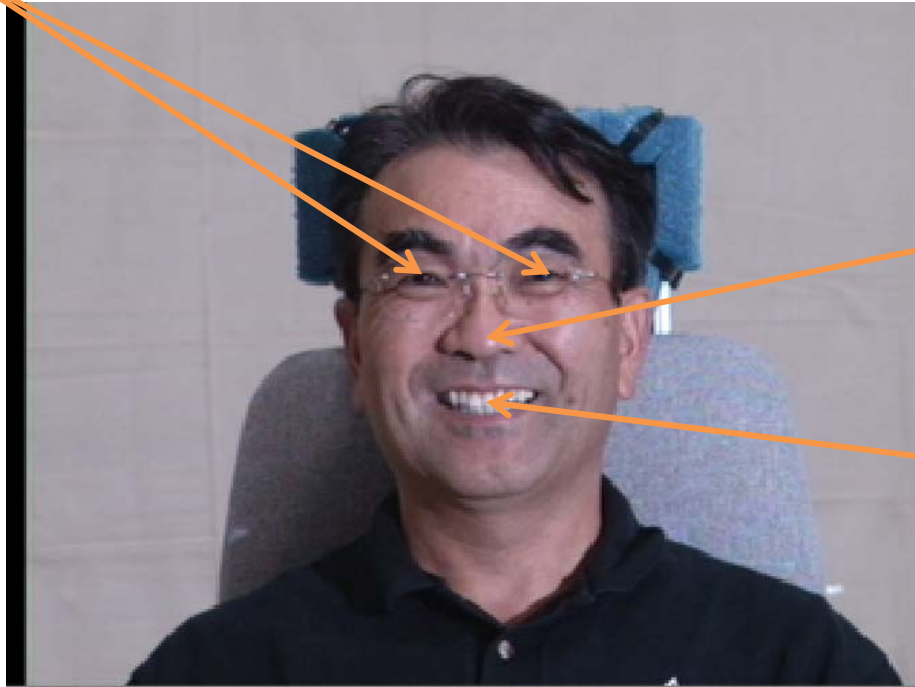


Facial Feature Detection with Optimal Pixel Reduction SVM

Minh Hoai Nguyen, Joan Perez, Fernando de la Torre
Robotics Institute, Carnegie Mellon University

Facial Feature Detection

Eyes



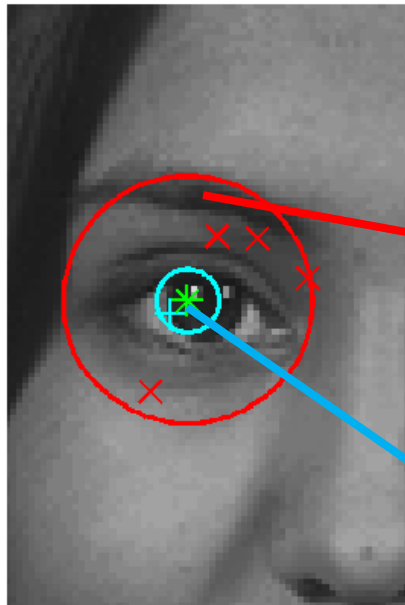
Nose

Mouth

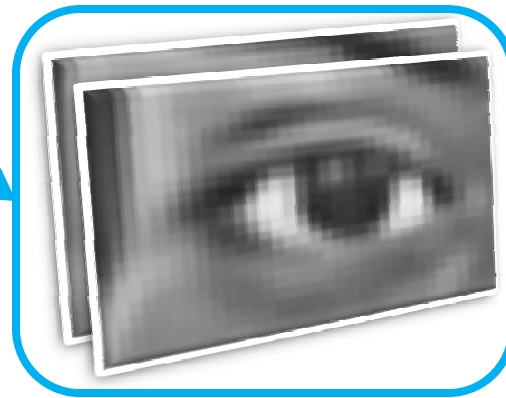
Facial feature detection – previous work

- **Representation:**
 - Intensity
 - Gradient [Kothari & Mitchell 96]
 - Wavelets [Huang & Wechsler 99]
 - Gabor filters [Fasel et al 02]
 - Active infrared [Zhu & Zhi 05]
- **Classifiers:**
 - Neural Network [Reinders et al 96, Huang & Wechsler 99]
 - Boosting [Cristiancce & Cootes 03, Everingham & Zisserman 06, Ma et al 04, Wang et al 05, Niu et al 06]
 - Bayesian/Regression [Zhou & Jiang 04, Everingham & Zisserman 06]
 - SVM [Jee et al 04, Tang et al 05, Jin et al 06, Campadelli et al 06]
- **Multiple facial features**
 - Active shape models [Cootes et al 98, Cristiancce & Cootes 03-06]
 - Pair of eyes [Ma et al 04, Jee et al 04, Tang et al 05, Jin et al 06]

