

改进的基于增强型 HOG 的行人检测算法^①



李文书, 韩 洋, 阮梦慧, 王志骁

(浙江理工大学 信息学院, 杭州 310018)

通讯作者: 李文书, E-mail: charlie@zstu.edu.cn

摘 要: 行人检测在人工智能系统、车辆辅助驾驶系统和智能监控等领域具有重要的应用, 是当前的研究热点. 针对 HOG 特征不明显、支持向量机 (SVM) 分类器计算复杂度高, 导致识别率低和检测速度慢的问题, 本文提出了一种改进的基于增强型 HOG 的行人检测算法. 该算法首先预处理原始图像并提取其 HOG 特征, 然后增强该特征生成增强型 HOG, 经 XGBoost 分类器进行行人检测. 在 INRIA 数据集上进行测试, 实验结果表明所提算法识别率高达 95.49%, 有效地提高了行人检测性能.

关键词: 行人检测; HOG 特征; 增强型 HOG; XGBoost 分类器

引用格式: 李文书, 韩洋, 阮梦慧, 王志骁. 改进的基于增强型 HOG 的行人检测算法. 计算机系统应用, 2020, 29(10): 199-204. <http://www.c-s-a.org.cn/1003-3254/7587.html>

Improved Pedestrian Detection Method Based on Enhanced HOG

LI Wen-Shu, HAN Yang, RUAN Meng-Hui, WANG Zhi-Xiao

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Pedestrian detection is a current research hotspot, which has important applications in the fields of artificial intelligence system, vehicle assistant driving system, and intelligent monitoring. In the process of pedestrian detection based on HOG feature, the HOG feature is not obvious, the SVM classifier has high computational complexity, resulting in low recognition rate and high missed detection rate, this study proposes an improved enhanced HOG feature combined with the eXtreme Gradient Boosting (XGBoost) classifier for pedestrian detection. Firstly, the original image is preprocessed to get saliency map and HOG features. Then, the contrast of HOG features is enhanced and the pedestrian detection analysis is carried out with XGBoost classifier. Tested with the INRIA dataset, the experimental results show that the proposed algorithm has a significant improvement in recognition rate and detection speed.

Key words: pedestrian detection; HOG feature; Enhanced HOG (EHOG); XGBoost classifier

近年来, 行人检测在众多领域得到了广泛的发展, 日渐成为计算机视觉和模式识别领域中的关键问题. 然而, 由于行人身材和姿态多样性、光线多变、背景复杂等原因, 使得行人检测依旧是计算机视觉研究的热点与难点^[1-3].

自 2005 年, Dalal 等人^[4,5]提出了方向梯度直方图

(HOG) 特征后, 行人检测技术进入快速发展阶段. HOG 特征用于描述图像中像素值的梯度信息, 该特征对行人边缘信息描述充分, 对光照变化不敏感, 目前仍是行人检测领域中应用最广泛的特征算子. Dalal 等人提出的 HOG 特征结合支持向量机 (SVM) 分类器的行人检测方法在 INRIA 数据集上检测精度达 90%, 但是检测

① 基金项目: 国家重点研发计划重点专项 (2018YFB1004901); 国家自然科学基金 (31771224, 60702069); 浙江省自然科学基金 (LY17C090011)

Foundation item: Major Project of National Key Research and Development Program of China (2018YFB1004901); National Natural Science Foundation of China (31771224, 60702069); Natural Science Foundation of Zhejiang Province (LY17C090011)

收稿时间: 2020-02-14; 修改时间: 2020-03-13; 采用时间: 2020-03-20; csa 在线出版时间: 2020-09-30

速度非常慢. 在 Dalal 等人的研究成果基础上, 学者们纷纷提出了改进方法.

为了提高检测精度, 陈锐等人^[6]提出了一种利用主成分分析法 (PCA) 对 HOG 特征进行降维后与局部二值模式 (LBP) 特征进行融合, 再使用 SVM 分类器进行分类的行人检测框架, 组合特征的检测方法比单一特征的检测方法更精确^[7,8]. 李盈盈等人^[9]提出了一种更精确的行人检测方法, 该方法将 LBP 特征、CSS 特征^[10]和 HOG 特征进行结合, 采用 Adaboost^[11]分类器进行行人检测, 提高了在 INRIA 数据集检测的识别率, 但由于提取特征种类增加导致检测速度变得极其慢.

