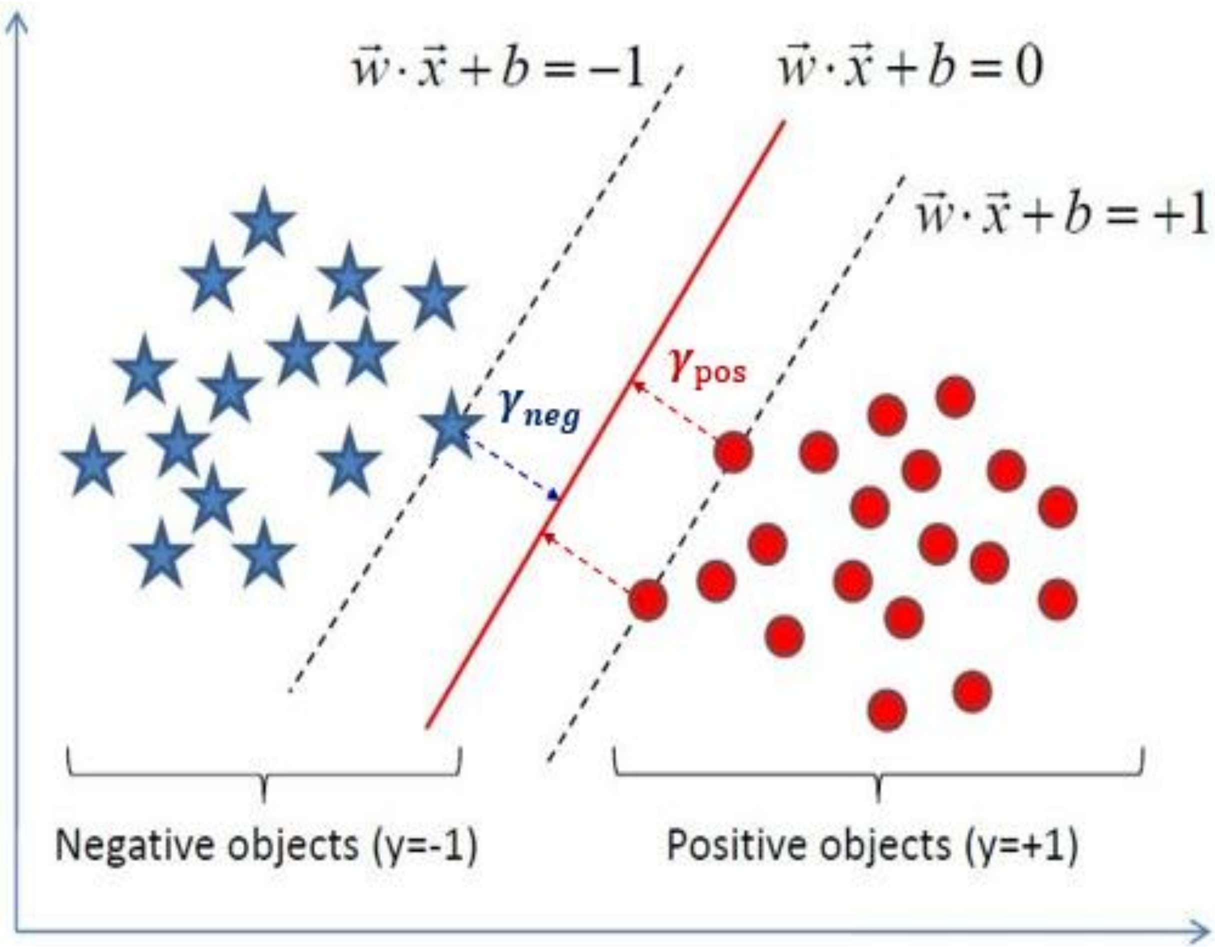
支持向量机应用

黄晟
huangsheng@cqu.edu.cn
办公室：信息大楼B701

A. SVM模型回顾

SVM的二分类问题



最大化分类间隔:

$$\max_{\gamma,w,b} \quad \gamma := \frac{2}{\|w\|}$$

$$\text{s.t.} \quad y_i(w^{\mathrm{T}}x_i + b) \geq 1$$

等价于

$$\min_{w,b} \quad \frac{1}{2}\|w\|^2$$

$$\text{s.t.} \quad y_i(w^{\mathrm{T}}x_i + b) \geq 1$$

$$\gamma = \gamma_{\text{pos}} + \gamma_{\text{neg}}$$

问题转化

- 拉格朗日乘子法：原问题➔对偶问题

$$\text{原问题 } \boldsymbol{p} = \min_{w,b} \max_{\lambda\geq 0} \mathcal{L}(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\lambda})$$