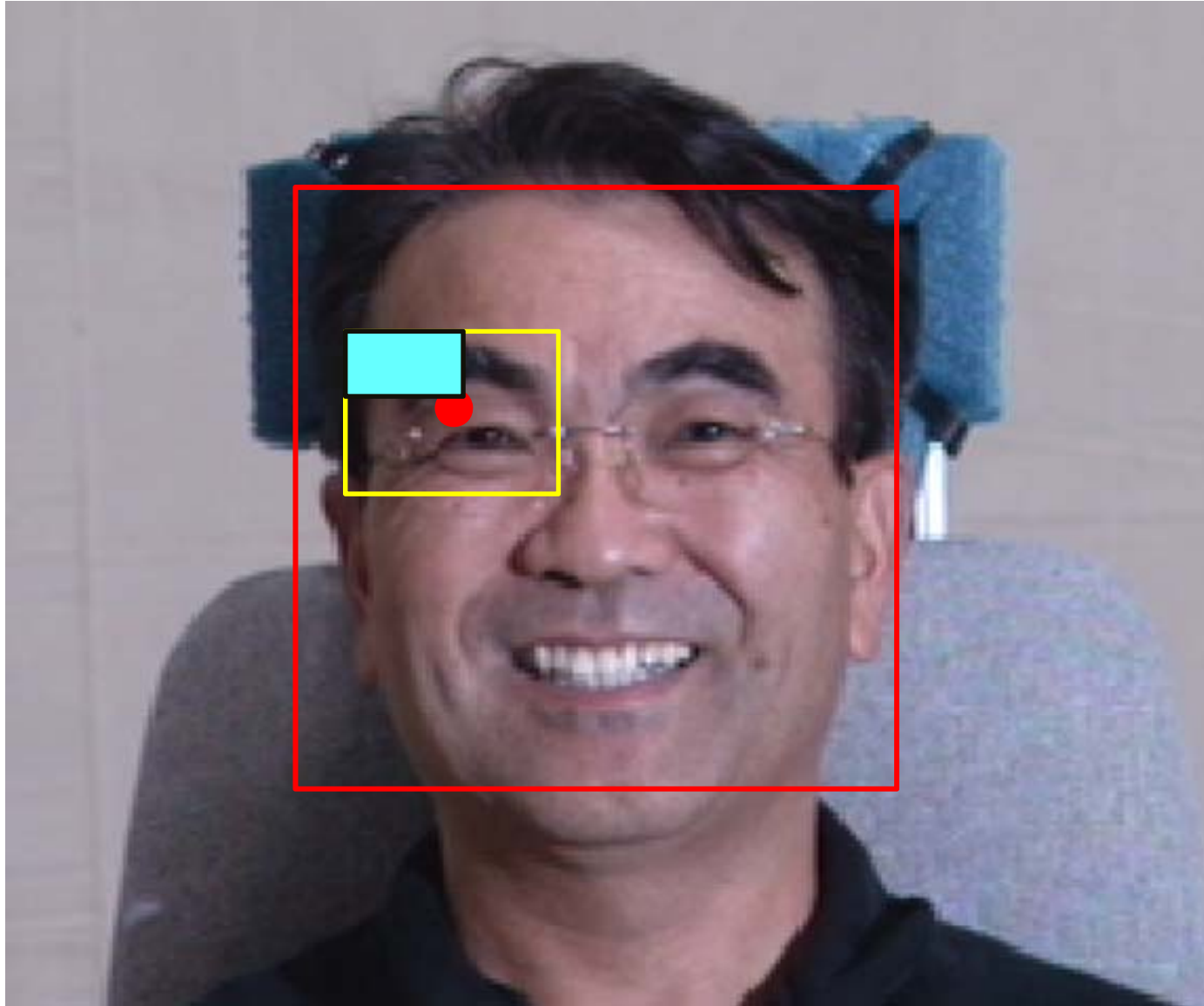
Training an SVM for eye detection



1. Get ground truth
2. Define Positive/Negative Areas
3. Randomly Select Samples
4. Extract Patches

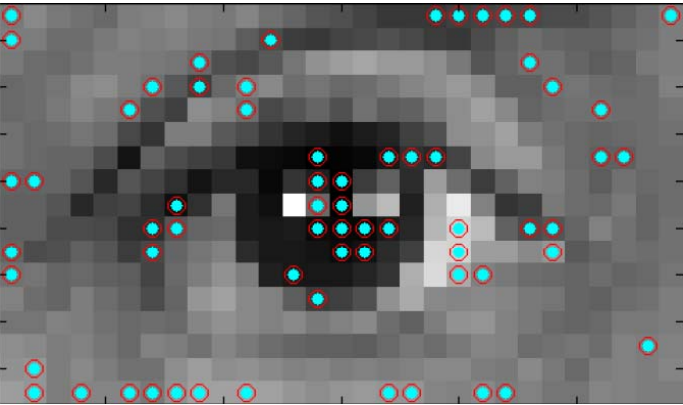
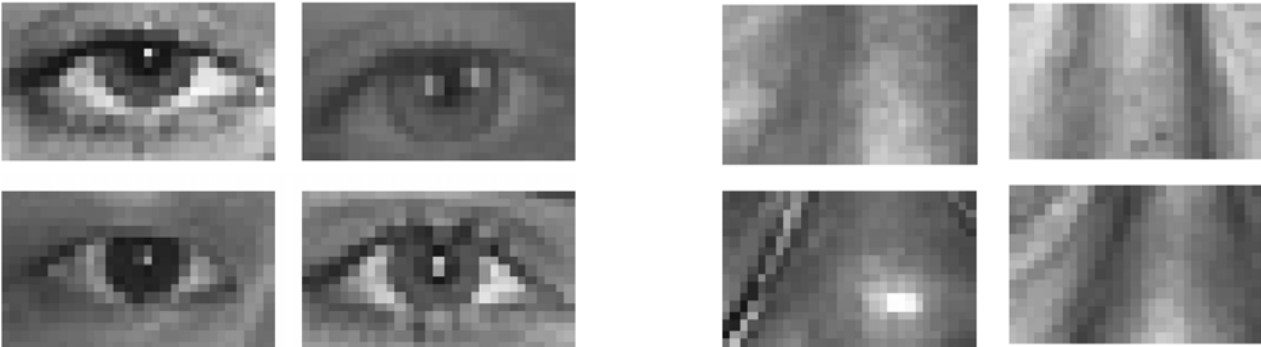