为了提高检测速度, Chen PY 等人^[12]尝试从硬件层面改善, 提出了一种用于 HOG 特征提取的低成本高速硬件实现, 仿真实验表明, 它需要更少的硬件成本并且实现更快的工作速度. 龚露鸣等人^[13]提出一种用于视频流的快速行人检测方法. 该方法首先利用基于高斯模型的背景差分法进行运动目标提取, 然后在对运动目标进行人体识别, 通过缩小检测区域的方式提高了行人检测的速度, 达到可在视频流进行实时行人检测的目的.

此外, 近年来深度学习的算法在行人检测上的应用也越来越广泛, 基于深度学习的行人检测算法^[14-17]在速度和准确率上普遍优于传统机器学习算法, 但对硬件性能要求较高.

综上所述, 现有的 HOG 特征的研究和改进都是从外部寻找其他手段提高检测精度和检测速度, 而本文是从图像本身和 HOG 特征内部出发增强图像和 HOG 特征的表达能力.

1 HOG 特征

HOG 特征即方向梯度直方图特征, 是一种通过目标边缘密度的分布来对目标形状进行描述的特征算子, 被广泛应用于计算机视觉和图像处理中. HOG 特征提取的方式是依次提取目标局域特征再进行串联, 所有的提取工作只在目标局部区域进行, 避免了光照和图像几何形变的影响. 因此, HOG 特征具有良好的几何不变性和光照不变性. 此外, HOG 特征通过串联局部信息描述目标整体形状的方式对目标姿态的变化具有很好的宽容度. 在行人检测中, 只要人体姿态大致不变, 小幅度的肢体动作不会对检测效果产生影响. 因此, HOG 特征特别适用于行人检测.

HOG 特征的基本单位为细胞单元 (cell), 每个 cell 由 $n \times n$ 个像素组成, $N \times N$ 个 cell 组成一个块 (block), 具体特征提取过程如图 1 所示. HOG 特征的提取以 block 尺寸的滑动窗口为掩码, 进行滑窗扫描提取, 具体步骤为.

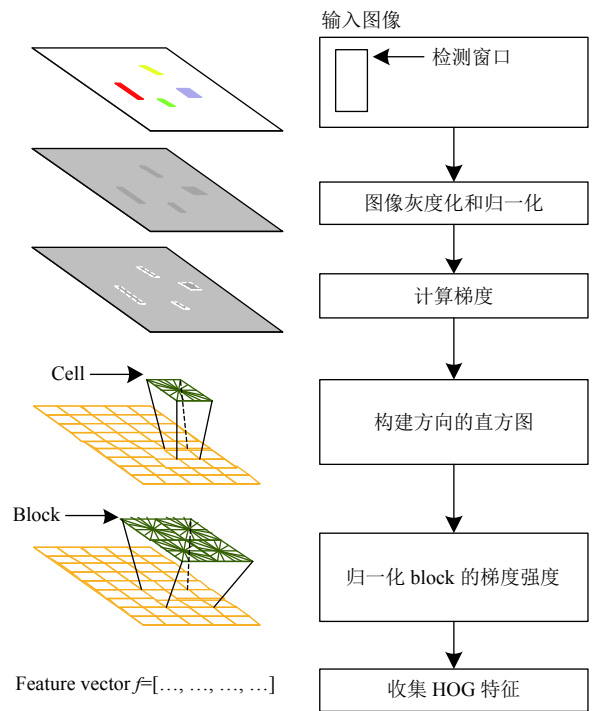


图 1 HOG 特征提取流程图

1) 将彩色图像转换为灰度图像, 使用 Gamma 校正法^[18,19]对图像进行归一化处理, 计算经 Gamma 校正后的图像中像素的水平 and 垂直梯度;

2) 其次将梯度方向分为九个区间, 如图 2 所示, 建立梯度直方图统计 cell 在每个区间梯度幅值的分布情况, 生成的 9 维特征向量为该 cell 的特征描述子;

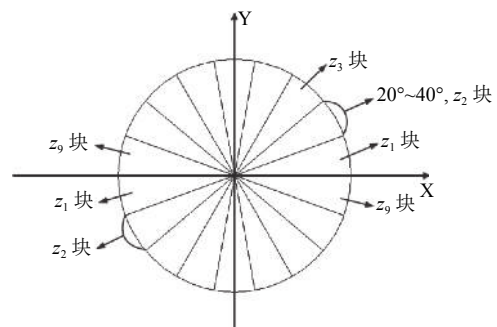


图 2 梯度方向划分图

3) 将 block 内所有 cell 的特征描述子串联起来并进行归一化处理, 生成该 block 的特征描述子;

4) 将所有 block 的特征描述子串联起来得到输入图像的 HOG 特征向量 f .

