文章编号:1674-2974(2020)08-0042-07 DOI:10.16339/j.cnki.hdxbzkb.2020.08.005

一种城市道路场景下行人危险度评估方法

曾令秋,马济森†,韩庆文,叶蕾
(重庆大学 计算机学院,重庆 400044)

摘 要: 针对行人在交通场景对车辆驾驶造成的影响和辅助驾驶需要对行人进行避险的问题,提出一种基于车载单目摄像机的行人危险度评估方法.基于中国城市的特色环境,将行车环境划分为三类:普通道路、人行横道和有辅警道路,对每类场景采用不同的评估方法.采用卷积神经网络,检测视频中道路上的行人、辅警、信号灯和人行道等信息;检测行人关键点并使用多目标跟踪方法,生成骨架姿态时间序列,通过 LSTM(长短时记忆神经网络)分析姿态序列获得行人行为和趋势;最后综合视频信息、行人信息和场景信息,构建行人危险评估模型,实现行人危险度评估.实验结果表明,提出的模型可以有效地评估行人危险度,辅助驾驶员安全行车,场景分类使危险模型评估结果更符合行人实际危险度.

关键词: 行人安全;行人行为分析;辅助驾驶;多场景分析
中图分类号: TP399 **文献标志码:** A

A Pedestrian Hazard Assessment Method in Urban Road Scene

ZENG Lingqiu, MA Jisen†, HAN Qingwen, YE Lei
(College of Computer Sceinec, Chongqing University, Chongqing 400044, China)

Abstract: In view of the influence of pedestrians on vehicle driving in traffic scenes and the problem that the assisted driving needs to avoid the danger for pedestrians, a pedestrian hazard assessment method based on vehicle-mounted monocular cameras is proposed. Based on the characteristic environment of Chinese cities, this paper divided the driving environment into three scenarios: regular roads, crosswalks, auxiliary roads. Different risk assessment methods are used for each type of scenarios. A convolutional neural network is used to detect and identify pedestrians, auxiliary police, signal lights and sidewalks on the road in the video. Then, it detects the key points of the pedestrian and uses the multi-target tracking method to generate the time series of the pedestrian skeleton. Pedestrian behavior and trend are obtained through LSTM (Long-Short Term Memory Neural Network) analysis of posture sequences. Finally, a pedestrian hazard assessment model is built to realize the pedestrian hazard assessment in multi road scene by synthesizing the video information, pedestrian information and scene information. The experimental results show that the scene classification makes the assessment results of the hazard model more consistent with the actual pedestrian hazard. The proposed model can effectively evaluate the pedestrian hazard and assist the drivers to drive safely.

Key words: pedestrian safety; pedestrian behavior analysis; assisted driving; multi scenario analysis

---

* **收稿日期:** 2019-12-19
基金项目: 国家自然科学基金资助项目(61601066),National Natural Science Foundation of China(61601066)
作者简介: 曾令秋(1975—),男,重庆人,重庆大学副教授
† 通讯联系人,E-mail:459819629@qq.com

现代智能交通安全辅助驾驶系统(Safety Driving Assist System,SDAS)应用通信与控制系统技术和计算机视觉原理,可以对感知范围的行人进行分析,辅助驾驶员发现危险情况,有效降低碰撞率,对于车载驾驶系统有着十分重要的作用. 其中的关键环节包括分析感知范围内的环境信息、评估场景危险程度等.

驾驶员在驾驶过程中会对视野中的行人行为的危险程度进行主观判断和分析, 主观判断分析会存在一定错误,容易导致意外发生,这时需要辅助驾驶系统主动采取避让措施.通过观察行人姿势的变化来判断和预测行人可能的行为, 评估汽车与行人之间的碰撞率, 并做出相应的反应. 研究发现,对行人的多种属性综合考虑, 可以提高行人意图分析的准确率.