Eye detection with SVM

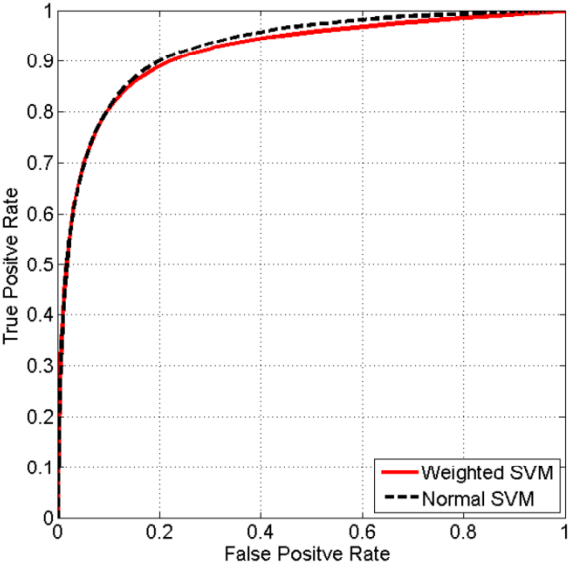


Increasing speed while maintaining accuracy

- Reduce #pixels
- Reduce #SVs

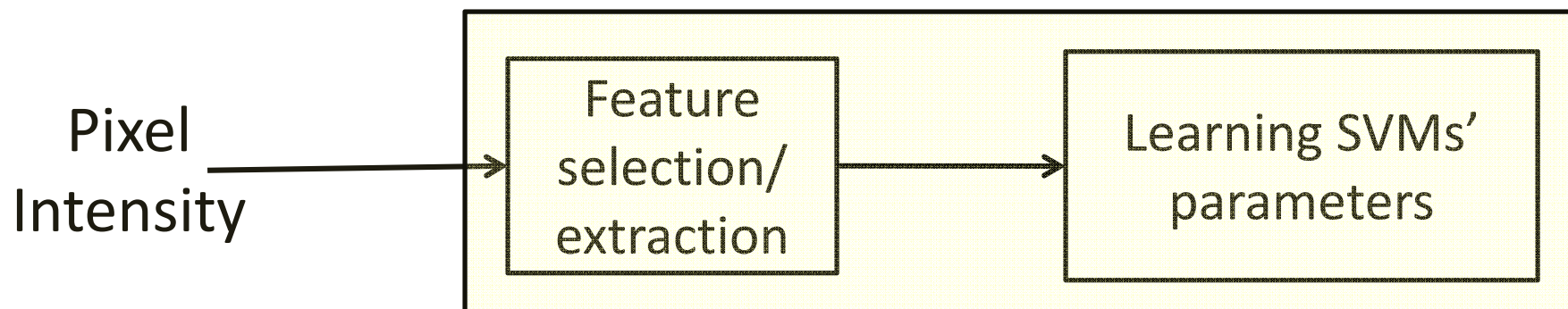


64 pixels (13%) used

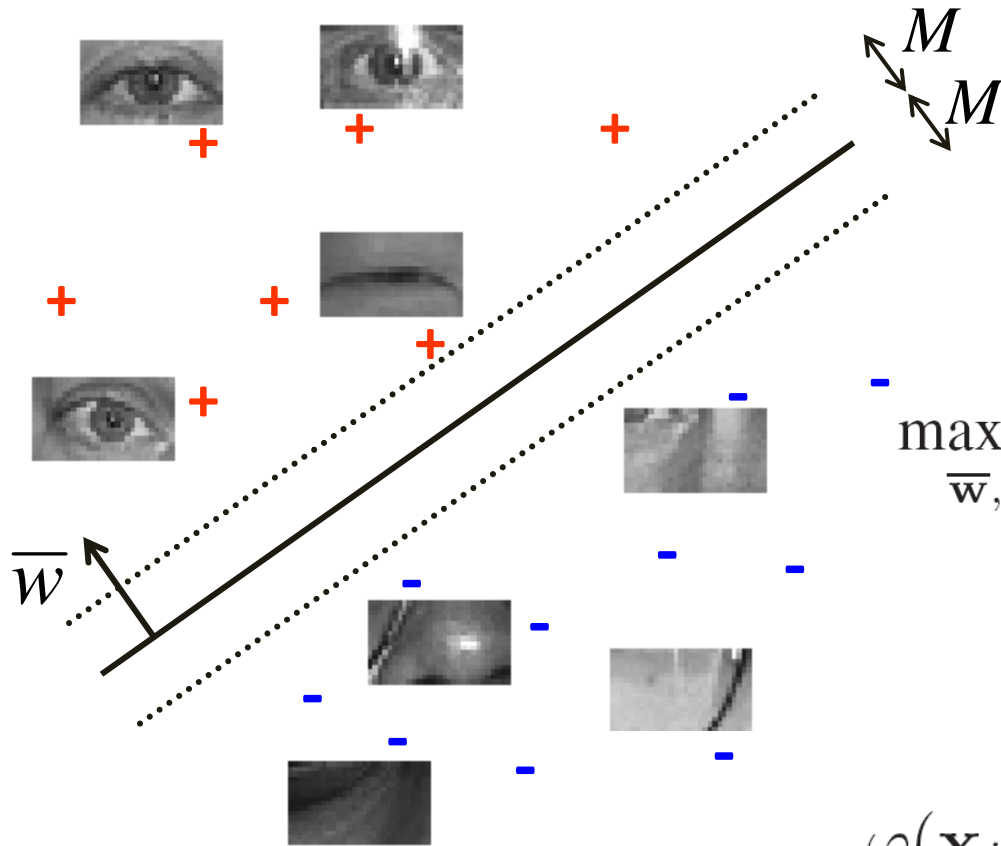


Feature construction in SVM

- Using LOOCV bounds [Weston et al 01, Chapelle et al 02]
- Learning kernel matrix [Cristianini et al 01 , Lanckriet et al 04, Hoi et al 06, de la Torre & Vinyals 07]
- Greedy Selection [Hermes & Buhmann 00, Avidan 04, Mangasarian & Wild 07]
- RELIEF [Kira & Rendell 92, Sun & Li 06]
- Adding Regularization [Stoeckel & Fung 05, Dundar et al 06, Chan et al 07]