2 增强型 HOG

2.1 预处理操作

计算每个像素在整个图像上的全局对比度, 即该像素与图像中其他所有像素在颜色上的距离之和作为该像素的增强值, 得到新的图像.

图像 I 中某个像素 I_k 的增强值计算公式如下:

$$SalS(I_k) = \sum_{\forall I_l \in I} \|I_k - I_l\| = \|I_k - I_1\| + \dots + \|I_k - I_N\| \quad (1)$$

其中, I_k 为灰度值, 取值范围为 $[0, 255]$; N 表示图像中像素的数量. 给定一张图像, 每个像素的颜色值已知, 假定 $I_k = a_m$, f_n 表示图像中第 n 个像素的频数. 则上式可重构:

$$SalS(I_k) = f_0 \|a_m - a_0\| + \dots + f_{255} \|a_m - a_{255}\| \quad (2)$$

$$SalS(a_m) = \sum_{n=0}^{255} f_n \|a_m - a_n\| \quad (3)$$

$$g_{b,ave} = \frac{\sum_{c=1}^n g_{c,sum}}{n} \quad (4)$$

$$g_c = \begin{cases} a_1 g_c(x, y), & g_c(x, y) < \frac{1}{5} g_{b,ave} \\ a_2 g_c(x, y), & \frac{1}{5} g_{b,ave} \leq g_c(x, y) \leq \frac{2}{5} g_{b,ave} \\ a_3 g_c(x, y), & \frac{2}{5} g_{b,ave} \leq g_c(x, y) \leq \frac{3}{5} g_{b,ave} \\ a_4 g_c(x, y), & \frac{3}{5} g_{b,ave} \leq g_c(x, y) \leq \frac{4}{5} g_{b,ave} \\ a_5 g_c(x, y), & g_c(x, y) > \frac{4}{5} g_{b,ave} \end{cases} \quad (5)$$

其中, $g_{b,ave}$ 表示当前 block 内 cell 幅值的均值, $g_{c,sum}$ 表示当前 cell 内梯度值, n 为 block 内 cell 的个数. 式 (5) 为 block 内 cell 的幅值具体调整方式, 其中 $g_c(x, y)$ 表示当前 block 内 (x, y) 位置 cell 的梯度值; a_1, a_2, a_3, a_4 和 a_5 为调整系数, 经实验分析后在本文取 $a_1 = 0.6, a_2 = 0.8, a_3 = 1, a_4 = 1.2$ 和 $a_5 = 1.4$.

提取经过预处理后图像的 HOG 特征, 从增强特征对比度角度对 HOG 特征进行改进, 提出了增强的 HOG (Enhanced HOG, EHOg) 特征. 增强 HOG 特征对比度

的方法是在提取 HOG 特征后, 按照式 (4) 计算出当前 block 的梯度值均值. 再按照式 (5) 对该 block 内每个 cell 的梯度值做出调整. 之后串连调整过的 cell 特征描述子得到 block 的特征描述子. 最后将所有 block 的特征描述子串连得到 EHOg 特征.

经过特征增强处理, 梯度直方图中梯度值小的 bin 得到进一步压缩, 梯度值大的 bin 得到进一步放大, 直方图的方差增大, 对比度显著增强. 为进一步分析 EHOg 特征相较 HOG 特征的变化, 本文提取 INRIA 数据集中一张图片的 HOG 特征和 EHOg 特征进行可视化分析, 如图 3 所示.

