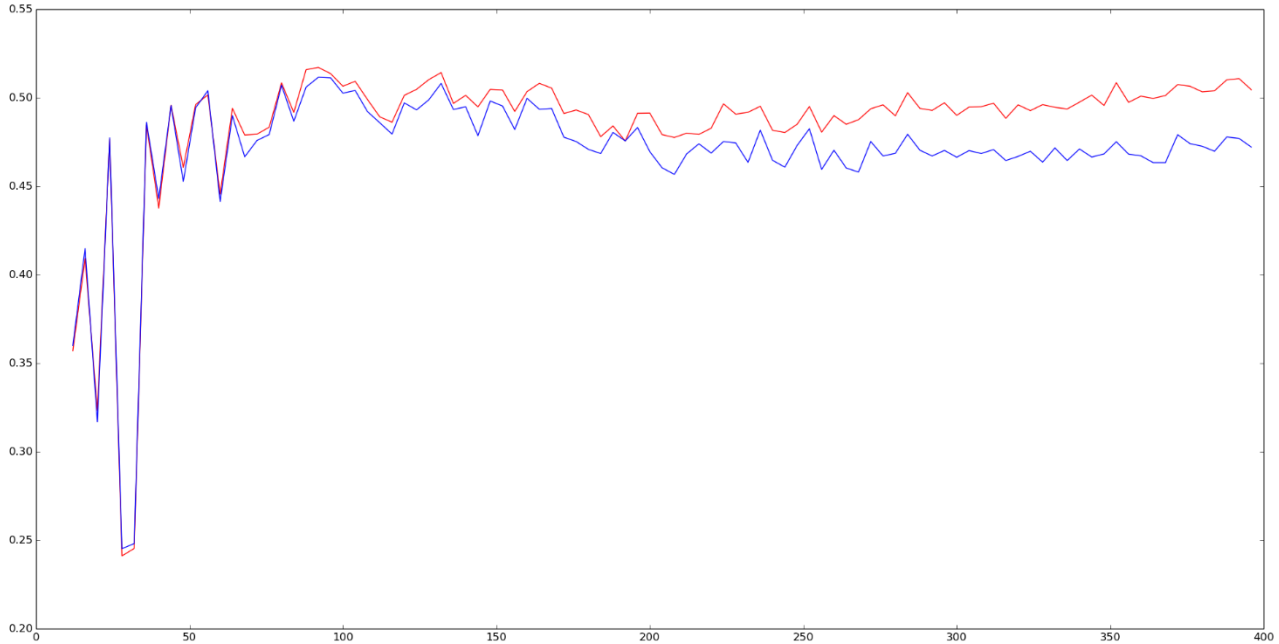
AdaBoost

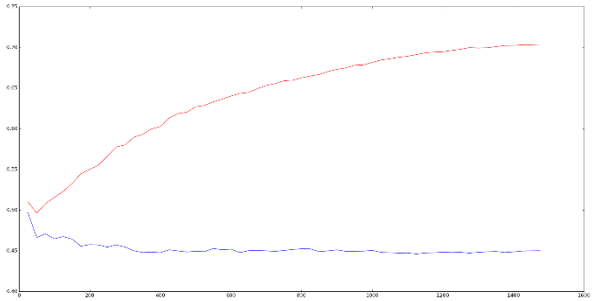
Learning Vector Quantization



Learning Vector Quantization







Learning

Learning Vector Quantization (LVQ)

