

文章编号: 1671-7449(2013)01-0073-06

基于 Adaboost 算法选取和组合 SVM 的行人检测算法^{*}

张丽红, 李 林

(山西大学 物理电子工程学院, 山西 太原 030006)

摘 要: 行人检测是计算机视觉中一个重要的研究方向, 为了提高行人的识别精度, 将支持向量机 (Support Vector Machine, SVM) 和 Adaboost 算法结合起来, SVM 是基于结构风险最小化准则的新型机器学习算法, 适合小样本学习并且能够有效地抑制过拟合问题, Adaboost 基于最小化训练错误率, 一般使用易训练的分类器作为弱分类器。由于 SVM 比较难训练, 因此将样本集划分形成多个训练集, 然后利用正样本和不同的负样本组成不同训练集反复训练, 最后通过 Adaboost 对训练集生成的 SVM 模型筛选出具有最小错误率的 SVM 分类器并且采用投票机制形成最终的强分类器。实验结果表明, 在 FPPW (false positive per window) 为 10^{-5} 时检测率能够达到 30%, 检测效果优于单个 SVM 算法训练出来的分类器模型, 用行人测试库测试, 该方法取得了较好的检测效果并且具有较强的鲁棒性。

关键词: Adaboost; 支持向量机; 行人检测; 方向梯度直方图; 分类器

中图分类号: TP301.6 文献标识码: A doi: 10.3969/j.issn.1671-7449.2013.01.014

A Pedestrian Detection Method Combining SVM with Adaboost Algorithm

ZHANG Lihong, LI Lin

(School of Physics and Electronic Engineering, Shanxi University, Taiyuan 030006, China)

Abstract: Pedestrian detection remains an important task in computer vision. In order to further improve the precision of pedestrian detection, a method of combining support vector machine (SVM) with Adaboost algorithm is employed in this paper. SVM is a new machine learning method based on structural risk minimization, which is applicable to the case of small sample size learning and can restrain overfitting efficiently. Adaboost algorithm is based on minimizing training error rate, and its component classifiers generally should be easy to train. Considering SVM is hard to train, we divided the sample set into many small training groups. Then it generated many new training data by combining positive samples with different negative samples which were used for repeated training. Finally, it selected many weak classifiers with low training error rate from SVM training data. Those weak classifiers were assembled to become a strong classifier by voting mechanism. Experimental results demonstrated that the detection rate is about 30% at 10^{-5} FPPW (false positive per window) which is much more better than the classifier model trained by single SVM algorithm. This algorithm has strong robustness and is capable of obtaining relatively good detection results in pedestrian testing dataset.

Key words: Adaboost; support vector machine; pedestrian detection; histograms of oriented gradients; classifier

* 收稿日期: 2012-09-16

基金项目: 山西省高校高新技术产业化项目(2010002)资助

作者简介: 张丽红(1968-), 女, 副教授, 硕士生导师, 主要从事图像处理 and 模式识别研究。

基于机器视觉的行人检测是一个很重要的研究领域,有着广泛的应用价值,其涉及到人机交互、车辆辅助驾驶和安全监控等方面^[1].但是由于人体的非刚性以及复杂背景的影响且在实际应用中要求极高的实时性,这些限制因素导致行人检测成为一项艰难的研究课题^[2-4].

目前有许多行人检测的方法,比如模板匹配方法和基于统计学习方法,模板匹配是一种比较简单且很早就应用了的方法,人的头部轮廓有很好的不变性,在这种情况下,模板匹配能够较好地描述目标, Hausdorff 距离是一种适合度量模板匹配的方法,被广泛应用于基于边缘特征的匹配^[5]. 基于学习的方法有两个很重要的方面,一个是表征目标的特征描述算子,另一个是采用的学习算法,众所周知的特征有边缘集特征^[6]、Harr-like 小波特征集^[7]、SIFT-like 特征^[8]、局部二值模式 (LBP, Local Binary Pattern)^[9] 以及梯度方向直方图 (HOG, Histogram of Oriented Gradient)^[10] 等,这些都是具有代表性的轮廓特征. 学习算法有 Adaboost^[11] 和 SVM 等, SVM 的主要优点是对目标模式变化具有一定的鲁棒性.

在机器学习中集成算法近年来得到了很多的关注,它是将许多分类精度一般的分类器按一定权重组合起来的. 有两个最常见的集成学习算法是 Bagging^[12-13] 和 Boosting^[14-15], Bagging 是一种并行训练过程,反复抽样每个个体分类器所采用的训练样本,样本都是从训练集中按等概率抽取的,因此 Bagging 的各子网能够很好地覆盖训练样本空间,有着良好的稳定性. Boosting 的一个改进型是 Adaboost,它注重分类错误的样本,将个体子网分类错误的训练样本的权值提高,降低分类正确的样本权值,在每次迭代过程中根据分类错误率自动调整样本权重.