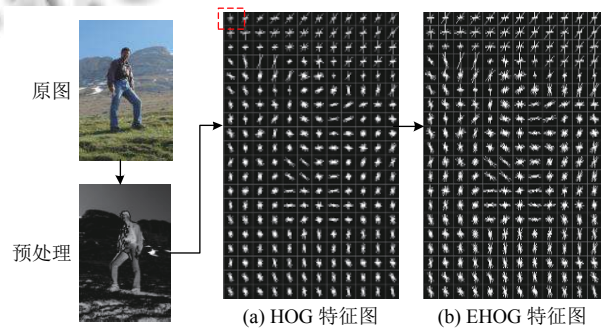


图 3 HOG 和 EHOg 特征可视化比较图

图 3 中, 每个红色虚线区域代表一个 cell 单元, 由图 3(a) 与图 3(b) 比较可以很直观地看出经过特征增强处理, cell 单元内各方向梯度值在保持原有趋势前提下, 对比度显著增大, cell 单元特征描述子的局部描述能力得到增强. 图 4 从数据角度很好的反应了这一点, HOG 特征整体变化趋势与 EHOg 特征保持一致但对比度得到增强, 如维度 17 位置对应特征分别为 0.005 和 0.008, 仅相差 0.003; 最大特征值与最小特征值之差由 0.335 提高到 0.470, 较未进行特征增强之前增长了 40.3%.

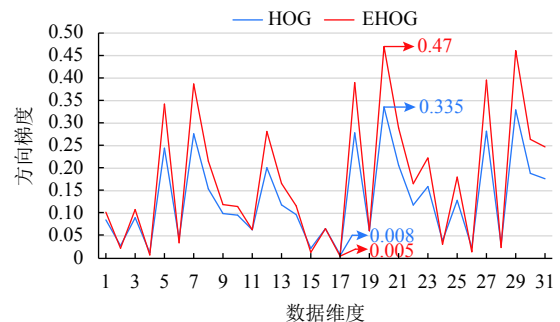


图 4 HOG 特征和 EHOg 特征值图

算法 1. EHOH 特征获取

```

Input: imgGray: 原始图像对应的灰度图
Output: fOut: EHOH 特征
hist ← imgGray 中出现像素灰度对应的频数
设置一个用于保存 0 到 255 各灰度值对应的增强值的对象 dist
for i ← 0 to 255 do
dist[i] ← 根据 hist 利用式 (3) 求出增强值
end
rows ← imgGray 的高度
cols ← imgGray 的宽度
设置一个用于保存结果增强后结果的与 imgGray 大小相同的单通道
图像 imgEnhance
for i ← 0 to cols - 1 do
for j ← 0 to rows - 1 do
gray ← imgGray 位于 (i,j) 位置的灰度值
将 imgEnhance 中位于 (i,j) 位置的像素设置为 dist[gray] 的值
end
end
对 imgEnhance 进行归一化处理
szCell ← 用于提取 HOG 的 cell 单元大小
szBlock ← 用于提取 HOG 的滑动窗口大小, cell 的数量
szStep ← 用于提取 HOG 的步长, 为 szCell 的整数倍
f ← 根据 szBlock、szStep 和 szCell 在 imgEnhance 上滑动计算收集
所有 HOG 特征 numWindow ← (rows - szBlock×szCell)×(cols -
szBlock×szCell)/(szStep×szStep) + 1
for i ← 0 to numWindow - 1 do
gCellSum ← 根据 f 中的信息计算当前窗口内 cell 梯度总和
gBlock ← 根据 gCellSum 利用式 (4) 计算出当前窗口内 cell 梯度均值
for j ← 0 to szBlock×szBlock - 1 do
gCellCur ← 格局 f 中的信息获取当前位置 cell 的梯度值
gCellAdp ← 根据 gBlock 和 gCellCur 利用式 (5) 计算出当前 cell 的
调整值
将 f 中对应的 cell 位置的设置值为 gCellAdp
end
end
fOut ← f

```

3 实验分析

本文实验的相关参数设置如下: cell 大小为像素, block 大小为 cell, 滑动窗口采用固定大小为, 扫描步长为 8 个像素. 实验采用 INRIA 数据集进行实验分析, 部分示例图像如图 5 所示.

