

# 视频异常行为识别与分级预警系统

杨 谦 何小海\* 蒋 俊 吴晓红

(四川大学电子信息学院 成都 610065)

**摘 要** 参考格灵深瞳分级评价体系并引入迟滞比较器相关思想,结合最近邻与SVM双层分类器学习,建立了针对目标入侵、目标高速运动、目标遗留物与人群聚集逃离、人群打架斗殴、人群骚乱六种常见目标异常行为的自动分类与分级预警系统。①提出并实现了一套较完备的异常行为分级预警系统;②在行为分析之前以人群密度与能量为特征引入最近邻分类器实现个体行为与群体行为的预分类;③通过引入迟滞比较器实现高速运动行为的稳定预警;且该方法具有一定普及意义。分别在标准库和自行拍摄的视频上进行实验验证。实验证明,该系统能够稳定实现对上述六种普遍异常行为的分类分级预警,实现了群体分析与个体分析、检测与识别、分类与预警的一体化。

**关键词** 分级预警系统 双层分类器学习 最近邻 SVM 迟滞比较

中图分类号 TP391.41;

文献标志码 A

近几年,移动目标异常行为的识别成为计算机视觉领域最为活跃的研究方向之一<sup>[1-4]</sup>。与此同时,一套完备的智能监控系统在银行、重要仓库、军事基地、边海防线等敏感区域都有着较大的需求。欧盟的ADVISOR系统<sup>[5]</sup>就是一个应用于公共交通的安全管理系统;它涵盖了个体行为分析,也包括群体行为分析,还对人机交互方面有了部分涉及。然而,目前大多数监控系统都还停留在普通网络视频监控的概念上,智能化的程度不高。一些监控系统只能识别出部分简单的异常行为,光照对系统精度影响较大;且处理速度、精度都不尽如人意。随着市场上对视频监控系统需求的爆炸性增大,开发出一套稳定、快速、多功能的智能监控系统迫在眉睫。

能否准确实时实现分级预警是衡量智能监控系统的重要标准之一。截至目前,国内外鲜有针对异常行为敏感程度的评价体系或模型。部分国内外学者和安全研究专家开始尝试用风险管理<sup>[6-8]</sup>的方法对道路交通安全评价体系和模型,取得了一定的进展。Leonelli采用模糊数学法<sup>[6]</sup>来评估不确定性风险影响因素;Fabiano将危险货物运输过程中的风险影响因素<sup>[7]</sup>划分为道路固有特征、天气条件和交通状况等类,提出了面向事故现场的评价模型。然而,由于在视频监控中的异常行为种类众多,且之间没有太多联系,因此建模困难<sup>[8]</sup>。目前较为先进的格

灵深瞳无人监控安防系统<sup>[9]</sup>也只是将不同的异常行为通过人为指定预警级别实现分级预警。

提出并实现的异常行为识别与分级预警系统结构图1所示。

## 1 系统设计

### 1.1 视频分类

移动目标异常行为的识别算法大致可以分为两类:①基于物理建模的行为识别方法;②基于分类器学习的行为识别方法。基于物理建模的方法主要用来识别物理特征较明显的行为,因此只需对此类行为建立合适的物理模型即可用于行为分类与判断。同时,基于物理建模的行为识别方法由于需要对目标准确的检测与跟踪,常常用于场景较简单的情况。基于分类器学习的方法则主要用于特征多样、复杂的行为分类,特别是对于场景复杂的群体行为,在这类场景中无法检测跟踪独立个体,因此往往先提取全局特征,然后运用机器学习的方法分类学习。

结合上述两类行为识别方法的不同适用场景,本文提出一种视频分类算法:首先依据Davies等人提出的人群密度估计算法<sup>[10]</sup>及Guogang Xiong等人提出的能量模型<sup>[11]</sup>,计算出待处理视频的人群密度与能量,通过事先训练好的最近邻分类器判断场景的复杂度,从而实现视频的分类。

### 1.2 个体行为分析

如果视频分类结果为简单场景视频,系统采用VIBE算法<sup>[12]</sup>对目标实时检测跟踪,之后对入侵、遗留物、高速运动这三种行为分别建模,从而实现三种单人行为的识别。