- Weighting features in kernel space [Cao et al 07]



SVM & parameterized kernel



$$\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^{d \times 1}$$

$$y_1, \dots, y_n \in \{-1, 1\}$$

$$\text{maximize } M$$

$$\bar{\mathbf{w}}, \bar{b}, M$$

$$\text{s.t. } y_i(\bar{\mathbf{w}}^T \varphi(\mathbf{x}_i) + \bar{b}) \geq M \quad \forall i$$

$$\|\bar{\mathbf{w}}\|_2 = 1.$$

Mapping to
kernel space

$$\varphi(\mathbf{x}_i) = \varphi(\mathbf{x}_i, \mathbf{p})$$

$$\text{maximize } M$$

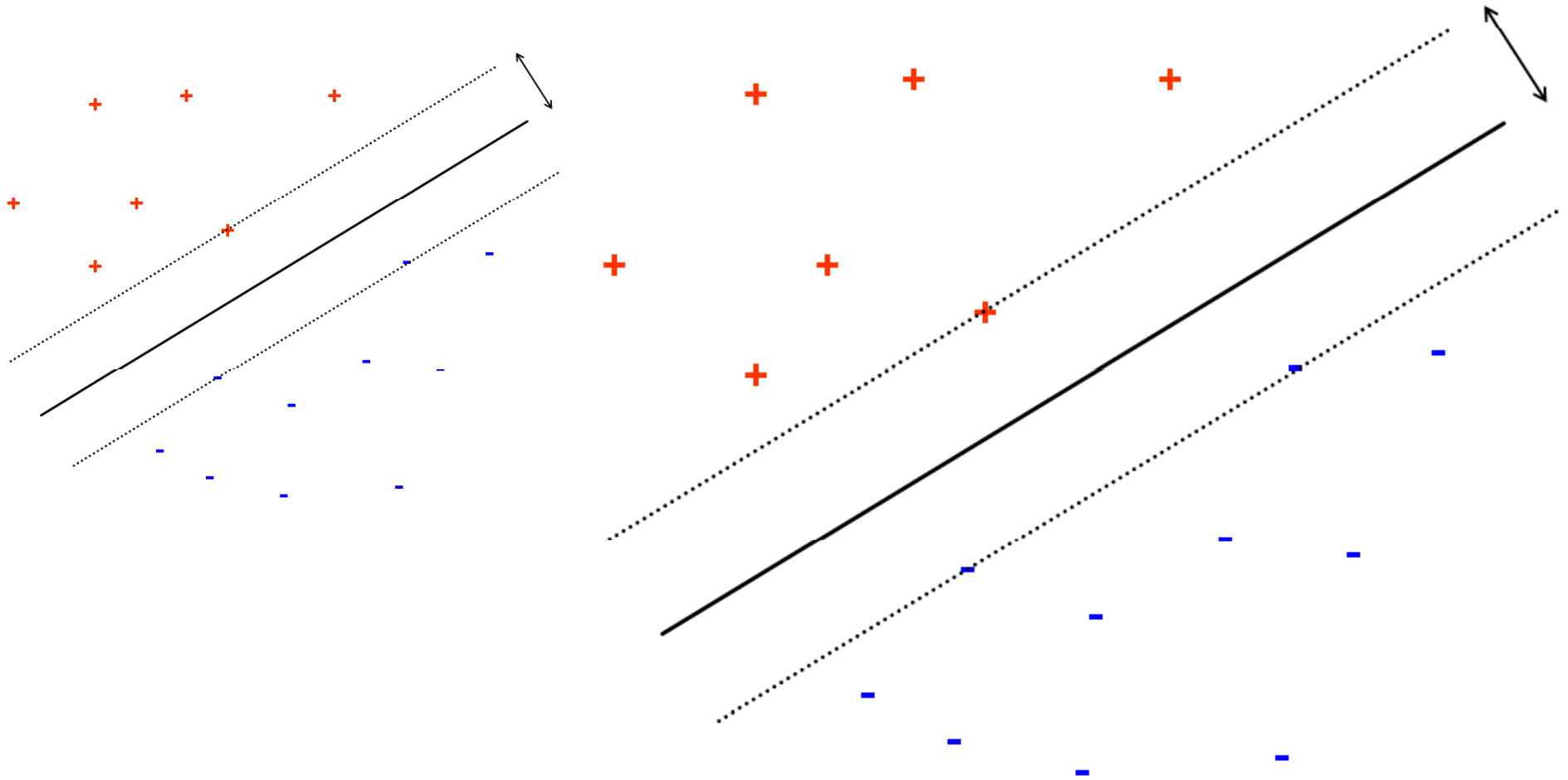
$$\bar{\mathbf{w}}, \bar{b}, M, \mathbf{p}$$