INRIA 数据集中行人姿态多样, 背景复杂多变, 在行人检测领域该数据集认可度较高. 训练样本中正样本数量为 2416 个, 负样本数量为 1218 个; 测试样本中正样本数量为 1126 个, 负样本数量为 453 个. 硬件的运行环境设置为: Windows 7 操作系统, Intel(R) Core(TM) i7-4770 CPU, 8 GB 内存. 软件平台为 Visual Studio 2015.



(a) 训练图像



(b) 测试图像

图 5 INRIA 数据集的部分示例图像

为验证本文提出的 EHOH 特征相较其他特征的优越性, 将 EHOH 特征和其他几种常用特征结合 XGBoost 分类器^[20,21]进行实验对比, 实验结果见表 1.

表 1 不同特征结合 XGBoost 分类器的比较

特征种类	识别率(%)	检测时间(ms)
Haar	77.40	53.27
LBP	83.64	49.01
HOG	90.78	25.61
EHOH	95.49	30.85

由表 1 可知 HOG 特征和本文提出 EHOH 在识别率和检测时间上明显优于其他特征. 相较 HOG 特征, 使用 Haar 和 LBP 特征训练所得的分类器识别率相对较低, 分别只有 77.40%, 83.64%. 使用 EHOH 特征的识别率为 95.49%, 较 Haar 和 LBP 特征分别提高了 18.09% 和 11.85%, 较传统 HOG 特征提高了 4.71%. 表 1 中最后一列是不同特征在与 XGBoost 分类器结合时所需的检测时间, 其中 HOG 特征所需最少, 为 25.61 ms, 而本文提出的 EHOH 特征所需检测时间为 30.85 ms 与其接近, 但是识别率有明显的提升.

为了进一步检测 XGBoost 分类器较其他分类器的

优越性, 本文将 EHOH 特征结合 SVM 分类器进行实验对比, 实验结果见表 2 和图 6。

表 2 不同特征结合 SVM 分类器检测时间

特征种类	检测时间(ms)
Haar	74.63
LBP	69.51
HOG	57.81
EHOH	64.31

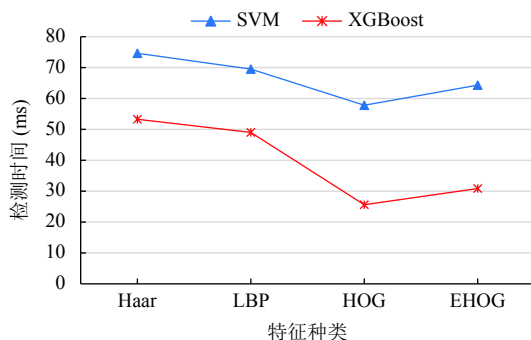


图 6 不同特征在 SVM 和 XGBoost 分类器上检测时间的比较

由图 6 可以看出, 不同特征结合 XGBoost 分类器普遍比 SVM 分类器所需的检测时间要少, 这是因为 XGBoost 分类器的多线程并行操作, 有效地降低了行人检测时间。由表 2 可知, HOG 特征在结合 SVM 分类器时所需的检测时间, 远远少于 Haar 特征和 LBP 特征, 而 EHOH 特征较 HOG 特征检测时间略有增加, 是因为提取 EHOH 特征之前对原始图像进行了预处理增加了时间开销, 但较传统 SVM 分类器, XGBoost 分类器很明显的降低了检测时间。

4 结束语

行人检测具有极广泛的应用: 智能辅助驾驶, 智能监控, 行人分析以及智能机器人等领域。从 2005 年以来行人检测进入了一个快速的发展阶段, 但是也存在很多问题还有待解决, 主要还是在性能和速度方面还不能达到一个权衡。近年, 以谷歌为首的自动驾驶技术的研发正如火如荼地进行, 这也迫切需要能对行人进行快速有效的检测, 以保证自动驾驶期间对行人的安全不会产生威胁。