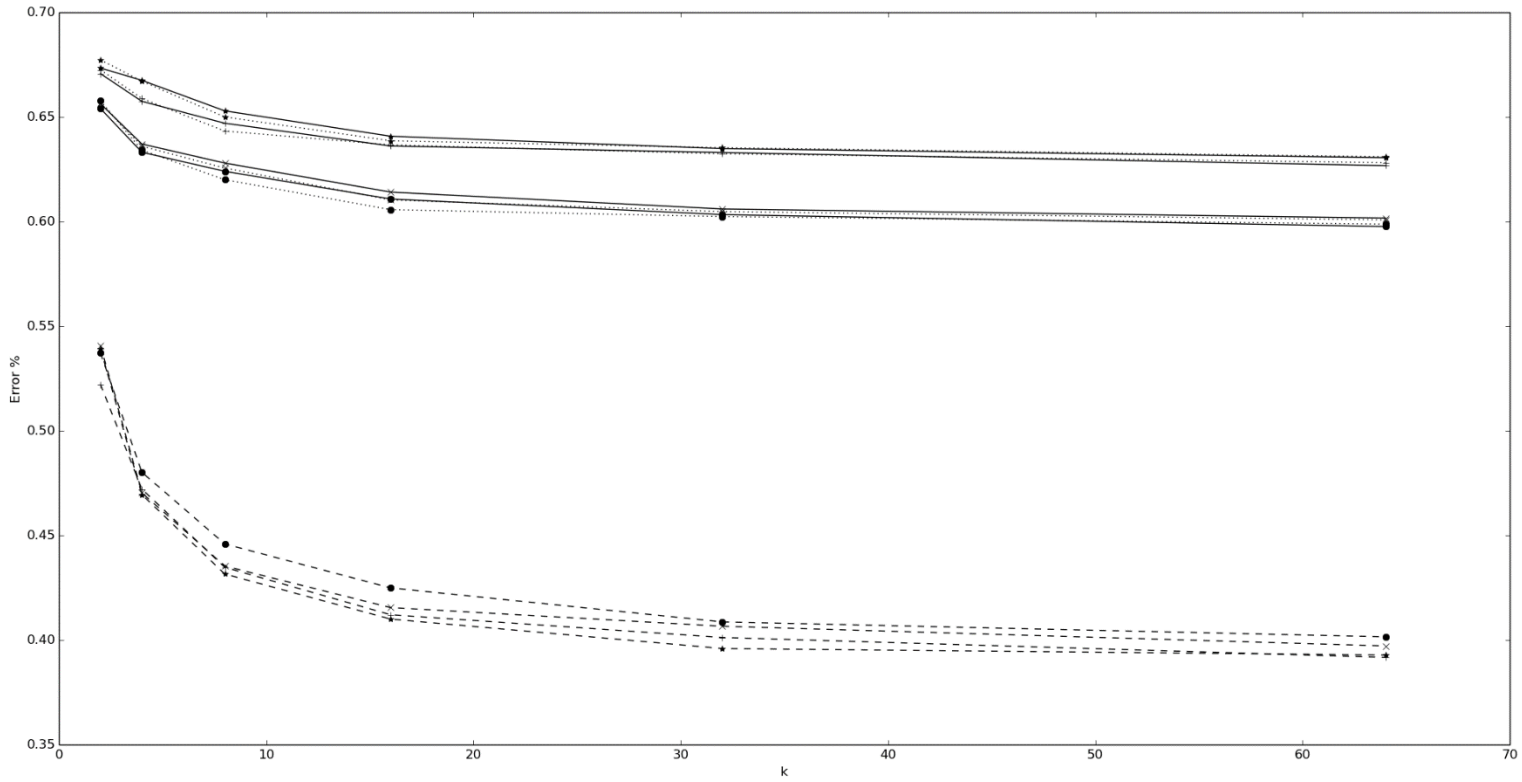
k-NN

SVM

Mixture of Gaussians

AdaBoost

k Nearest Neighbors (k-NN) Classifier



k Nearest Neighbors (k-NN) Classifier

	without LDA		with LDA	
	Training	Testing	Training	Testing
k-NN $(k = 1)$	100%	33.6%	100%	45.0%
k-NN $(k = 5)$	57.4%	35.7%	67.7%	49.2%
k-NN $(k = 10)$	51.7%	37.9%	63.7%	52.8%
k-NN $(k = 50)$	44.5%	40.0%	61.5%	55.8%
k-NN $(k = 100)$	43.3%	39.8%	62.1%	55.0%
k-NN $(k = 500)$	41.4%	38.8%	61.1%	55.2%

Learning

Learning Vector Quantization (LVQ)

k-NN

SVM

Mixture of Gaussians

AdaBoost

Support Vector Machine (SVM)

	VOC2007	IndoorScenes	VOC2012	"Easy"
lin	40.5%	55.5%	45.4%	65.6%
poly ($d = 2$)	36.3%	40.5%	37.2%	44.0%
poly ($d = 3$)	39.8%	53.9%	43.8%	61.6%
poly ($d = 4$)	33.4%	37.0%	34.4%	38.4%
poly ($d = 5$)	38.9%	47.3%	38.5%	58.4%
poly ($d = 6$)	31.4%	36.9%	31.1%	43.2%
rbf ($\gamma = 2^0$)	40.5%	55.6%	46.1%	24.0%
rbf ($\gamma = 2^{-1}$)	40.4%	55.7%	45.9%	24.8%
rbf ($\gamma = 2^{-2}$)	40.9%	55.6%	46.0%	25.6%
rbf ($\gamma = 2^{-3}$)	40.5%	55.4%	45.8%	32.0%
rbf ($\gamma = 0$)	41.1%	55.7%	46.0%	25.6%
sigm ($\gamma = 2^0$)	30.2%	36.3%	29.2%	36.8%
sigm ($\gamma = 2^{-1}$)	30.9%	36.5%	32.0%	41.6%
sigm ($\gamma = 2^{-2}$)	33.4%	39.2%	34.5%	41.6%
sigm ($\gamma = 2^{-3}$)	35.1%	43.4%	37.9%	41.6%
sigm ($\gamma = 0$)	25.8%	25.8%	25.3%	24.0%

