

# 夜间复杂场景下红外图像行人检测算法研究

赵双, 陈树越, 王巧月

(常州大学 信息科学与工程学院, 江苏 常州 213164)

**摘要:** 针对夜间红外图像中行人与背景灰度差异小且存在遮挡等问题, 提出了一种夜间复杂场景下的红外行人检测算法。首先利用行人语义融合方法生成对目标全覆盖的显著图, 与原图融合得到感兴趣区域, 然后构造基于改进的方向梯度直方图特征的两分支分类器, 同时提出一种遮挡判别算法, 根据分类器模糊分数判断是否遮挡, 设计一种头部模板实现最终的行人检测。在 LSI 远红外行人数据集和自主采集的冬、夏季节夜间行人数据集上进行实验, 结果表明: 在不同环境下, 所提出的方法均可快速鲁棒地检测出行人, 可较显著地降低漏检率, 检测率可达到 94.20%。

**关键词:** 红外图像; 行人检测; 显著性; 复杂情况; 方向梯度直方图特征

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1001-8891(2021)06-0575-08

## Infrared Pedestrian Detection in Complex Night Scenes

ZHAO Shuang, CHEN Shuyue, WANG Qiaoyue

(School of Information Science and Engineering, Changzhou University, Changzhou 213164, China)

**Abstract:** An infrared pedestrian detection algorithm is proposed to solve the problem of small differences between pedestrians and backgrounds in gray scale images and the occurrence of occlusion in infrared images at night. First, a significant graph with the full coverage of the target is generated by the pedestrian semantic fusion method, and the region of interest is obtained by combining it with the original graph. Then, a two-branch classifier based on the improved histogram of the gradient feature is constructed. The fuzzy score of the classifier is used to determine the occurrence of occlusion and call the head template for the final detection. Experiments based on the LSI far infrared pedestrian dataset and independent datasets of pedestrians captured at night in winter and summer prove that the proposed method is robust and quick in detecting pedestrians under different environments. It can significantly reduce the rate of missed detection and realize a detection rate of 94.20%.

**Key words:** infrared image, pedestrian detection, saliency, complex scenes, HOG feature

## 0 引言

行人检测是计算机视觉领域极具挑战性的问题, 广泛应用在公共安全、智能监控、自动驾驶等领域中<sup>[1-2]</sup>。与传统的基于可见光的行人检测任务相比, 远红外图像不受光照的影响, 可以检测物体的热分布, 从而在夜间对物体进行空间显示。热成像受温度影响大, 环境温度升高会带来灰度值反转问题, 当行人与环境相接时, 轮廓边缘无法保持完整, 导致感兴趣区域提取失败进而无法检出行人、路灯、车辆向阳面等灰度值高的目标易产生误检<sup>[3]</sup>。与此同时, 还

要应对行人个体和姿态、视角及尺度的差异问题, 复杂的背景使行人极易出现遮挡情况导致误判, 现阶段的研究多关注完整体态的行人检测, 而忽略了遮挡的情况。

传统的行人检测常采用分割及连通域分析, 模板匹配等方法, Dalal<sup>[4]</sup>在 2005 年提出梯度方向直方图特征 (histogram of gradient, HOG), 率先使用梯度描述行人, 利用滑动窗口遍历图像, 得到了很高的检测率, 但计算资源消耗大。随后描述纹理的局部二值模式 (local binary patterns, LBP) 特征、描述线状、形状、边缘的 Haar-like 特征, Shapelet 特征和 Edgelet

收稿日期: 2019-07-02; 修订日期: 2019-11-25.

作者简介: 赵双 (1995-), 女, 硕士, 主要研究方向: 模式识别 (行人检测)。E-mail: zhss1030@163.com.

通信作者: 陈树越 (1963-), 男, 教授, 主要研究方向: 计算机视觉、图像处理。E-mail: cshyue2000@163.com.