约束复杂不好解!

$$\text{对偶问题 } \boldsymbol{d} = \max_{\lambda\geq 0} \min_{w,b} \mathcal{L}(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\lambda})$$

约束简单容易解! (等式约束完全就是无约束问题)

$$\text{其中} \mathcal{L}(\boldsymbol{w}, \boldsymbol{b}, \boldsymbol{\lambda}) = \frac{1}{2}\|\boldsymbol{w}\|^2 + \sum_{i=1}^{m} \lambda_i \left(1 - y_i(\boldsymbol{w}^{\mathrm{T}} \boldsymbol{x}_i + \boldsymbol{b})\right)$$

原问题 $\neq$ 对偶问题，但当KKT条件成立
,原问题的解 $=$ 对偶问题的解

对偶问题

- **对偶问题求解(分两步解):** $\max\limits_{\boldsymbol{\lambda}\geq 0} \min\limits_{\boldsymbol{w},\boldsymbol{b}} \mathcal{L}(\boldsymbol{w},\boldsymbol{b},\boldsymbol{\lambda})$
  - 先求$\mathcal{L}(\boldsymbol{\lambda}) = \min\limits_{\boldsymbol{w},\boldsymbol{b}} \mathcal{L}(\boldsymbol{w},\boldsymbol{b},\boldsymbol{\lambda})$ **[1]**
  - 再求 $\max\limits_{\boldsymbol{\lambda}\geq 0} \mathcal{L}(\boldsymbol{\lambda})$ **[2]**
- 第一步—求解问题[1]: $\frac{\boldsymbol{\delta}\mathcal{L}(\boldsymbol{w},\boldsymbol{b},\boldsymbol{\lambda})}{\boldsymbol{\delta w}} = \mathbf{0}, \frac{\boldsymbol{\delta}\mathcal{L}(\boldsymbol{w},\boldsymbol{b},\boldsymbol{\lambda})}{\boldsymbol{\delta b}} = \mathbf{0}$
  - 得$\boldsymbol{w} = \sum_{i=1}^{m} \lambda_i y_i x_i$, $\sum_{i=1}^{m} \lambda_i y_i = \mathbf{0}$, 反代入$\mathcal{L}(\boldsymbol{w},\boldsymbol{b},\boldsymbol{\lambda})$
  - **可得带约束目标函数**$\mathcal{L}(\lambda)$，**s.t.** $\lambda \geq 0$与$\sum_{i=1}^{m} \lambda_i y_i = \mathbf{0}$

对偶问题

- **第二步—求解问题[2]：** $\max\limits_{\lambda\geq 0} \mathcal{L}(\boldsymbol{\lambda})$

$$\max_{\lambda} \quad \mathcal{L}(\boldsymbol{\lambda}) := \sum_{i=1}^{m} \lambda_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i \lambda_j y_i y_j \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \lambda_i y_i = 0 \qquad i = 1, \ldots, m$$

$$\lambda_i \geq 0 \qquad i = 1, \ldots, m$$

① 把约束$\sum_{i=1}^{m} \lambda_i y_i = 0$代入目标函数，化去$\boldsymbol{\lambda}_i$，然后对所有$\boldsymbol{\lambda}$求偏导赋零，$\frac{\delta\mathcal{L}}{\delta\lambda_{t\neq i}} = 0$，即可求解出所有$\boldsymbol{\lambda}_i$。

② 检查$\boldsymbol{\lambda}_i$约束情况，如不满足则讨论边界情况，给出最优$\boldsymbol{\lambda}_i$，从而算出w与b。