$$\text{s.t. } y_i(\bar{\mathbf{w}}^T \varphi(\mathbf{x}_i, \mathbf{p}) + \bar{b}) \geq M \quad \forall i$$

$$\|\bar{\mathbf{w}}\|_2 = 1.$$



Comparison of two feature spaces



$$\varphi(\mathbf{x}_i, \mathbf{p}_1) = \frac{1}{2} \times \varphi(\mathbf{x}_i, \mathbf{p}_2)$$

Normalized Margin

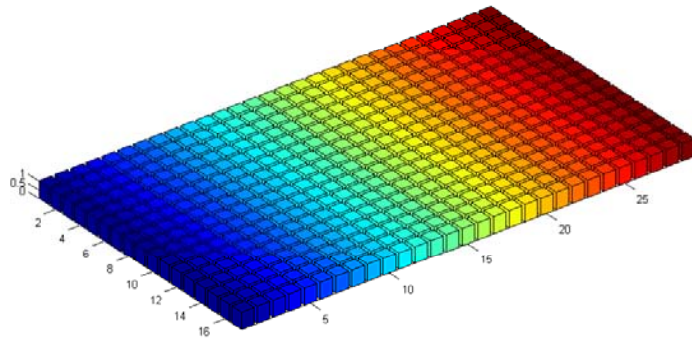
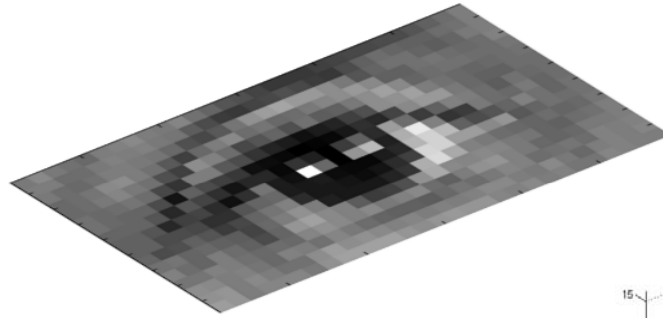
$$\begin{aligned}
 & \underset{\bar{\mathbf{w}}, \bar{b}, M, \mathbf{p}}{\text{maximize}} && M \\
 & \text{s.t.} && y_i(\bar{\mathbf{w}}^T \varphi(\mathbf{x}_i, \mathbf{p}) + \bar{b}) \geq M \quad \forall i \\
 & && \|\bar{\mathbf{w}}\|_2 = 1.
 \end{aligned}$$

Scale, translation, rotation invariant!

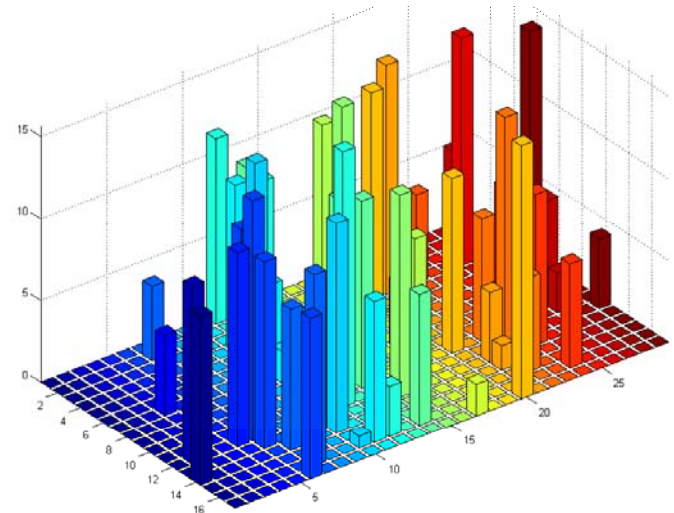
Let $\mathbf{w} = \bar{\mathbf{w}}/M$, $b = \bar{b}/M$

$$\begin{aligned}
 & \underset{\mathbf{w}, b, \mathbf{p}, \xi}{\text{minimize}} && \frac{1}{2} \phi(\mathbf{p}) \|\mathbf{w}\|_2^2 + C \sum_i \xi_i \\
 & \text{s.t.} && y_i(\mathbf{w}^T \varphi(\mathbf{x}_i, \mathbf{p}) + b) \geq 1 - \xi_i \quad \forall i \\
 & && \xi_i \geq 0 \quad \forall i
 \end{aligned}$$