Learning

Learning Vector Quantization (LVQ)

k-NN

SVM

Mixture of Gaussians

AdaBoost

AdaBoost

	without LDA		with LDA	
	Training	Testing	Training	Testing
AdaBoost (5)	40.9%	40.0%	58.1%	50.8%
AdaBoost (10)	48.0%	46.6%	59.1%	51.6%
AdaBoost (25)	52.3%	49.9%	60.7%	51.4%
AdaBoost (50)	54.3%	51.9%	60.8%	51.4%
AdaBoost (100)	56.7%	52.1%	61.3%	51.6%
AdaBoost (250)	59.9%	52.1%	61.6%	51.8%

Pipeline

Feature Extraction

Feature Selection

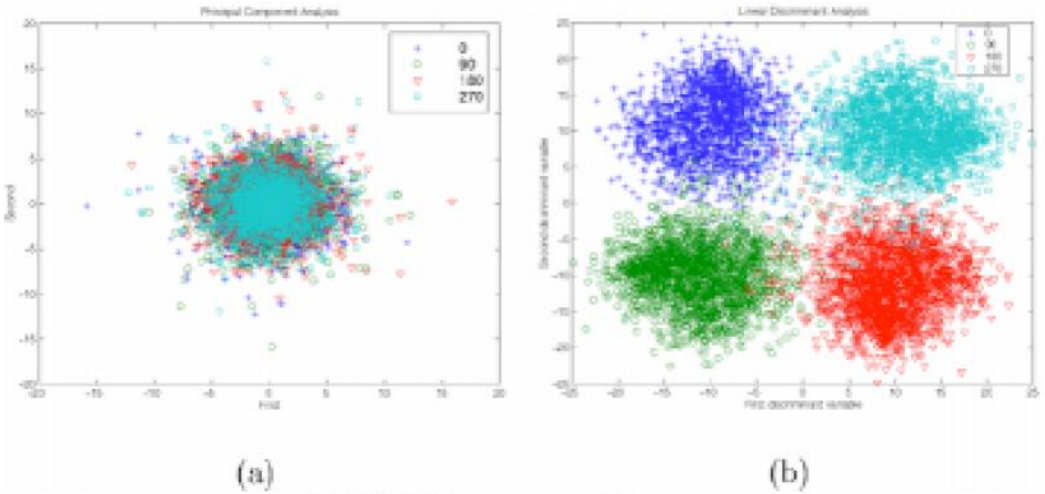
Learning

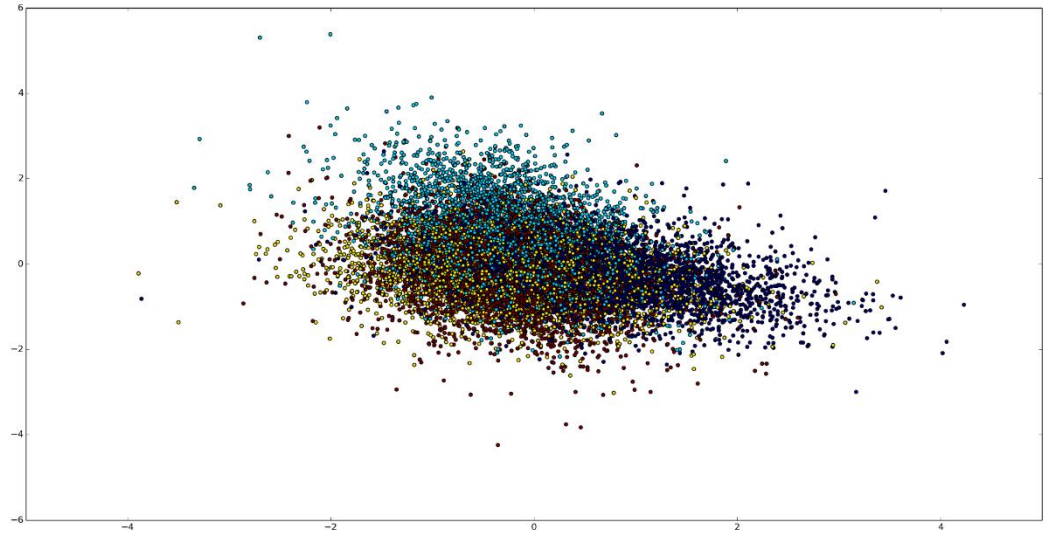
Results

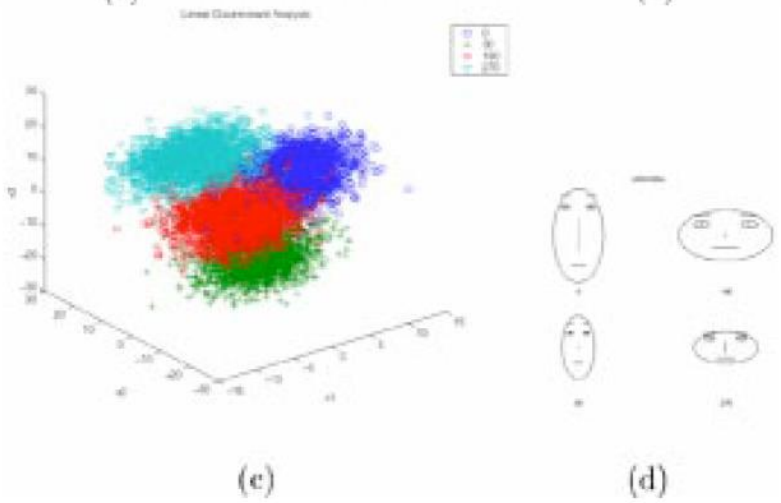
Results