求解b

- 解出λ后，利用公式可以非常容易算出w.
- 求解b:
  - 对偶问题等价需满足KKT条件：
  - 互补(选择)条件：$\boldsymbol{\lambda_i(y_i(w^{\mathrm{T}}x_i+b)-1)=0}$
  - 即$\boldsymbol{\lambda_i \neq 0}$，则必有$\boldsymbol{y_i(w^{\mathrm{T}}x_i+b)-1=0}$
  - 则根据此关系可得：$\boldsymbol{b=y_i-w^{\mathrm{T}}x_i}$

SMO(Sequential Minimal Optimization)

- SMO引入的动机：
  - SVM本质任然是个多元规划问题，当λ_i数目非常多的时候，求解依然很耗时。
- SMO核心思想：
  - 多元规划问题才分成多个简单规划问题进行迭代求解。
- SMO算法流程：
  ① 启发地选择一对λ作为优化参数，其他λ设为固定。
  ② 求解此二元规划问题，解出λ，更新λ参数列表并计算w与b。
  ③ 检查所有λ是否满足KKT条件。
  ④ 如不满足，则跳转至步骤1，如满足，则输出w与b。

核化与软间隔SVM

- Standard：

$$\max_{\boldsymbol{\lambda}} \quad \mathcal{L}(\boldsymbol{\lambda}) \coloneqq \sum_{i=1}^{m} \lambda_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i \lambda_j y_i y_j \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \lambda_i y_i = 0, \quad \lambda_i \geq 0 \qquad i = 1, \dots, m$$

- Kernelized：

$$\max_{\boldsymbol{\lambda}} \quad \mathcal{L}(\boldsymbol{\lambda}) \coloneqq \sum_{i=1}^{m} \lambda_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i \lambda_j y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \lambda_i y_i = 0, \quad \lambda_i \geq 0 \qquad i = 1, \dots, m$$

- Soft margin:

$$\max_{\boldsymbol{\lambda}} \quad \mathcal{L}(\boldsymbol{\lambda}) \coloneqq \sum_{i=1}^{m} \lambda_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i \lambda_j y_i y_j \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \lambda_i y_i = 0, \quad C \geq \lambda_i \geq 0 \qquad i = 1, \dots, m$$

B. 基于SVM的行人检测

Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//*IEEE Computer Society Conference on Computer Vision and Pattern Recognition*（**CVPR**), 2005, 1: 886-893. (**Google citation:** 39301)

Histograms of **oriented gradients** for human detection CCF A [PDF] inria.fr

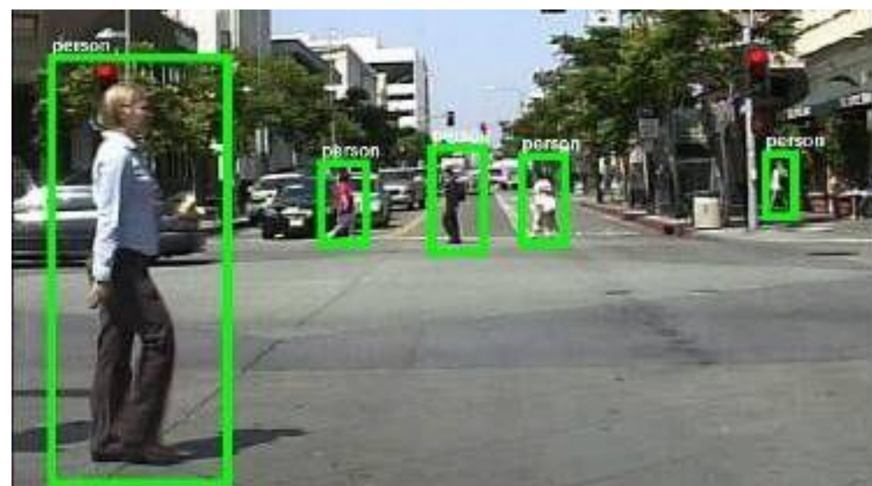
N Dalal, B Triggs - 2005 IEEE computer society conference on …, 2005 - ieeexplore.ieee.org

We study the question of feature sets for robust visual object recognition; adopting linear SVM based human detection as a test case. After reviewing existing edge and **gradient** …