2015年1月6日收到 国家自然科学基金委员会和中国工程物理研究院联合基金(11176018)资助

第一作者简介:杨 谦(1991—),男,硕士研究生。研究方向:通信与信息系统、智能监控等。E-mail: 873132246@qq.com。

\* 通信作者简介:何小海。E-mail: hxx@scu.edu.cn。



图1 系统结构图

Fig. 1 The system structure diagram

### 1.3 群体行为分析

若视频分类结果为场景复杂的群体行为视频,则考虑运用模式识别中机器学习的方法对行为分类。在这类复杂场景下,对每个目标的检测跟踪变得异常困难,只能对整个视频帧提取特征,最后运用机器学习的手段实现分类。Heng Wang 等人提出用视频稠密轨迹特征<sup>[13]</sup>表征行为特征,并运用特征字典实现特定行为的识别分类;Piotr Dollar 等人通过对视频帧时间域与空间域的滤波得到时空特征点<sup>[14]</sup>,对特征点周围的时空立方体做梯度描述形成特征向量,最后通过 SVM<sup>[15]</sup>机器学习实现特定行为的识别。该方法所提取的特征点常出现在运动剧烈的区域且经过 PCA 降维大大减少了算法运算量,因此本文采用此方法实现群体行为的识别。

### 1.4 行为分级与预警

为了实现系统稳定分级预警,本文将格灵深瞳系统的分级思想与迟滞比较器<sup>[18]</sup>引入分级预警体系中,实现了异常行为的抗干扰分级预警。

## 2 系统算法设计与实现

### 2.1 人群密度与能量为特征的最近邻分类器

#### 2.1.1 人群密度估计

本文采用 Davies<sup>[10]</sup>提出的通过数像素个数的人群密度统计算法。具体步骤如下:

1) 连续选取两帧图像,转为单通道灰度图后再对其进行  $3 \times 3$  高斯滤波。

2) 对上述步骤得到的灰度图像做 canny 边缘检测,对所得结果相减得到边缘差。

3) 对步骤 2) 得到的边缘差图像用形态学腐蚀膨胀运算消除噪声,然后以该图像为遮罩对原始图

像二值化,统计灰度不为 0 的点的个数作为人群密度的度量。

#### 2.1.2 人群能量估计

Guogang Xiong 等人<sup>[11]</sup>提出的能量计算公式如下:

$$E_p = \sum_{i=1}^X \sum_{j=1}^Y m_{ij} g_{img}(H + Y - y_{ij}) \quad (1)$$

式(1)中,  $m_{ij}$  判断某像素点是否为前景点,  $m_{ij} \in \{0, 1\}$ , 若为前景点取值 1, 否则取值 0;  $g_{img}$  是一个固定常量,其值的大小取决于特定的场景,本文中取 10;  $H$  是前景到摄像头最短距离的粗略估计,对于固定的摄像头而言,  $H$  也是常量;  $X, Y$  分别代表图像的宽高;  $y_{ij}$  代表图像在  $y$  轴上的坐标。由上可知,由公式(1)计算出的能量与摄像头到目标的距离无关。将  $e$  定义为每单位人的平均能量,可以得出人群能量  $N$  的估计公式:

$$N = \frac{kE_p}{e} \quad (2)$$

式(2)中,  $k$  表征了图像中的遮挡程度。

由上述人群密度和能量模型分别计算得出的结果最终会融合成一个特征向量,作为本文系统第一层分类器的输入。选取最近邻分类器对视频分类。

#### 2.1.3 最近邻分类器

最近邻分类器的理论依据是  $K$  最近邻算法(KNN),它采用向量空间模型对输入样本进行分类<sup>[16]</sup>。若样本类别相同,样本间的相似度较高。因此本文通过计算与已知类别样本之间的相似度,来估计得到未知类别样本可能的类别。

分类过程如下:

1)  $k$  的取值与已知类别的样本数量有关;

在本文实现的系统中,用于训练学习的样本数为 50 个,  $k$  取 7。

2) 以欧式距离作为样本间的距离度量方法,得出待分类数据和所有已知类别样本中距离最近的 7 个样本。

3) 统计这 7 个样本中各个类别的数量。把其中数量最多的已知样本的类别作为待测样本类别。

最近邻算法的缺点是计算量大,特别当输入为高维特征向量时。本文阐述的系统中,人群密度与能量已足够用来区分个体行为与群体行为,特征维度低,可以快速地用来对个体行为与群体行为准确分类。