基金项目: 江苏省研究生科研创新基金项目 (KYCX19\_1770)。

特征等接连被提出并应用到行人检测中。很多学者对 HOG 特征做出了改进, 比如将金字塔思想应用于尺度缩放的 PHOG 特征<sup>[5]</sup>, 结合边缘信息的 HEOG 特征<sup>[6]</sup>, 或者以 HOG 特征为基础进行特征融合, 如 HOG-LBP<sup>[7]</sup>, HOGI 特征<sup>[8]</sup>等。

常见的感兴趣区域 (region of interest, ROI) 提取方法是通过阈值分割, 这种方法受图像对比度影响大, 效果较差, 基于视觉显著性进行 ROI 提取可以最大程度模拟司机的人眼注视, 将有限的计算资源分配给容易引起观察者注意的区域, 达到辅助驾驶的目的。显著性检测分为两类: 自下而上基于数据驱动的显著性区域凸显, 和自上而下任务驱动的目标凸显。

Koch<sup>[9]</sup>引入了显著图 (saliency map) 的概念来表示复杂场景的突出位置, 他使用“赢者全取” (winner-take-all) 的神经网络来选择图像中最显著的位置, 在此基础上, 许多由视觉注意机制启发的显著性模型被提出, 如应用广泛的 FT 模型<sup>[10]</sup>, SR 模型<sup>[11]</sup>等。

与相关研究相比, 本文更关注在黑暗复杂环境下探测行人, 考虑了遮挡、尺度不一等情况, 利用行人先验语义进行 ROI 提取, 构造改进的 HOG 二分支分类器进行行人检测, 同时利用遮挡判别算法, 设计头部检测算子对被遮挡行人进行二次识别, 实验证明本文方法检测效果良好, 克服了行人的多变性, 背景的复杂性。

## 1 红外行人候选区域提取

### 1.1 系统检测框架

文中提出的检测方法由 3 个模块组成, 包括 ROI 提取、基于 HOG 特征的二分支分类和遮挡特征重识别, 整体流程图如图 1 所示。

### 1.2 光谱残差 (SR, spectrum residual) 模型

ROI 是图像中最吸引视觉的显著关键信息, 提取 ROI 代替原始图像, 可以减少大量无用的背景和噪声点的影响。本文在光谱残差模型的基础上融入红外行人的高层语义特征, 得到更准确的感兴趣区域。Hou 等人<sup>[11]</sup>认为不同的图像数据在对数尺度上有相似分布趋势, 因此在不同的 log 谱中只需要关注差异性部分就可以得到显著性区域。首先对输入图像  $H(x)$  进行二维离散傅里叶变换, 将其从空间域转化到频率域, 光谱残差  $R(f)$  定义为:

$$R(f) = |\lg[A(f)]| - h_n f \cdot |\lg[A(f)]| \quad (1)$$

式中:  $A(f)$  为实部;  $h_n(f)$  为局部平均滤波器。

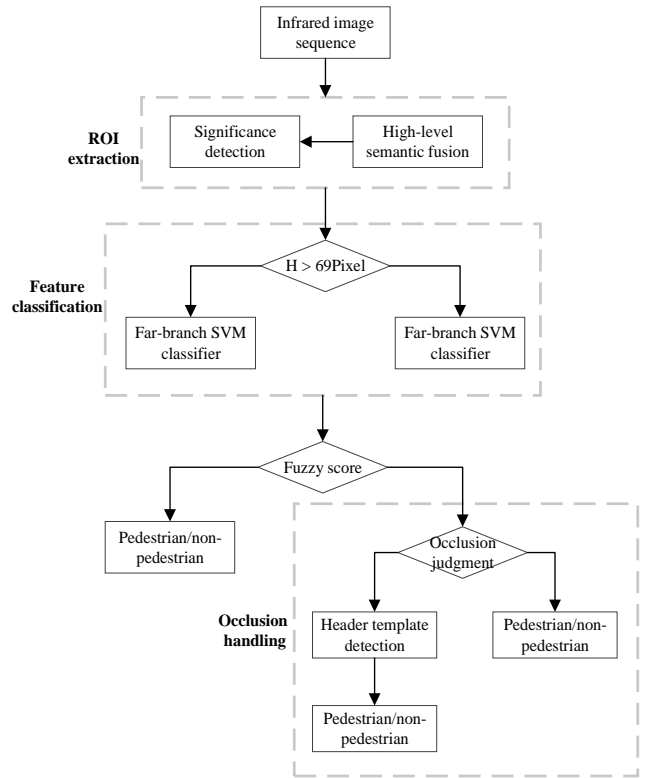


图 1 行人检测流程图

Fig.1 Pedestrian detection flow chart

利用式(2)对谱残差  $R(f)$  和相位  $P(f)$  进行二维离散傅里叶逆变换得到全局特征显著图  $S(x)$ , 所得的显著性图与人眼感知具有一致性。

$$S(x) = |F^{-1} \{ \exp[R(f) + P(f)] \}|^2 \quad (2)$$

### 1.3 行人高层语义特征融合

在利用 SR 分析得到显著性图的基础上, 本文提出了针对性的红外行人的高层语义特征作为先验知识融合, 相当于在进行显著性检测时增加了对图像的认知功能, 提高分析准确度。行人的高层语义特征是指在待检测的图片中, 一个区域或者一个像素点被行人目标包含的可能性的, 求解这个可能性作为依托来求图像的目标性, 可以获得图像中所有行人目标的大概分布情况, 对图像的显著性检测有语义上的指导。