本文利用集成算法思想,根据最小分类错误率将许多分类精度一般的 SVM 按照一定的权重组合起来,通过划分训练样本形成多个训练集从而能够避免弱分类器不能采用较难训练的学习算法的这个问题.

1 算法基本理论

1.1 特征提取

本文所选择的特征是 Dalal 等人提出的梯度方向直方图 (HOG, Histograms of Oriented Gradients),它通过对局部区域内目标边缘的分布进行提取,可以很好地表示目标的形状. 本文选择的检测窗口大小是 64×128 , 用 I 表示一幅图像, $I(x, y)$ 表示图像在像素点 (x, y) 处的灰度值,那么该图像的 HOG 可通过如下步骤求得:

1) 利用梯度模板 $[-1, 0, 1]$ 计算图像中各点的梯度值

$$\begin{aligned} G_x(x, y) &= I(x+1, y) - I(x-1, y); \\ G_y(x, y) &= I(x, y+1) - I(x, y-1), \end{aligned} \quad (1)$$

式中: x, y 表示图像中某像素点的横纵坐标; $G_x(x, y)$ 和 $G_y(x, y)$ 分别表示该点在 x 方向和 y 方向的梯度值.

2) 计算该点的梯度强度 $M(x, y)$ 和梯度方向 $\theta(x, y)$

$$M(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)}, \quad (2)$$

$$\theta(x, y) = \arctan(G_x(x, y) / G_y(x, y)) + \frac{\pi}{2}. \quad (3)$$

为了提高 HOG 的抗噪性,将 $\theta(x, y)$ 限定在 $[0, \pi]$ 范围内.

3) 把 $[0, \pi]$ 均匀分成 9 个区间,利用三线性插值将每个梯度方向角按一定比例分配到临近的区间上.

4) L_1 范数归一化处理消除光照噪声等影响,收集到每个小块的特征 $B_i = (f_1, f_2, \dots, f_{36})$.

5) 按重叠率为 0.5 移动小块,重复 1) 到 4),串联检测窗口中所有的特征向量,最后生成的人体特征向量表示为 $F = (B_1, B_2, \dots, B_{105}) = (f_1, f_2, \dots, f_{3780})$.

1.2 支持向量机

支持向量机 (SVM) 是一种基于统计学习理论的模式识别方法^[16],基本思想是通过一个非线性变换

将输入空间的数据 x 映射到一个高维空间中, 然后在新空间求最优线性分类面, 这样非线性变换通过核函数 $K(x_i, y_j)$ 来实现^[17]. 最后的分类判别公式为

$$f(x) = \operatorname{sgn}(w \cdot x + b) = \operatorname{sgn}\left\{\sum \lambda_i y_i K(x_i, y_j) + b\right\},$$

(4)

式中: 样本点 (x_i, y_i) , 样本类别标识 $y_i = \{-1, +1\}$, 不为 0 的 λ 所对应的样本就是支持向量(support vector); b 是常量, 可以由任一支持向量求得.

常用的核函数有:

1) 线性核函数

$$K(x, y) = x \cdot y.$$

(5)

2) 多项式核函数

$$K(x, y) = (x \cdot y + 1)^d, \quad d = 1, 2, \dots.$$

(6)

3) 径向基核函数

$$K(x, y) = \exp(-\gamma \|x - y\|^2).$$

(7)

4) Sigmoid 核函数

$$K(x, y) = \tanh(b(x \cdot y) - c).$$

(8)

相对其他核函数来说, 线性核函数具有很快的运算速度, 因此本文实验中选用的是线性 SVM.

1.3 Adaboost 算法

在 Adaboost 算法中, 已知包含正样本和负样本的特征集 (x_i, y_i) 作为训练集, 初始化时将正负样本的权值 w 设为相同. 然后针对每个特征训练得到一个弱分类器, 用弱分类器去测试训练样本得到每个弱分类器的错分类器 ϵ , 根据错分率 ϵ 更新样本特征的权值 w , 生成新的训练样本空间并用来训练下一个过程 t 中的分类器, 更新样本权重的原则是提高分类错误的训练样本的权重, 降低分类正确的样本权重, 在每次循环 t 结束时根据错分率 ϵ 得到最优的弱分类器. 最后将选取出的 T 个弱分类器按一定方式组合起来就形成了最终的强分类器. 算法流程为:

- 1) 已知样本 $(x_1, y_1), \dots, (x_n, y_n)$, 其中 $y_i = -1, 1$ 分别表示负样本和正样本.
- 2) 初始化正负样本权重 $w_1, i = 1/(2m), 1/(2l)$, m 和 l 分别代表正负样本的个数.
- 3) 循环 $t = 1, \dots, T$.
- ① 归一化权重: $w_{t,i} \leftarrow \frac{w_{t,j}}{\sum_{j=1}^n w_{t,j}}.$
- ② 对每个特征 j 训练一个弱分类器 h_j , 错分率 $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|.$
- ③ 选择最小的错分率 ϵ_j 对应的弱分类器 h_j .
- ④ 更新权重: $w_{t+1,i} = w_{t,j} \beta_t^{1-e}.$
- 当 x_i 正确分类时 $e = 0$, 错误分类时 $e = 1, \beta_t = \frac{\epsilon_t}{1 - \epsilon_t}.$
- 4) 最后的强分类器:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0 & \text{其他,} \end{cases}$$

式中: $\alpha_t = \log \frac{1}{\beta_t}.$

2 本文的算法: SVM 及 Adaboost 组合方法

当用 Boosting 算法联合生成强分类器时, 需要弱分类器是很容易训练得到的, 本文选用线性 SVM

作为弱分类器, SVM 并不容易训练得到, 本文采用将训练样本集 (x_i, y_i) 按个数逐渐递增的方式分成 T 份, 每份中允许包含重叠的样本, 每份训练样本产生 $tcount$ 个 SVM 模型, 根据这 $tcount$ 个 SVM 模型用在测试集上面得到错分率选择最优的一个 SVM 模型, 最后将选取了 T 个最优的弱分类器按一定的方式组合起来形成最终的强分类器. 改进后的算法流程为:

- 1) 已知样本 $(x_1, y_1), \dots, (x_n, y_n)$, 其中 $y_i = -1, 1$ 分别表示负样本和正样本.
- 2) 初始化正负样本权重 $w_{1,i} = 1/(2m), 1/(2l)$, m 和 l 分别代表正负样本的个数.
- 3) 将样本集 (x_i, y_i) 按逐渐递增的方式分成 T 份.
- 4) 循环 $t = 1, \dots, T$.
- ① 归一化权重:

$$w_{t,i} \leftarrow \frac{w_{t,j}}{\sum_{j=1}^n w_{t,j}}.$$

- ② 选择第 t 份的样本集.

- a. 将该样本平均分为 $tcount$ 份, 总共训练得到 $tcount$ 个 SVM 模型.
- b. 对每个 SVM 弱分类器 h_j , 计算错分率 $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.

- ③ 从 $tcount$ 个 SVM 弱分类中选择最小的错分率 ϵ_j 对应的弱分类器 h_j .
- ④ 更新权重:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e}.$$

当 x_i 正确分类时 $e = 0$, 错误分类时 $e = 1, \beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$.

- 5) 最后的强分类器:

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0 & \text{其他,} \end{cases}$$

式中: $\alpha_t = \log \frac{1}{\beta_t}$.

通过将训练集分成许多部分, 训练得到不同的 SVM 分类器模型, 再从这些分类器中选择最优的组合起来形成最终的强分类器. 这样改进有 3 点好处: ① 提高了检测的准确性. 实验证明本文中改进方法的误检率降低了, 相应地提高了检测的准确性, 下文中的实验结果和分析将作进一步详细阐述; ② 避免了难训练的分类器不能做弱分类器这个问题. SVM 属于难训练的一种分类器^[17], 不适合作弱分类器, 经过本文算法改进, 使得 SVM 能够作弱分类器; ③ 大大缩短了训练的时间. 若没有将样本分组训练弱分类器, 直接送入 Adaboost 框架中训练需要花费 3 d 时间, 由于改进了算法, 本文改进方法大大缩短了训练时间, 采用最终的强分类器采用本文算法训练耗时 8 h.

3 实验结果和分析

在 INRIA 数据集上进行了验证, 在 VC 和 OpenCV 开发环境中实现, 与现有方法进行比较.

3.1 本文算法的检测结果

抽取了 12 180 个负样本特征和 2 416 正样本特征训练 20 层得到最终的强分类器, 在双核 2.2 GHz 32 b 的 CPU 和 2 G 内存上面本文算法耗时 8 h.

图 1 中的图片是从 INRIA 库抽取代表各种困难情况下检测的事例. 从图 1 中可以看出, 本文算法能够正确检测大部分行人目标, 如图 1(a) 中是严重弯曲的行人, 运用本文的算法能够很好地检测出来, 图 1(c) 中处在明亮和阴暗交界处的行人目标能够很好地被检测到, 对光照具有一定的鲁棒性. 采用本文

算法能够检测出 80% 以上行人目标.