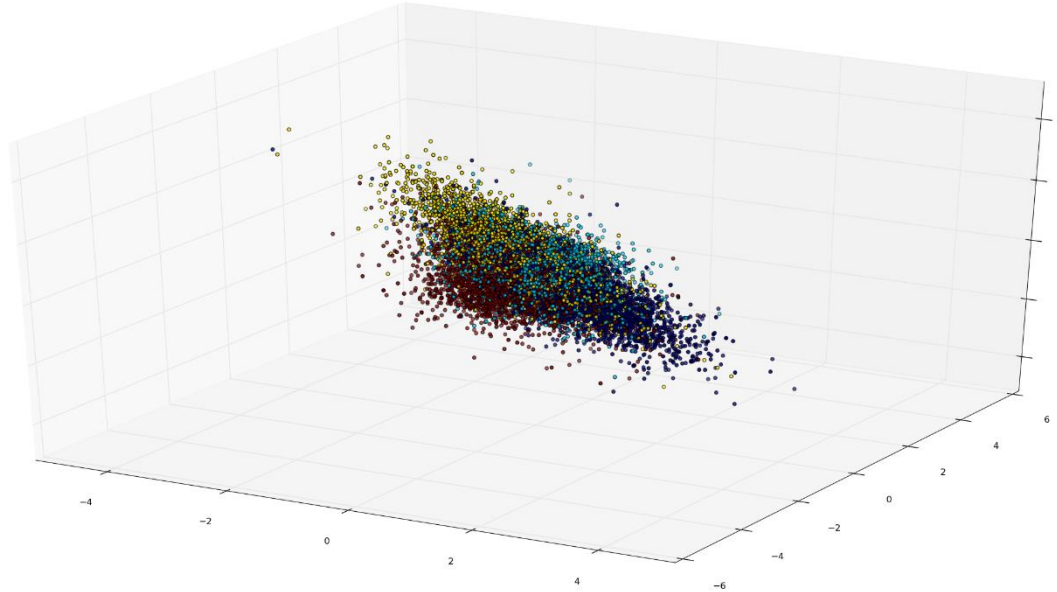
	without LDA		with LDA	
	Training	Testing	Training	Testing
k-NN ($k = 50$)	44.5%	40.0%	61.5%	55.8%
SVM (linear kernel)	100%	41.5%	61.3%	55.4%
LVQ ($q = 100$)	-	-	50.6%	50.2%
AdaBoost (50)	54.3%	51.9%	60.8%	51.4%
Mixture of Gaussian	-	-	49.6%	44.0%

Results



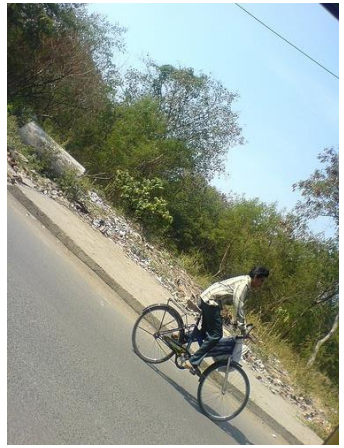






Some images are difficult…







…even for humans.







Improving Feature Selection

EDH

Optimizing N_{CM} and N_{EDH}

Removing center blocks does work!

Improving Classifiers

Semantic cues.

Specialized classifiers.

Thank you.