首先对输入的图像进行块分割, 本文选择 Alexe<sup>[12]</sup>的方法对原始图像进行分割, 然后对每个分割得到的窗口进行评分, 量化每个窗口包含行人目标的可能性。在红外图像中, 行人的亮度普遍高于周边环境, 头部区域亮度最高, 低频信号强, 边缘纹理比内部纹理明显, 所以本文以亮度对比度和边缘信息作为先验语义, 融入显著性检测中。

1) 行人亮度对比度 (brightness contrast, BC)。

亮度特征是红外图像最重要的特征之一, 如果一个包围框完整包含行人部分, 则得到较高的分数, 如果窗口只包含部分目标时分数降低。如图 2(a)所示, 实线框代表目标窗口  $m$ , 在计算  $m$  的亮度对比度时, 先将  $m$  沿 4 个方向延伸一定尺度  $\theta$  得到一个矩形外环  $S$ , 如图 2 中虚线框所示, 用外环框减去目标框, 得到  $m$  所对应的背景区域, 此时有:

$$\frac{|S(m, \theta)|}{|m|} = \theta^2 - 1 \quad (3)$$

计算窗口  $m$  和背景的亮度直方图的卡方距离来定义窗口的亮度对比度:

$$BC(m, \theta) = \chi^2[h(x), h(S(m, \theta))] \quad (4)$$

式中:  $BC(m, \theta)$  表示亮度对比度特征值;  $\chi^2(\cdot)$  表示卡方距离;  $h(\cdot)$  表示窗口的亮度直方图。

2) 边缘线密度 (edge density, ED)。相比可见光图像, 红外图像 (如图 2(b)所示) 考虑的是物体的温度信息, 对杂乱的背景抑制效果良好, 当温度差距较大时, 行人边缘线明显, 当温差小时, 效果不显著。行人边缘线密度指的是窗口边界附近的边缘密度, 通常情况下目标边界边缘的数量随着目标边界周长增大而增大, 如图 2(c)所示。首先将得到的目标窗口  $m$  沿 4 个方向向内收缩  $\lambda$  个尺度, 得到:

$$\frac{|I(m, \lambda)|}{|m|} = \lambda^2 - 1 \quad (5)$$

式中:  $I(m, \lambda)$  是  $m$  的矩形内环。则 ED 计算公式如下:

$$ED(m, \lambda) = \frac{\sum_{p \in I(m, \lambda)} I_{ED}(p)}{\text{Len}(I(m, \lambda))} \quad (6)$$

实验发现 Canny 算子的边缘线检测效果最佳,  $I_{ED}(p) \in \{0, 1\}$  是利用 Canny 算子进行边缘检测后得到的二值化图像,  $\text{Len}(\cdot)$  表示矩形内环的周长。

对图像的每一个窗口都从这两个角度进行打分, 按式(7)计算图像中每个像素点的目标性:



(a) 亮度对比度 (b) 红外图像 (c) 边缘线密度  
(a) Brightness contrast (b) Infrared image (c) Edge density

图 2 目标性计算示意图

Fig. 2 Schematic diagram of target calculation

$$P(x) = \frac{\sum_{m \in M_x} S(m)}{\text{Max}_p} \quad (7)$$

式中:  $M_x$  是包含像素点  $x$  的所有窗口;  $S(m)$  表示对行人亮度对比度和边缘线密度特征打分得到的分数;  $\text{Max}_p$  是归一化因子, 将得到的目标性得分归一化到  $[0, 1]$  区间。按照公式  $Q(x) = \alpha * P(x) + (1 - \alpha) * S(x)$  进行线性融合得到最终的显著性估计, 对得到的二值化结果图像寻找它的最小外包矩形, 定位我们所需要的感兴趣区域, 减少行人模型的搜索范围, 同时排除复杂背景的干扰, 减少误检的可能, 本文参数  $\alpha = 0.6$ 。