Weighting the pixels



Uniform weights



Non-uniform weights

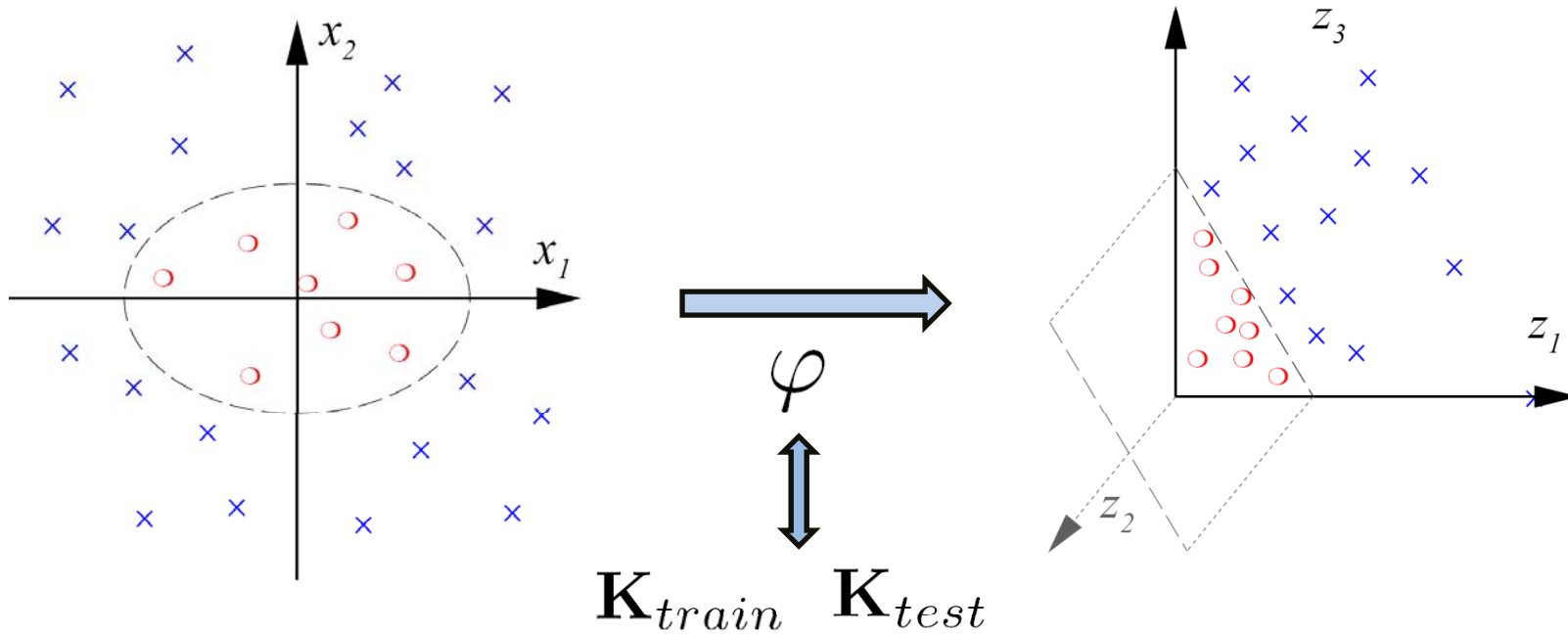
$$\varphi(\mathbf{p}) = \text{diag}(\mathbf{p})^{1/2}$$
$$p_i \geq 0 \quad \forall i$$

Learning feature weights

$$\begin{aligned} & \underset{\mathbf{v}, b, \mathbf{p}, \boldsymbol{\xi}}{\text{minimize}} && \frac{1}{2} \sum_{k=1}^d \frac{v_k^2}{p_k} + C \sum_{i=1}^n \xi_i \\ & \text{s.t.} && y_i (\mathbf{v}^T \mathbf{x}_i + b) \geq 1 - \xi_i \quad \forall i \\ & && \sum_{k=1}^d p_k \sum_{i,j} \frac{1 + y_i y_j}{2} (x_{ik} - x_{jk})^2 = 1 \\ & && \xi_i \geq 0 \quad \forall i, p_k \geq 0 \quad \forall k. \end{aligned}$$

Convex!

Feature weighting in feature space



$$\tilde{\varphi} : \mathbb{R}^d \rightarrow \mathbb{R}^n, \tilde{\varphi}(\mathbf{x}) = \mathbf{K}_{train}^{-\frac{1}{2}} \varphi(\mathbf{X})^T \varphi(\mathbf{x})$$

$$\tilde{\mathbf{K}}_{train} = \mathbf{K}_{train} \quad \tilde{\mathbf{K}}_{test} = \mathbf{K}_{test}$$

$$\tilde{\varphi}(\mathbf{x}; \mathbf{p}) = \text{diag}(\mathbf{p})^{1/2} \varphi(\mathbf{x})$$