☆ Save 99 Cite Cited by 39301 Related articles All 93 versions

行人检测

- 行人检测(human detection/pedestrian detection)
  - 从图片或视频中检测行人目标，典型的二分类问题。
  - 计算机视觉的基础问题，应用十分广泛，如视觉导航、视频监控、智能交通等领域。
  - 非常容易推广到其他检测问题，如人脸检测、车辆检测等。



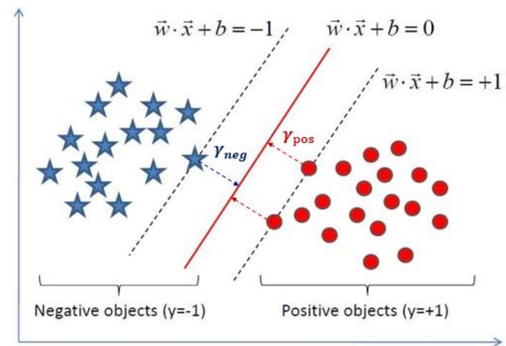
行人检测流程

- 基本流程：

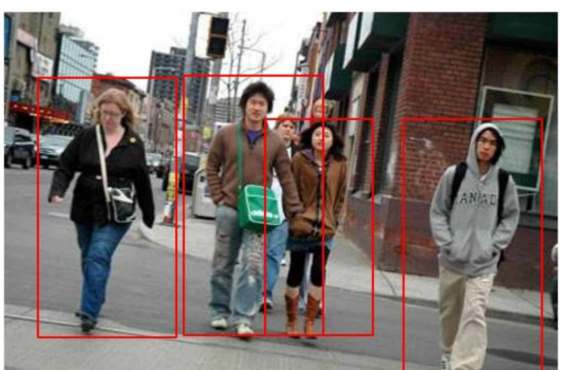
  输入图片➔选取候选窗口➔特征提取➔分类➔窗口融合➔输出结果

proposals

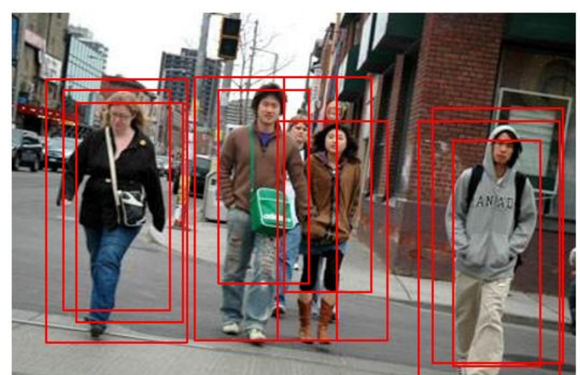
HOG features



SVM Classification



fusion



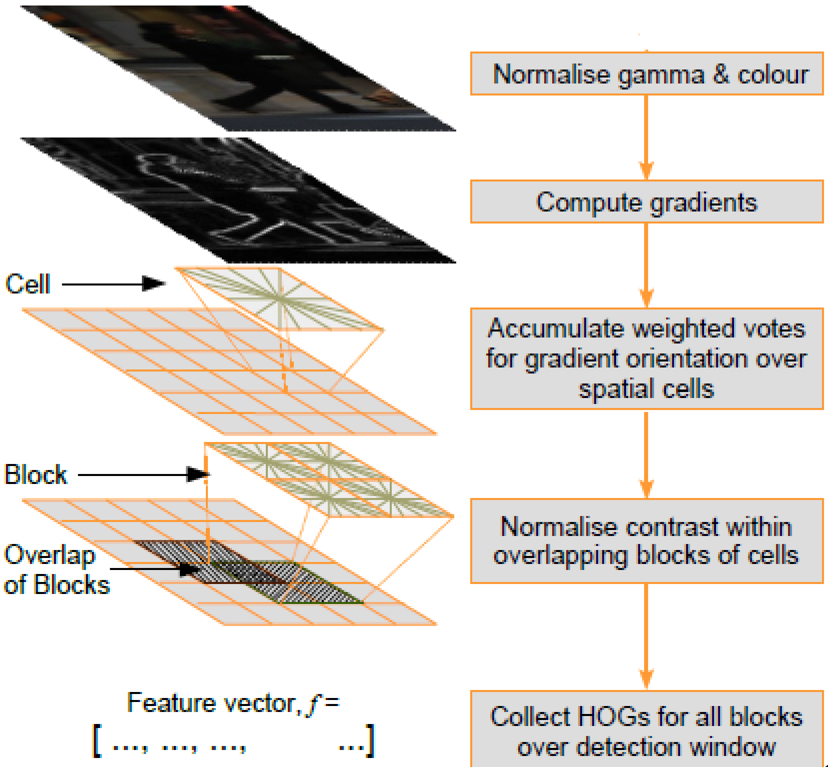
Sliding window

- 输入图片➔选取候选窗口

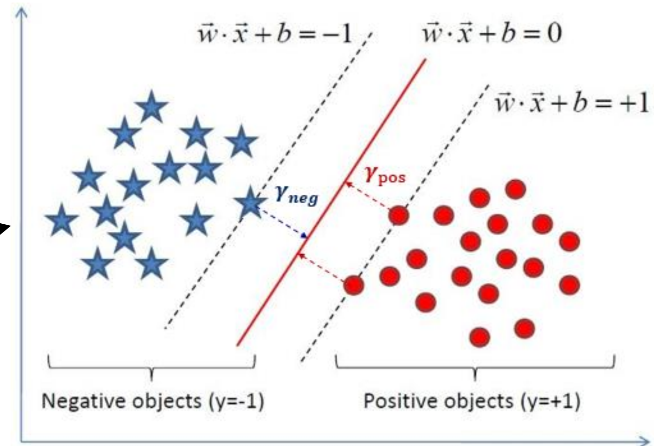
不同尺度图片下，利用滑动窗口(Sliding window)法, 裁剪出一系列128x64尺寸候选窗口(proposals)。