在实验过程中, 窗口划分总个数  $W_m$  的选值决定了“行人”区域的标定时间和准确性, 如果  $W_m$  值太小, 得到的目标性区域结果存在较大偏差, 如果  $W_m$  值太大, 则显著性检测的计算时间过长。

## 2 快速行人特征提取与分类

### 2.1 基于改进的 HOG 特征的二分支构造

HOG 特征可以很好地表征图像梯度的方向分布, 具体计算步骤如下:

1) 输入大小为  $64 \times 128$  的检测窗口, 每 64 个像素作为一个细胞单元, 每 4 个细胞单元作为一个区块, 步长设置为 8 个像素, 水平方向得到 7 个扫描窗口, 垂直方向得到 15 个扫描窗口, 如图 3(a)所示。定义每个细胞的组距为  $2\pi/9$ , 统计落在 9 个不同方向区间内的梯度大小, 如图 3(b)所示。利用式(8)、(9)计算每个像素的梯度幅值  $m(x, y)$  和方向  $\theta(x, y)$ :

$$m(x, y) = \sqrt{[I(x, y+1) - I(x, y-1)]^2 + [I(x, y+1) - I(x, y-1)]^2} \quad (8)$$

$$\theta(x, y) = \arctan \left[ \frac{I(x, y+1) - I(x, y-1)}{I(x, y+1) - I(x, y-1)} \right] \quad (9)$$

式中:  $I(x, y)$  是图像在点  $(x, y)$  处的像素值, Dalal<sup>[4]</sup>验证了利用最简单的一阶梯度算子  $[1, 0, -1]$  进行水平和垂直方向的梯度计算效果最佳。

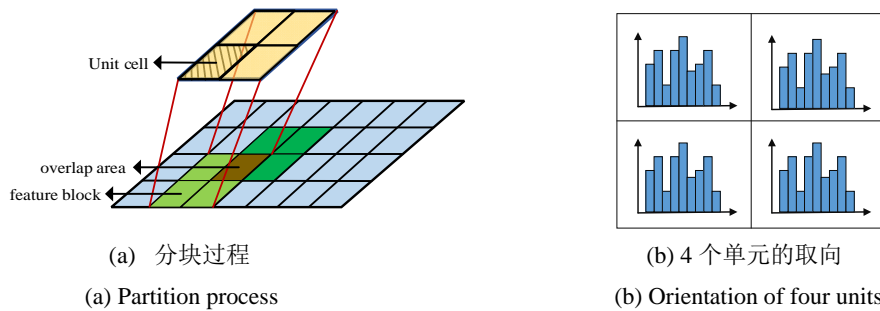


图3 HOG 特征计算过程

Fig.3 HOG feature calculation process

2) 利用 SVM 分类器构建超平面进行分类, 对于给定的  $\{x_i, y_i\}$ ,  $i=1, \dots, N$ ,  $y_i \in \{1, -1\}$ , 分类器通过寻找最优的超平面  $\mathbf{W}^T X + b = 0$  实现行人的分类识别。

传统的 HOG 特征提取所需的计算量非常大, 分辨率为  $64 \times 128$  的检测窗口需要计算 105 个块的梯度直方图特征, 得到 3780 维特征向量, 同时多尺度缩放会导致相邻检测窗口间存在大量的重复计算区域, 针对此种情况, 本文提出了一种改进的 HOG 特征并构造了基于 ROI 高度的二分支分类结构 (TBHOG, two-branches HOG), 具体步骤如下:

① 将得到的感兴趣区域扩充至扫描步长的整数倍, 用邻域像素的灰度平均值作为扩充像素的灰度值;

② 根据 ROI 的高度  $H$  进行分类,  $H$  定义为最小外接矩形垂直方向上的像素数, 统计所有得到的 ROI 区域, 平均高度为 69.1 pixel, 所以本文选择 69 像素作为分水岭, 高度小于 69 像素作为近景目标, 高度大于 69 像素作为远景目标, 分别训练 SVM 分类器, 具体参数如表 1 所示。

③ 将得到的感兴趣区域分别缩放至不同的窗口大小, 输入离线训练的 SVM 分类器进行判别。

表 1 两分支的 SVM 分类器训练参数

Table 1 SVM classifier training parameters of the two branches

	Window size	Block size	Cell size	Step	Bin	Feature dimension
Near target	$48 \times 96$	$16 \times 16$	$8 \times 8$	8	9	1980
Distant target	$24 \times 48$	$8 \times 8$	$4 \times 4$	4	9	1980