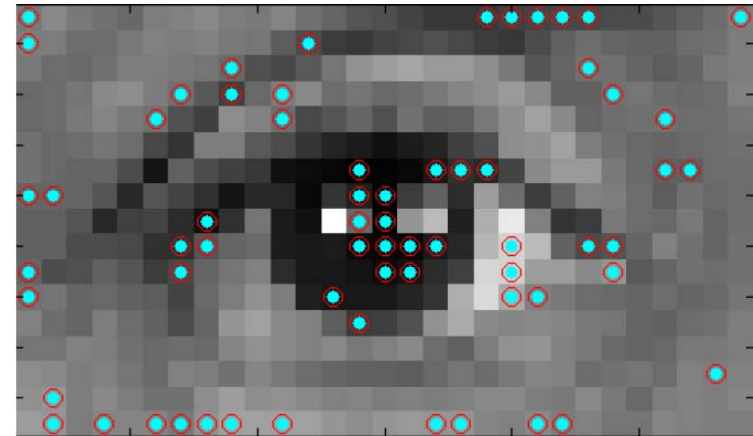
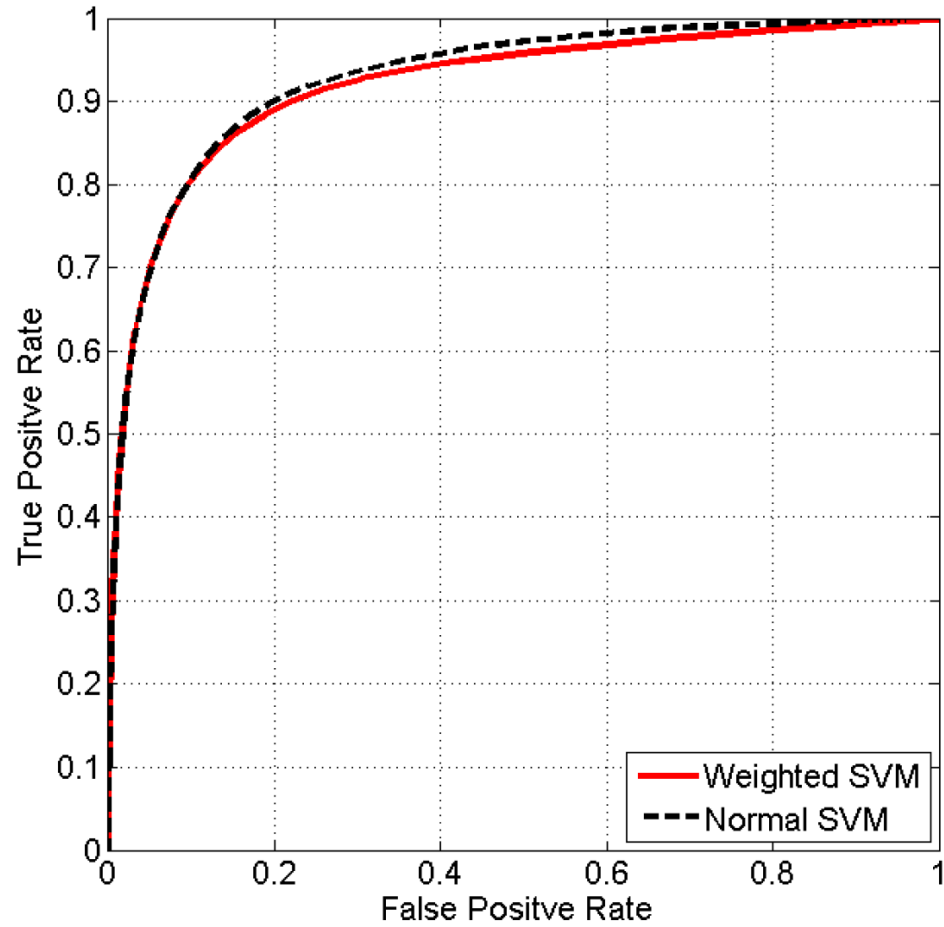
$$p_i \geq 0 \quad \forall i$$

Experiment 1 – Eye Detection

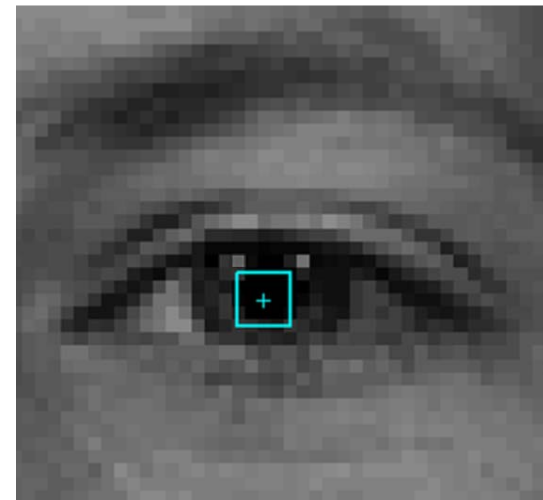
- **Task:**
 - Detect left eye position
- **Data:**
 - FERET
 - 2963 frontal faces
 - 4 hand-labeled landmarks
 - 256*384 pixel
 - Training: 60%, testing: 40%