## 2.2 个体行为分析

当环境较为简单,特别是单人行为时,适合采用基于检测跟踪的行为分析方法。本文首先采用 VIBE 算法检测出前景目标<sup>[12]</sup>并跟踪,然后依据事先建立好的物理模型实现特定行为的识别:①通过目标跟踪框与预设的被保护区域对比分析得出目标是否入侵被保护区域;②利用三维测距模型对目标实时测速,并与事先设定的阈值比较得出目标是否高速运动;③利用不同更新率的双高斯模型实现遗留物的检测。

## 2.3 群体行为分析

在复杂环境下(人群密集的场景或者运动信息复杂的区域),本文通过提取时空特征立方体的梯度特征,加载到事先训练好的 SVM 分类器实现三种群体异常行为(打架斗殴、人群骚乱、聚集逃离)的分类。

基于三维时空特征立方体的群体行为分析步骤如下:

1) 时空兴趣点检测。本文选取 Dollar 的 Cuboids 的特征提取方法<sup>[14]</sup>,特征点是对视频帧时间域和空间域滤波的结果。Cuboids 特征检测算子响应函数如下公式所示:

$$R = (I \otimes g \otimes h_{ev})^2 + (I \otimes g \otimes h_{od})^2 \quad (3)$$

式(3)中  $I$  为输入的视频序列,  $g$  是对输入视频的空间域平滑滤波的二维高斯函数  $g(x, y, \sigma)$ ,  $h_{ev}$ ,  $h_{od}$  则是对视频序列时间域滤波的 Gabor 滤波器,其定义公式如下所示:

$$h_{ev}(t; \tau, \omega) = -\cos(2\pi t\omega) e^{-\frac{t^2}{\tau^2}} \quad (4)$$

$$h_{od}(t; \tau, \omega) = -\sin(2\pi t\omega) e^{-\frac{t^2}{\tau^2}} \quad (5)$$

选取  $\omega = 4/\tau$ 。

从上述响应函数可以看出,周期性的行为会得到较大的响应。时间域与空间域的局部最大值同样会得到较大的响应值。而那些简单平滑的运动(如平移)响应函数输出值很小。

2) 时空兴趣点描述子。为了减少微小平移、光照变化等因素的不利影响,与简单的直方图统计或者归一化的亮度向量不同,本文通过求取每个特征点周围时空立方体在各个方向的梯度组成描述矢量。由于采用上述方法得到的描述矢量规模庞大,不利于后面的学习分类。因此,本文采取 Pearson 于 1901 年提出的 PCA 思想对描述矢量降维。之后以欧式距离为度量,采用 K-Means++ 算法聚类,得到  $N$  个聚类中心(本文  $N$  取 50)。这 50 个码本构成一个行为码本。利用这一行为码本对测试视频描述矢量中的每个值进行归类,从而得到归类后的描述矢量。本文以归类后的描述矢量的统计直方图作为视频序列最终的描述矢量。

3) 由步骤 2 可知,每个视频最终映射为一个统计直方图。为了提高算法的识别效果并减小计算复杂度,本文采用支持向量机算法对直方图进行分类与识别。

## 2.4 改进的异常行为分级预警

### 2.4.1 格灵深瞳

格灵深瞳<sup>[9]</sup>是计算机视觉领域致力于人工智能的科技公司。其主要宗旨在于让计算机像人一样主动获取视觉信息并进行精确的实时分析,从而释放人工智能最大的潜能为人所用。2014 年 4 月起首款产品“深瞳智能无人监控系统”工程样机开发完成。基于其人为指定行为分级的思想,指定本文所识别的六种异常行为(入侵、高速运动、遗留物、聚集逃离、打架斗殴、人群骚乱)级别为 LEV-EL1-6。本文系统所搭建的针对 6 种异常行为的分级系统结构如图 2 所示。