## 2.2 遮挡处理

上述行人检测方法的出发点是将行人作为一个完整的整体设计, 在 LSI 远红外 (FIR) 数据集的实验发现, 对比传统的 HOG 方法, 该方法的准确率有

所提升, 但是对本文自行拍摄的数据集来说, 效果不够理想, 因为拍摄的视频包含更多的遮挡情况, 例如行人被树木等背景遮挡或行人互相遮挡。

研究发现, 如果行人发生遮挡, 块的 HOG 特征通过 SVM 分类器得到负的内积, 每个窗口的线性 SVM 得分是 HOG 特征与向量  $\mathbf{w}$  的内积, 向量  $\mathbf{w}$  是学习到的所有支持向量的加权和<sup>[13]</sup>。本文选择利用整体分类器对得到的特征向量进行评价, 如果整体分类器得到的置信度属于一个模糊范围, 则判断是否发生遮挡情况<sup>[14]</sup>, 如果确定发生遮挡, 则利用设计的头部模板进行行人头部的检测, 对是否包含行人进行确认。具体步骤如下:

1) 将待测窗口按图 4 分块,  $C_i$  是每个块得到的特征值, 维数为 36, HOG 特征计算过程中得到 55 个块, 定义窗口特征描述子  $x = (C_1, C_2, \dots, C_{55})^T$ , 通过训练好的 Linear SVM 分类器, 根据式(10)计算返回值  $H(x)$ :

$$H(x) = \varepsilon + \mathbf{w}^T \cdot x \quad (10)$$



图 4 检测窗口分块

Fig.4 Detection window chunking

理论上  $H(x)$  值越高, 说明窗口中有行人的可能性越高, 反之亦然, 如果分类器的响应接近于 0, 就认为分类是模糊的。本文定义模糊区间为  $(-1.5, 0.5)$ , 认为可能存在遮挡。

2) 对于每个  $C_i$ ,  $i \in \{1, 2, \dots, 55\}$ , 定义一个局部分类器  $h(i) = \varepsilon_i + \mathbf{w}_i^T \cdot C_i$ , 其中  $\mathbf{w}_i$  是支持向量展开后所对

应的向量, 均为 36 维,  $\varepsilon_i$  为每个块的检测误差。

3) 通过对局部响应  $h(i)$  进行阈值化得到离散的标签  $t_i$ ,  $t_i=1$  表示是  $C_i$  行人的一部分,  $t_i=-1$  表示  $C_i$  是被遮挡或是背景的一部分, 得到一个二值映射  $(t_1, t_2, \dots, t_{55})$ 。

4) 利用均值漂移分割算法进一步得到块标签  $(t'_1, t'_2, \dots, t'_{55})$ , 本文将均值漂移权重设置  $|h(C_i)|$ , 将具有多数负响应的窗口作为遮挡窗口, 得到遮挡确认。

5) 无论行人是否遮挡或出现不同的朝向、姿势, 行人的头部轮廓变化基本可以保持不变, 头肩部呈现出类似于“ $\Omega$ ”的形状<sup>[15]</sup>, 如图 5 所示, 我们设计了  $16 \times 16$  的头部模板进行分类识别, 使用离线 SVM 分类器训练。



图 5 头部检测模型

Fig.5 Head detection model

### 3 系统设计与结果评估

#### 3.1 数据集选择

实验运行环境为 Intel(R) Core(TM) i3-4170 CPU, 平台为 MATLAB R2018b 和 Python OpenCV 库, 实验使用 3 个数据集: LSI 远红外行人数据集 (<https://e-archivo.uc3m.es/handle/10016/17370>) 和利用 FLIR E40 热像仪采集的冬季 (Test-W)、夏季 (Test-S) 道路/行人 (Test-L) 视频, 基本信息如表 2 所示。

表 2 测试数据的基本信息

Table 2 Basic information of test data

Dataset	Size/frame	Date	Location
Test-L	2000	-	-
Test-W1	1578	2018/1/2	Street
Test-W2	979	2018/1/3	Campus
Test-S1	1123	2019/5/29	Street
Test-S2	1089	2019/5/30	Campus

LSI 数据集来自在室外城市场景中驾驶的车辆, 包含用于训练的 53598 张红外图像和用于测试的 27994 张图像。自主数据采集速率为 31 帧/s, 分辨率为  $320 \times 240$  像素。数据采集的时间均为晚上 22:00 左右, 涵盖了温度变化、遮挡、尺度变化等一系列夜间行人检测常见的问题, 示例如图 6。每隔 3 帧标注一次, 保证了实验的真实有效性。