ROCs



64 pixels (13%) used



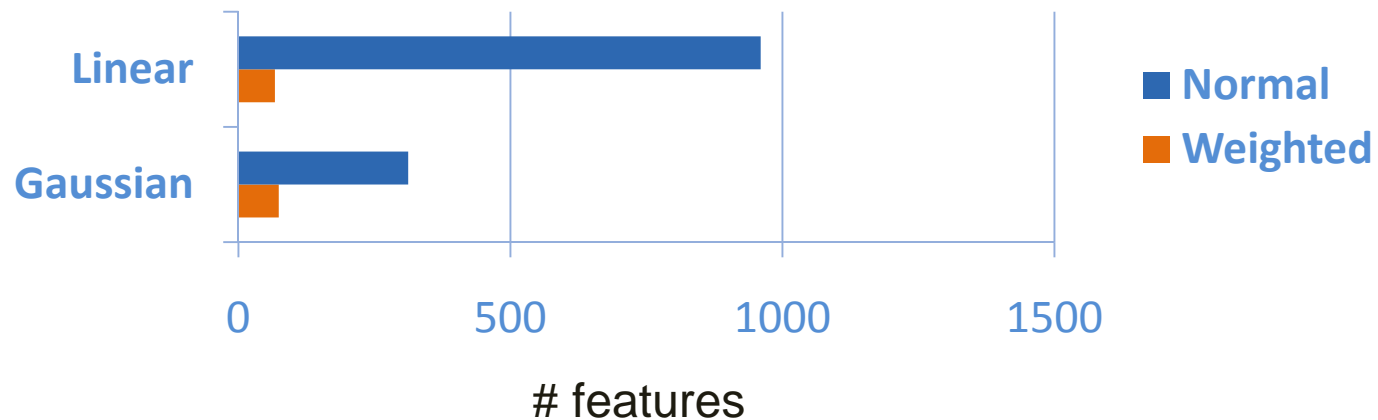
Experiment 2 - Pose classification

- **Task:** pose classification
- **Data:**
 - CMU face images from UCI ML repository
 - 312 faces, resolution: 30*32
 - 20 subjects, at various expressions and poses



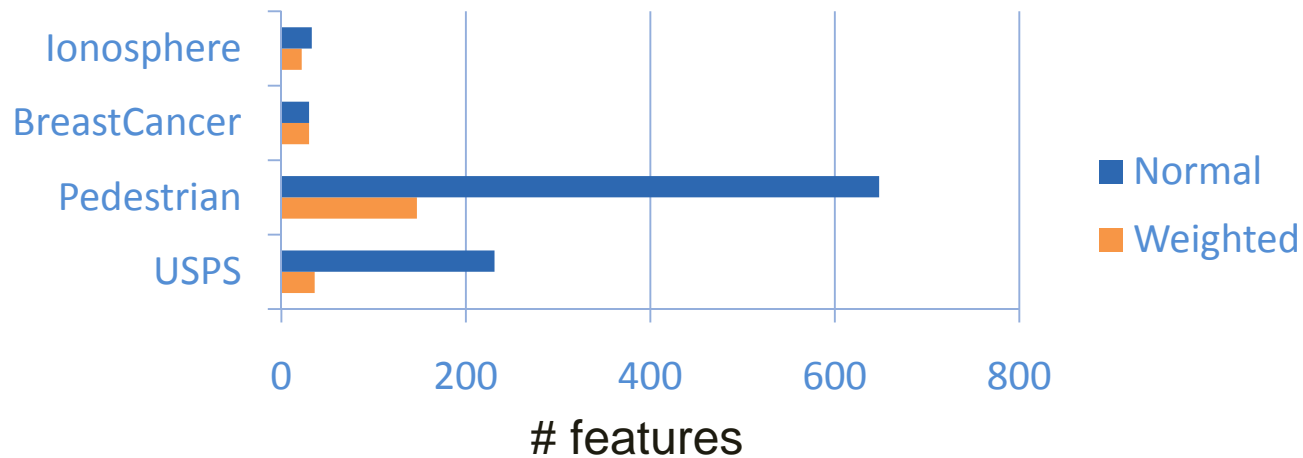
Pose classification

| Data | Method | Accuracy | | # features used | | # Support Vectors | |
|----------|--------------|----------|----------|-----------------|----------|-------------------|----------|
| | | Normal | Weighted | Normal | Weighted | Normal | Weighted |
| CMU Face | Linear SVM | 95.5% | 95.5% | 960 | 67 | 120 | 85 |
| CMU Face | Gaussian SVM | 97.48% | 98.06% | 312 | 74 | 186 | 73 |



Other standard datasets

| Data | Method | Accuracy | | # features used | | # SVs | |
|---------------|------------|----------|----------|-----------------|----------|--------|----------|
| | | Normal | Weighted | Normal | Weighted | Normal | Weighted |
| Ionosphere | Linear SVM | 86.89% | 88.50% | 33 | 22 | 79 | 65 |
| Breast Cancer | Linear SVM | 95.43% | 96.31% | 30 | 30 | 40 | 28 |
| Pedestrian | Linear SVM | 72.83% | 74.9% | 648 | 147 | 1291 | 1159 |
| USPS | Linear SVM | 97.75% | 97.5% | 231 | 36 | 50 | 40 |



Conclusions

- Learning a subset of pixels optimal for SVM
- Extended to the feature space
- Learning is a JOINT CONVEX optimization problem
- Maintain the classification performance using less support vectors and less features
- More complex features, but convexity will not be preserved