图 1 各种情况下的检测实例结果

Fig. 1 Detection results for several cases

3.2 分类器性能比较

比较 Dalal & Triggs 和该文的方法, Dalal 的方法中做了困难样本(hard sample)多次训练, 本文在进行分类器性能比较的时候省去了这一步, 在同等条件下得到这两种分类器的 DET 曲线, 如图 2 所示.

在图 2 中可以看出, 在 FPPW 是 10^{-5} 的时候, 本文改进方法的误检率 miss rate 和 Dalal 相比占很多优势, 在实际应用中要求在低的误检率时, 有较高的检测率($1 - \text{miss rate}$). 在 FPPW 在 10^{-5} 到 10^{-2} 之间, 本文改进方法的误检率均低于 Dalal & Triggs 方法, 在 FPPW 为 10^{-5} 时检测率能够达到 30%, 检测效果优于单个 SVM 算法训练出来的分类器模型.

4 结 论

在 Boosting 思想中难训练的分类器是不适合做弱分类器的, 本文采用分割训练样本形成新的训练集避免了这个问题, 将线性 SVM 作为 Adaboost 的弱分类器使用, 通过这两点改进得到最终的强分类器, 实验结果表明本文的算法具有很好的检测效果.

在行人检测中改进分类器性能的方法关键是选取合适的特征和学习算法, 本文是在学习算法上进行了两个方面的改进, 目前在生成样本特征方面出现多种特征融合的方法, 进行多种特征联合能够改善检测效果, 这将是进一步研究的方向.

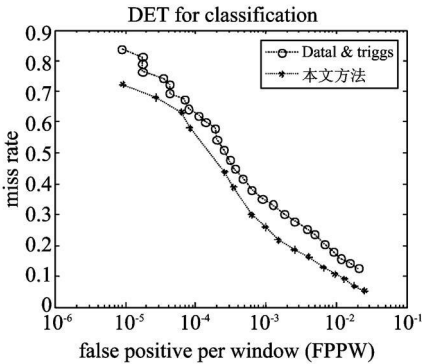


图 2 检测性能及比较

Fig. 2 Detection performance and comparison

参考文献:

- [1] 李豪杰, 林守勋, 张勇东. 基于视频的人体运动捕捉综述[J]. 计算机辅助设计与图形学学报, 2006, 18(11): 1645-1651.
Li Haojie, Lin Shouxun, Zhang Yongdong. A survey of video based human motion capture[J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(11): 1645-1651. (in Chinese)
- [2] Li L, Leung M K H. Unsupervised learning of human perspective context using ME- DT for efficient human detection in surveillance[C]. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition, 2008: 23- 28.
- [3] Wojek C, Walk S, Schiele B. Multi- Cue onboard pedestrian detection[C]. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition, 2010: 20- 25.
- [4] Xu R, Zhang B, Ye Q, et al. Cascade L1- norm classifier for pedestrian detection[C]. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition, 2010: 19- 22.
- [5] Anil J, Robert D, Mao J Statistical pattern recognition: A review[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22: 4- 37.
- [6] Wu B, Nevatia R. Detection of multiple, partially occluded humans in a single image by Bayesian combination of Edgelet part detectors[C]. In: Proc. of the IEEE Int'l Conf. on Computer Vision, 2009: 17- 21.
- [7] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]. Conference on Computer Vision and Pattern Recognition, 2001: 1- 4.
- [8] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2005, 27: 1615- 1630.
- [9] Wu B, Nevatia R. Optimizing discrimination- efficiency tradeoff in integrating heterogeneous local features for object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2008: 23- 28.
- [10] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. Conference on Computer Vision and Pattern Recognition, 2005: 886- 893.
- [11] Friedman J, Hastie T, Tibshirani R. Additive logistic regression: A statistical view of boosting[J]. The Annals of Statistics, 2000, 28: 337- 407.
- [12] Breiman L. Bagging predictors[J]. Machine Learning, 1996(24): 152-158.
- [13] Gonzalo M, Alberto S. Out- of- bag estimation of the optimal sample size in bagging[J]. Pattern Recognition, 2010, 43: 143- 152.
- [14] Schapire R E. The boosting approach to machine learning: an overview[C]. In: MSRI Workshop on Nonlinear Estimation and Classification, 2002: 21- 25.
- [15] Yongdai K, Byung Y K, Seong K. Component- wisely sparse boosting[J]. Journal of the Korean Society, 2011, 40: 487- 494.
- [16] Cui Jianguo, Li Zhonghai. The application of support vector machine in pattern recognition[C]. IEEE International Conference on Control and Automation, 2007: 3135- 3138.
- [17] Belousov A I, Verzakov S A, Von Frese J. A flexible classification approach with optimal generalization performance: support vector machines[J]. Chemometrics and Intelligent Laboratory Systems, 2002, 64: 15- 25.