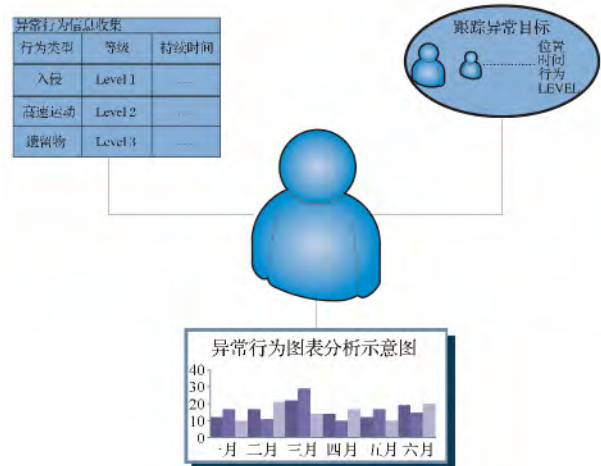


图 2 异常行为分级分析系统

Fig. 2 Abnormal behavior classification analysis system

### 2.4.2 迟滞比较器

引入迟滞比较器<sup>[17]</sup>可以使得电路系统具备较



强的抗干扰性能,迟滞比较器电路图与电压传输特性如图3所示。

由图3(b)可知:迟滞比较器存在回差电压,这使得电路抗干扰能力得到显著提高。回差  $\Delta U_{TH}$  ( $2U_{TH}$ ) 越大,电路的抗干扰能力越强。输入信号可能会受干扰或由于其他原因发生变化时,然而,只要这种输入信号的变化量不超过回差  $\Delta U_{TH}$ ,这种存在比较器的电路的输出电压就不会来回变化。

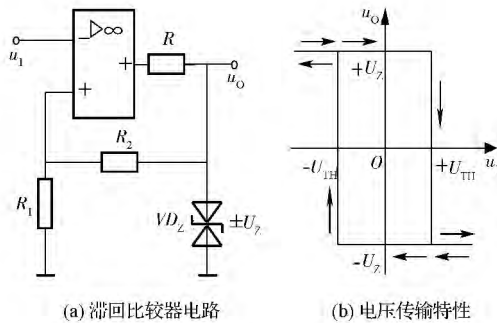


图3 迟滞比较器

Fig.3 Hysteresis comparator

在分级预警系统实现中,由于单人行为分析中常常会由于环境的突变甚至是目标的误匹配跟踪出现其状态的瞬变,因此引入一个“等级”参量,初始值设为4,以单人高速运动行为分析为例:若两帧之间检测到高速运,则该参量加1,否则减1,当该参量

达到+5时,提示高速运动并显示高速运动对应的异常行为等级,该高速运动标签一直存在直到参量减到了-5。采用迟滞比较的方法可以得到更加稳定的行为分类结果,更加可信且贴近实际应用。

### 3 实验结果与分析

在标准数据库 PETS 2009 加以自行拍摄的视频上进行测试,实验结果验证了所建立的异常行为分级预警系统的有效性。

图4第一排为实现的系统在自行拍摄的个体行为视频的测试结果。由这三幅图(从左到右)可以看出,系统可以准确识别出入侵、高速运动、遗留物这三种个体行为,并给予分级预警。

图4中间一排为实现的系统在自行拍摄的群体行为视频的测试结果;第三排为实现的系统在标准数据库 PETS 2009 上的测试结果。为了比较算法的有效性,采用精确度(ACC)作为算法的衡量指标,根据曲线理论知:

$$ACC = \frac{TP}{N} \quad (6)$$

式(6)中,TP为被正确识别的样本数,N为样本总数。ACC为由上式计算得出的精确度。在标准视频库(set1)与自行拍摄的视频(set2)中分别用三种算法测试,得出结果如表1所示。



图4 系统测试结果

Fig.4 Test results

表 1 算法效果对比

Table 1 The comparison of different algorithm

算法	Set1	Set2	平均
RBF + Cuboids(ours)	0.812	0.788	0.800
Linear + Cuboids	0.744	0.725	0.735
RBF + Harris3D	0.554	0.498	0.526

由表 1 可以看出,与采用线性核函数相比,采用 RBF 核得到的识别率更高;与 Harris3D 特征相比,采用时空立方体得到的识别率更高。由此可以看出本文采用的“Cuboids 特征 + BRF 核 SVM”算法的精确率最高。

为了对比分级时引入迟滞比较器前后对实验结果的影响,我们做了如下对比试验:

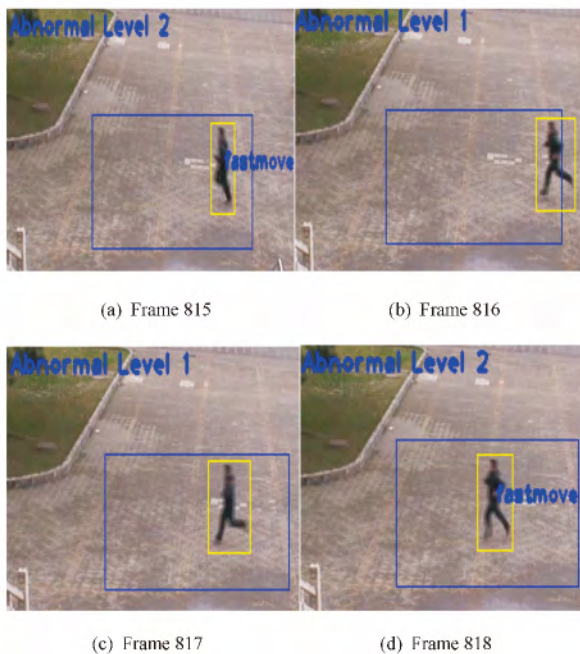


图 5 未引入迟滞比较器的分析结果

Fig. 5 Test results not applying Hysteresis comparator

未引入迟滞比较器时(图 5),视频在相对连续帧之间的行为分析结果会不断变化(即前一帧判断为高速运动异常行为、后一帧又判断为正常行为,并如此反复),这样反复不定的行为识别显然会对后续的预警产生不利的影响。引入迟滞比较器之后(图 6),连续帧的行为识别相对稳定,更有利于下一步的预警。实验证明了迟滞比较器可以实现行为的稳定识别与分级预警。

## 4 结论

研究并实现了一种可以广泛应用于视频监控中的基于双层分类器的异常行为识别与分级预警系统。该系统实现了入侵、高速运动、遗留物检测、聚集逃离、打架斗殴和人群骚乱这六种常见异常行为

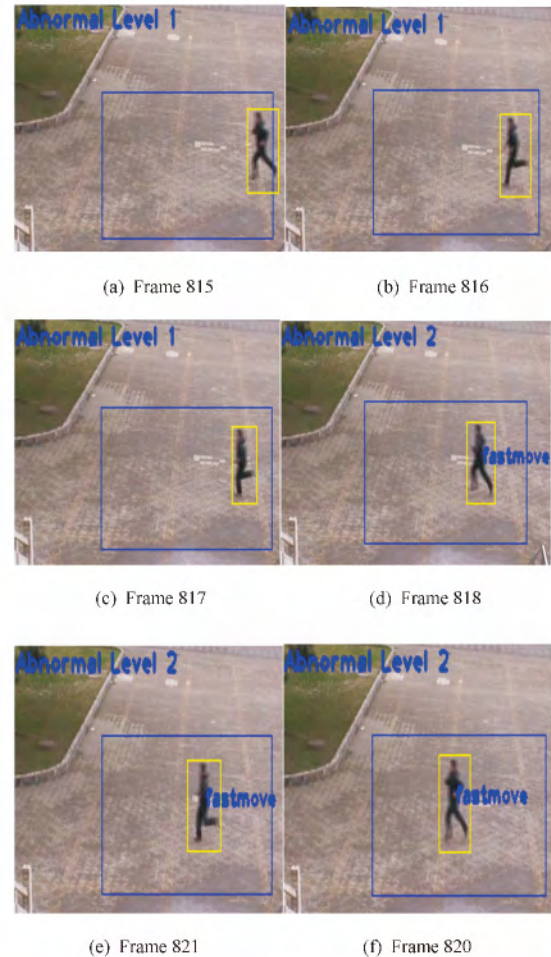


图 6 引入迟滞比较器后的分析结果

Fig. 6 Test results applying Hysteresis comparator

的识别与分级预警。实验证明该系统在行为分析分级时的有效性。群体行为分析时由于只采用了时空立方体梯度特征,特征较为单一,因此之后的研究致力于融合多特征行为分类,并试图在本文实现分级的基础上提出更加普遍、完善的分级体系。