#### 3.2 评价标准

红外图像中行人的检测本质上是一个二分类问题, 检测率 (accuracy, ACC) 是最常见的指标, 即:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

其中参数混淆矩阵如表 3 所示。

定义正确的检测需符合条件:

$$\alpha_0 = \frac{\text{area}(\text{BB}_{\text{ROI}} \cap \text{BB}_{\text{GT}})}{\text{area}(\text{BB}_{\text{ROI}} \cup \text{BB}_{\text{GT}})} > 0.6 \quad (12)$$

即预测的边界框  $\text{BB}_{\text{ROI}}$  和真实边界框  $\text{BB}_{\text{GT}}$  之间的重叠率  $\alpha_0$  必须超过 60%, 否则就认为是错误的检测结果 (虚警)。



图 6 样本示例: (a)(b)(c) 冬季拍摄; (d)(e)(f) 夏季拍摄

Fig.6 Sample diagram: (a)(b)(c) Shooting in winter; (d)(e)(f) Shooting in summer

表3 参数混淆矩阵

Table 3 Parameter confusion matrix

Truth	Predicted result	
	Positive example	Negative example
Positive example	TP	FN
Negative example	FP	TN

其中, TP (true positive) 表示指检测器正确预测的行人窗口, FP (false positive) 指将非行人预测为行人窗口的数目, TN (true negative) 指检测器正确分类的非行人窗口, FN (false negative) 指行人窗口发生漏检的数目。

考虑行人检测问题的实时性, 本文将每张图片的平均检测时间也作为评价标准, 即:

$$t = \frac{T}{N} \quad (13)$$

式中:  $T$  是检测耗费时间;  $N$  是图片张数。

### 3.3 参数选择

随机选取 500 幅 LSI 远红外数据集的测试图片, 手动标记行人区域, 统计不同目标性检测窗口数  $W_m$  下的行人区域数, 目标性检测示意图如图 7 所示。

随着  $W_m$  的增大, 检测结果如表 4 所示。

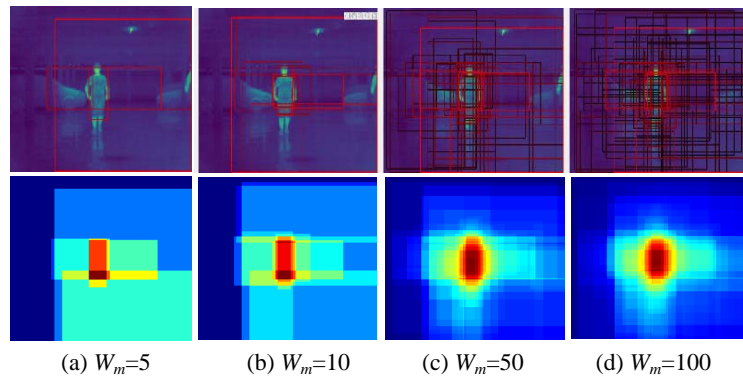


图7 目标性检测

Fig.7 Schematic diagram of target detection

表4  $W_m$  标记结果

Table 4  $W_m$  marking results

$W_m$	Picture number	Pedestrian number	Pedestrian mark	Labeling rate/%	Mark time/s
5	500	774	630	81.4	5.94
10	500	774	636	82.2	8.39
50	500	774	671	86.7	12.70
100	500	774	928	94.1	17.83
150	500	774	751	97.0	23.29
200	500	774	752	97.2	29.01
350	500	774	755	97.6	44.00
500	500	774	755	97.6	56.38

从表 4 可以看出, 随着  $W_m$  的增大, 显著性检测框与真实标注框的重合度也在上升,  $W_m=150$  时标记率可以达到 97%, 随后增速放缓但检测时间持续增加, 所以本文选择目标性窗口数为 150。

#### 3.3.1 与基本算法的检测效果对比

为了验证本文算法的有效性, 本文在 LSI 远红外数据集上进行了如下对比实验, 以经典的 HOG 特征作为基准, 分别测试了二分支分类结构 (TBHOG)、是否提取感兴趣区域、是否采用遮挡处理算法, 检测结果如表 5 所示。

检测结果表明与 HOG 特征相比, 本文方法准确率提高了 5.5%, 检测时间缩短了 23.3%, 可见本文方法对夜间行人识别有更好的实用性。准确率的提高主要来自于 ROI 提取和遮挡检测算法, 而二分支分类结构显著降低了检测时间。