本文从预处理原始图像和增强特征对比度两方面对其做出改进, 提出了增强型 HOG, 并基于 EHOH 特征结合 XGBoost 分类器的行人检测方法。实验结果证

明, 针对不同特征 (Haar、LBP、HOG、EHOH), 本文提出的方法在识别率和检测速度方面有明显优越性。

本文下一步研究方向是, 针对遮挡问题改进检测算法, 在复杂环境下进一步提高检测精度提升检测速度。

参考文献

- 1 苏松志, 李绍滋, 陈淑媛, 等. 行人检测技术综述. 电子学报, 2012, 40(4): 814–820. [doi: 10.3969/j.issn.0372-2112.2012.04.031]
- 2 宋婉娟, 张剑. 融合回归网络和多尺度特征表示的实时行人检测. 电子测量与仪器学报, 2018, 32(7): 15–20.
- 3 李俊杰, 刘成林, 朱明. 基于多任务 CNN 的监控视频中异常行人快速检测. 计算机系统应用, 2018, 27(11): 78–83. [doi: 10.15888/j.cnki.csa.006607]
- 4 Dalal N, Triggs B. Histograms of oriented gradients for human detection. Proceedings of 2015 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05). San Diego, CA, USA. 2005. 886–893.
- 5 Maldonado S, López J. Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification. Applied Soft Computing, 2018, 67: 94–105. [doi: 10.1016/j.asoc.2018.02.051]
- 6 陈锐, 王敏, 陈肖. 基于 PCA 降维的 HOG 与 LBP 融合的行人检测. 信息技术, 2015, (2): 101–105.
- 7 Ait-Sahalia Y, Xiu DC. Principal component analysis of high-frequency data. Journal of the American Statistical Association, 2019, 114(525): 287–303. [doi: 10.1080/01621459.2017.1401542]
- 8 Xiao L, Zhang YJ, Wang Q, et al. Improved local texture features for pedestrian detection. Proceedings of the IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). Chongqing, China. 2018. 60–65.
- 9 Zhang PB, Yang ZX. A novel AdaBoost framework with robust threshold and structural optimization. IEEE Transactions on Cybernetics, 2018, 48(1): 64–76. [doi: 10.1109/TCYB.2016.2623900]
- 10 Walk S, Majer N, Schindler K, et al. New features and insights for pedestrian detection. Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, CA, USA. 2010. 1030–1037.
- 11 李盈盈, 李菲菲, 陈虬. 基于改进 HOG 共生概率特征的行人检测算法. 电子科技, 2018, 31(9): 4–8.

- 12 Chen PY, Huang CC, Lien CY, *et al.* An efficient hardware implementation of HOG feature extraction for human detection. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(2): 656–662. [doi: [10.1109/TITS.2013.2284666](https://doi.org/10.1109/TITS.2013.2284666)]
- 13 龚露鸣, 徐美华, 刘冬军. 基于混合高斯和 HOG + SVM 的行人检测模型. *上海大学学报 (自然科学版)*, 2018, 24(3): 341–351.
- 14 陈恩加, 唐向宏, 傅博文. Faster R-CNN 行人检测与再识别为一体的行人检索算法. *计算机辅助设计与图形学学报*, 2019, 31(2): 332–339.
- 15 Tian YL, Luo P, Wang XG, *et al.* Deep learning strong parts for pedestrian detection. *Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile. 2015. 1904–1912.
- 16 Ouyang WL, Zhou H, Li HS, *et al.* Jointly learning deep features, deformable parts, occlusion and classification for pedestrian detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(8): 1874–1887. [doi: [10.1109/TPAMI.2017.2738645](https://doi.org/10.1109/TPAMI.2017.2738645)]
- 17 王琳, 卫晨, 李伟山, 等. 结合金字塔池化模块的 YOLOv2 的井下行人检测. *计算机工程与应用*, 2019, 55(3): 133–139. [doi: [10.3778/j.issn.1002-8331.1710-0236](https://doi.org/10.3778/j.issn.1002-8331.1710-0236)]
- 18 杨金锋, 师一华, 熊兴隆. 光照处理中改进的 Gamma 校正方法. *中国民航学院学报*, 2006, 24(4): 39–42.
- 19 储清翠, 王华彬, 陶亮. 图像的局部自适应 Gamma 校正. *计算机工程与应用*, 2015, 51(7): 189–193, 208. [doi: [10.3778/j.issn.1002-8331.1305-0193](https://doi.org/10.3778/j.issn.1002-8331.1305-0193)]
- 20 Chen TQ, Guestrin C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, CA, USA. 2016. 785–794.
- 21 Chen WB, Fu K, Zuo JW, *et al.* Radar emitter classification for large data set based on weighted-xgboost. *IET Radar, Sonar & Navigation*, 2017, 11(8): 1203–1207.