## 参考文献

- 1 Mahadevan V, Li W, Bhalodia V, et al. Anomaly detection in crowded scenes. *Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2010: 1975—1981
- 2 Saligrama V, Chen Z. Video anomaly detection based on local statistical aggregates. *Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2012: 2112—2119
- 3 Aggarwal J K, Ryoo M S. Human activity analysis: a review. *ACM Computing Surveys (CSUR)*, 2011; 43(3): 16
- 4 Ikizler-Cinbis N, Sclaroff S. Object, scene and actions: Combining multiple features for human action recognition. *ECCV 2010*, Springer Berlin Heidelberg, 2010: 494—507
- 5 Wipke K B, Cuddy M R, Burch S D. ADVISOR 2. 1: a user-friendly advanced powertrain simulation using a combined backward/forward approach. *Vehicular Technology, IEEE Transactions on*, 1999; 48

- (6): 1751—1761
- 6 Leonelli P, Bonvicini S, Spadoni G. New detailed numerical procedures for calculating risk measures in hazardous materials transportation. *Journal of Loss Prevention in the Process Industries*, 1999; 12(6): 507—515
  - 7 Fabiano B, Curro F, Palazzi E, *et al.* A framework for risk assessment and decision-making strategies in dangerous good transportation. *Journal of Hazardous Materials*, 2002; 93(1): 1—15
  - 8 Feng S, Xu L D. Decision support for fuzzy comprehensive evaluation of urban development. *Fuzzy Sets and Systems*, 1999; 105(1): 1—12
  - 9 36kr. Deepglint. 2014-6-15. <http://www.36kr.net/gelingshentong>
  - 10 Davies A C, Yin J H, Velastin S A. Crowd monitoring using image processing. *Electronics & Communication Engineering Journal*, 1995; 7(1): 37—47
  - 11 Xiong G, Wu X, Chen Y L, *et al.* Abnormal crowd behavior detection based on the energy model. *IEEE International Conference on Information and Automation*, 2011: 495—500
  - 12 Barnich O, Van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences. *IEEE Transactions on Image Processing*, 2011; 20(6): 1709—1724
  - 13 Wang H, Klaser A, Schmid C, *et al.* Action recognition by dense trajectories. *Computer Vision and Pattern Recognition (CVPR)*, 2011: 3169—3176
  - 14 Dollár P, Rabaud V, Cottrell G, *et al.* Behavior recognition via sparse spatio-temporal features. *2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, IEEE, 2005: 65—72
  - 15 Burges C J C. A tutorial on support vector machines for pattern recognition. *Data mining and Knowledge Discovery*, 1998; 2(2): 121—167
  - 16 Coomans D, Massart D L. Alternative  $k$ -nearest neighbour rules in supervised pattern recognition: part I.  $k$ -nearest neighbour classification by using alternative voting rules. *Analytica Chimica Acta*, 1982; 136: 15—27
  - 17 Abdi B. Programmable hysteresis comparator: U. S. Patent 5,386,192, 1992-10-13

## Ranked-warning System on Identification of Abnormal Behavior in Video Monitoring

YANG Qian, HE Xiao-hai\*, JIANG Jun, WU Xiao-hong

( College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, P. R. China)

**[Abstract]** Taking into account grid deep spiritual pupil grade-evaluation system and main ideas of hysteresis comparator, the nearest neighbor classifier and SVM learning was combined with, which aims to establish an automatic Classification and Grading Warning System recognizing six common target abnormal behavior including invasion, fast-moving, remnants, fleeing, fights and riots. Following three aspects were contributed: firstly, a comprehensive system was proposed dealing with classification and warning of abnormal behavior. Secondly, before recognizing abnormal behavior the population density and the energy was characterized, which were input of the nearest neighbor classifier, achieving pre-classification of individual behavior and group behavior. Thirdly, The stable behavior warning by introducing hysteresis comparator was achieved, and this method has certain universal significance. Experiments were carried out on the standard library and video sets shoot. Experimental results show that the system can achieve high warning and classification stability of the six abnormal behavior, which integrates self-analysis and analysis groups, detection and identification, classification and warning together.

**[Key words]** ranked-warning system      double layer classifier learning      nearest-neighbor classifier      SVM learning      hysteresis comparator