#### 3.3.2 与其他算法的检测效果对比

在所有数据集上将本文的检测方法与当前主流的红外行人检测方法对比: HOG+SVM 方法, LBP+SVM 方法及融合方法, 聚合通道 (aggregate channel features, ACF) 方法<sup>[16]</sup>, 基于深度学习的 RetinaNet

算法<sup>[17]</sup>, 该算法由 Python 的 image AI 库集成, 是一个开源的目标检测项目。对比示意如图 8 所示, 检测结果如表 6 所示。

表 5 与基本算法检测结果对比

Table 5 Comparison of detection results with the basic algorithm

Algorithm	ACC/%	t/ms
HOG	89.27	236.4
TBHOG	86.19	122.8
TBHOG+ROI extraction	91.44	159.0
TBHOG+Occlusion handling	93.23	168.4
TBHOG+ ROI extraction +occlusion handling	94.20	181.6

传统的特征提取采用滑动窗口法并利用尺度缩放策略进行检测, 计算量大且准确率取决于特征设计的合理性<sup>[18]</sup>, 由表 6 可以得到以下结论:

1) 采用本文提出的算法, 平均检测率比使用单一 HOG 特征提高了 11.1%, 比使用融合特征如 HOG-LBP 和 ACF 均有所提高, 说明了方法的有效性与针对性。

2) 本文所提算法平均检测速度相对于基于 HOG 的行人检测算法提升了 27.6%, 说明了采用 ROI 提取及两分支分类器结构的方法, 相较于传统尺度缩放策略能够极大地加快检测速度, 虽然仍不能满足实时性

表 6 与其他算法检测效果对比

Table 6 Comparison of detection effect with other algorithms

Algorithm	LSI FIR		test-W		test-S	
	ACC/%	t/ms	ACC/%	t/ms	ACC/%	t/ms
HOG	89.27	236.4	80.59	289.7	62.77	274.1
LBP	85.21	202.3	77.90	304.5	59.25	294.5
HOG-LBP	93.64	333.7	84.03	375.8	68.17	373.8
ACF	94.05	1008	85.54	1469	64.28	1742
RetinaNet	90.13	1364	82.21	1730	63.09	2297
Proposed	94.20	181.6	89.17	264.7	73.64	285.3

#### 4 结论

针对夜间行人检测背景复杂, 遮挡严重的情况, 本文在 SR 模型的基础上改进了 ROI 提取方法, 对得到的 ROI 构建基于改进的 HOG 特征的二分支分类器结构, 并提出一种遮挡处理算法, 性能评价表明, 与标准方法相比, 本方法在检测精度和速度上均得到很大提升, 为夜间复杂情况下的行人检测提供了一种新的可能性。但算法仍无法满足实时检测的要求, 且在

的要求, 但相比检测率较高的 ACF 和深度方法, 检测时间大幅缩短。

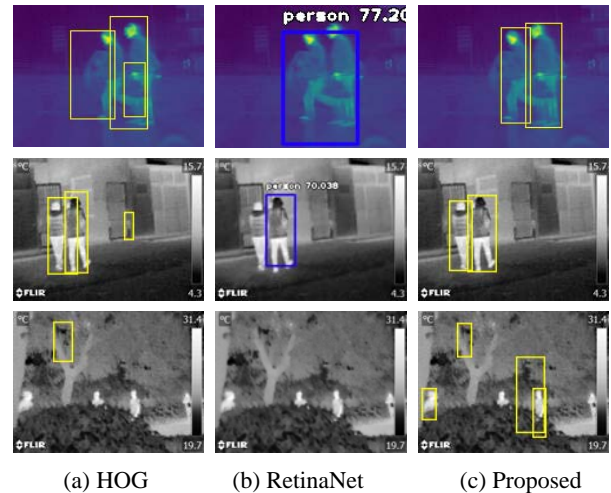


图 8 不同算法检测效果对比

Fig.8 Detection effects of different algorithms

3) 本文方法在 LSI 远红外数据集上检测率比次好的算法提高了 0.15 个百分点, 在 Test-W 数据集上比次好的算法提高了 3.6 个百分点, 在 Test-S 数据集上提高了 5.4 个百分点, 可能是因为实际道路背景复杂, 对行人检测不是很友好, 行人的位置和姿态与开源数据库差异大, 本文提出的遮挡检测算法在复杂情况下有效地降低了漏检。

行人较多的时候效果有限, 下一步计划采用特征性更强的描述子, 并结合深度学习理论, 尝试解决行人的尺度不一、粘连等问题, 进一步提高复杂情况下行人的检测率。

#### 参考文献:

[1] XU Z, ZHUANG J, LIU Q, et al. Nighttime FIR pedestrian detection benchmark dataset for ADAS[C]// *Proceedings of Pattern Recognition and Computer Vision*, 2018: 323-333.