候选窗口集

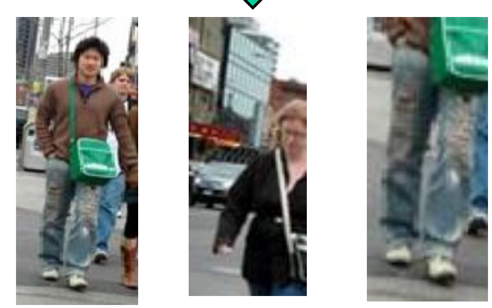
HOG+SVM



数据预处理与HOG特征提取



输出SVM Score



1.92 0.12 -0.88

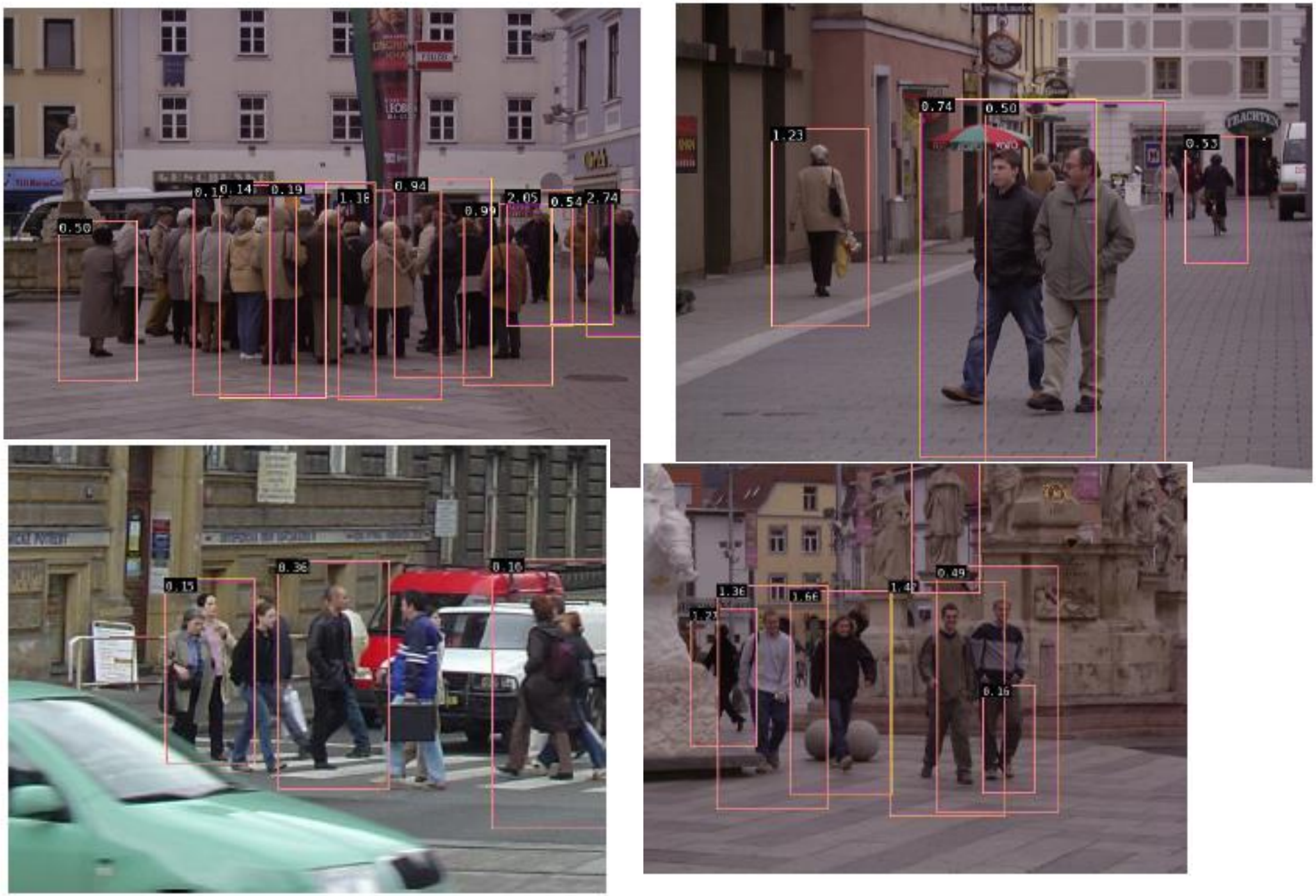
Detection window Fusion

- 窗口融合的两种思路
  - 非极大值抑制（Non-Maximum Suppression，NMS）
  - 聚类（Clustering）

(a) After multi-scale dense scan

(b) Fusion of multiple detections

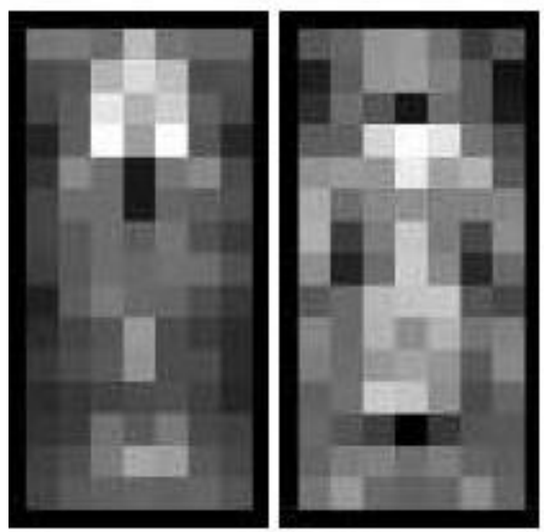
Results


0.50
0.14
0.19
1.16
0.94
0.99
2.05
0.54
2.74
1.23
0.74
0.50
0.53
0.15
0.36
0.16
1.36
1.66
1.47
0.49
0.16

思考

- 实际问题的解决≠数据+机器学习
  - 多种机器学习技术相互配合
- 数学不好就没法从事机器学习研究与开发？
  - HOG+SVM+NMS基本没有什么高深数学知识
- 支持向量机只是一个分类器？

局部最大正SVM权重



局部最小负SVM权重

练练手

- 求解P38页中的原始SVM优化问题。

  提示：matlab function: quadprog.

- 利用SMO算法求解例子中的问题。

  提示：SVM工具包---LIBSVM

- 在同一人工（或下载）数据集上比较 逻辑回归与支持向量机分类精度。
- 在SMO中，假设选定需要更新的变量为 λ_i 和 λ_j，试求 λ_i 和 λ_j 的更新公式。
- 尝试Github寻找HOG算法结合SVM实现静态图片中的行人检测问题。