- [2] TAO Y, FU D, SHU P. Pedestrian tracking for infrared image sequence based on trajectory manifold of spatio-temporal slice[J]. *Multimedia Tools and Applications*, 2017, **76**: 11021-11035.
- [3] 刘洋. 基于 LS-DYNA 的汽车正面碰撞计算机模拟仿真[D]. 西安: 西华大学, 2011.  
LIU Yang. Simulation on the Front Impact of Vehicle Based on LS-DYNA[D]. Xi'an: Xihua University, 2011.
- [4] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, **1**(1): 886-893.
- [5] Bosch A, Zisserman A, Munoz X. Representing shape with a spatial pyramid kernel[C]//Acm International Conference on Image & Video Retrieval, 2007: 401-408(doi: <https://doi.org/10.1145/1282280.1282340>).
- [6] Sangeetha D, Deepa P. A low-cost and high-performance architecture for robust human detection using histogram of edge oriented gradients[J]. *Microprocessors and Microsystems*, 2017, **53**: 106-119.
- [7] ZHENG C H, PEI W J, YAN Q, et al. Pedestrian detection based on gradient and texture feature integration[J]. *Neurocomputing*, 2017, **228**: 71-78.
- [8] 朱聪聪, 项志宇. 基于梯度方向和强度直方图的红外行人检测[J]. *计算机工程*, 2014, **40**(12): 195-198, 204.  
ZHU Congcong, XIANG Zhiyu. Infrared pedestrian detection based on histograms of oriented gradients and intensity[J]. *Computer Engineering*, 2014, **40**(12): 195-198, 204.
- [9] Itti L, Koch E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Trans.*, 1998, **20**(11): 1254-1259
- [10] Radhakrishna A, Sheila H, Francisco E, et al. Frequency-tuned salient region detection[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009: (doi: [10.1109/CVPR.2009.5206596](https://doi.org/10.1109/CVPR.2009.5206596)).
- [11] HOU X D, ZHANG L Q. Saliency detection: a spectral residual approach[C]//2007 IEEE Conference on Computer Vision and Pattern Recognition, 2007: (doi: [10.1109/CVPR.2007.383267](https://doi.org/10.1109/CVPR.2007.383267)).
- [12] Alexe B, Deselaers T, Ferrari V. Measuring the objectness of image windows[J]. *IEEE Transactions on Software Engineering*, 2012, **34**(11): 2189-2202.
- [13] WANG X, HAN T X, YAN S. An HOG-LBP human detector with partial occlusion handling[C]//12th International Conference on Computer Vision of IEEE, 2010: (doi: [10.1109/ICCV.2009.5459207](https://doi.org/10.1109/ICCV.2009.5459207)).
- [14] Javier M N, Vazquez D, Lopez A M, et al. Occlusion handling via random subspace classifiers for human detection[J]. *IEEE Transactions on Cybernetics*, 2013, **44**(3): V342-354.
- [15] Broggi A, Bertozzi M, Fascioli A, et al. Shape-based pedestrian detection[C]//IEEE Intelligent Vehicles Symposium, 2000: (doi: [10.1109/IVS.2000.898344](https://doi.org/10.1109/IVS.2000.898344)).
- [16] Brehar R, Vancea C, Nedevschi S. Pedestrian detection in infrared images using aggregated channel features[C]//IEEE International Conference on Intelligent Computer Communication & Processing, 2014: (doi: [10.1109/ICCP.2014.p6936964](https://doi.org/10.1109/ICCP.2014.p6936964)).
- [17] LIN T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, **99**: 2999-3007.
- [18] 车凯, 向郑涛, 陈宇峰, 等. 基于改进 Fast R-CNN 的红外图像行人检测研究[J]. *红外技术*, 2018, **40**(6): 578-584.  
CHE Kai, XIANG Zhengtao, CHEN Yufeng, et al. Research on infrared image pedestrian detection based on improved fast R-CNN[J]. *Infrared Technology*, 2018, **40**(6): 578-584.