基于计算机视觉算法对行人行为进行分析处理并判断行人意图是当前研究热点, 学者们在行人意图分析方面,开展了相关研究工作[1]. 目前关于行人意图分析的研究大多侧重于路径预测和固定场景研究[2-3],对推广到辅助驾驶系统有一定困难.部分学者倾向于通过分析行人运动姿态的变化规律, 判断行人是否有通过马路的意图[4]. 文献 [5]同样对行人姿态进行分析,得到 396 维的姿态特征,预测单目图像中行人过马路的意图. 文献 [6]基于详细的三维姿态信息对行人姿态骨架特征进行意图识别. 文献 [7]利用行人头肩骨骼关节的位移生成行人姿态序列来识别行人意图,对多人场景适应性较差. 文献 [8]基于身体重心和质心分析行人姿态和方向, 结合空间位置检测行人穿越街道的意图. 这些对行人意图的分析提供了许多不同的研究方向和思路.

由于目前关于行人意图方面的研究未考虑行车场景的划分,仅研究行人在一般道路上的过街意图,未针对道路类别进行区分, 不适用于多场景下的行人行为对车辆驾驶影响的分析, 如行人在人行横道过马路和一般道路上过马路的风险是不同的, 并且行人、辅警和道路共同约束了人车交互的场景. 因此, 本文重点研究不同道路场景中的行人行为危险评估, 通过对行人危险评估模型的应用场景进行分类,包括普通道路、人行横道和有辅警道路. 结合实际行车场景中的道路信息、驾驶信息和行人行为,综合评估行人行为危险程度.

1　数据预处理

目前,辅助驾驶领域的公开数据集有很多,本文选取 3 种广泛应用于计算机视觉领域的数据集,用于目标检测和行人、道路信息提取,分别是目标检测领域的 COCO2017[9]数据集、自动驾驶领域的 BDD100K[10]数据集和 JAAD[11]数据集. COCO2017 数据集和 BDD100K 数据集包含的场景类别数量和数据规模如表 1 所示.

COCO2017 数据集已对图像中的目标进行分割和位置标定,并对图像中的人体关键点进行标注,包含 17 种用于表示人体姿态的关键点信息.

BDD100K 数据集包含 10 万张图像, 每张图像来源于真实驾驶场景中的第 10 帧视频图像,并对道路目标、可行驶区域等进行了标注.

表 1　数据集

Tab.1　Dataset

数据集名称	场景类别数量	数据规模
COCO2017 目标检测	80	118 000
COCO2017 关键点检测	17	64 115
BDD100K	10	100 000

JAAD 数据集共有 346 段自然行驶状态的视频,视频每秒 30 帧, 分辨率为 1 920 × 1 080. 该数据集考虑到行人行为和交通场景的复杂性, 对多种因素进行注释,如不同的天气条件、行人反应、道路因素、交通场景和人口统计等等.

该数据集共标注 2 700 人,并对其中 686 人的行为状态、场景信息等细节进行了详细的标注. 本文利用 686 人(其中 211 人处于路口)的标注信息,进行图像筛选, 得到 132 700 张仅包含行人的图像 (共 686 段序列), 每张图像记录行人 ID、视频帧数、位置、遮挡、手势、反应、行动、过马路状态、过马路位置等信息. 本文对视频中的图像序列进行行人关键点的检测, 得到行人的关键点位置信息, 同时结合反应、手势、行走标注,生成行为的 3 类划分,即站立、行走、跑步. 对生成的 686 个行人关键点序列,进行筛选处理,删除无效检测数据并矫正较差数据,最终获取需要的行人行为的训练数据 90 644 张, 作为行人行为模型训练和验证数据.

数据选择流程如下:

1)根据数据集中的标注信息对有详细标注的行

人图像进行提取,并保留图像序列和相应标注信息.

2)通过关键点检测方法对无遮挡行人序列进行处理,获得每个序列行人的关键点信息.

3)对关键点信息进行筛选.

(a)选取有效数据,即头、肩、手、腿等明显关节处关键点信息无误检的数据.

(b)填补数据,根据前后准确帧校准数据.

(c)提取序列大于 8 的数据作为训练数据.

(d)对关键点坐标进行归一化处理.

4)对提取有效序列的行人标注其行为,根据详细标注,得到 90 644 张图,分为站立(15 261 张)、行走(72 123 张)、跑步(3 260 张)3 类.最终获得长度 9 的标注序列 43 274 条.

2 危险评估模型

本文提出的行人危险评估模型,总体框架如图 1 所示.

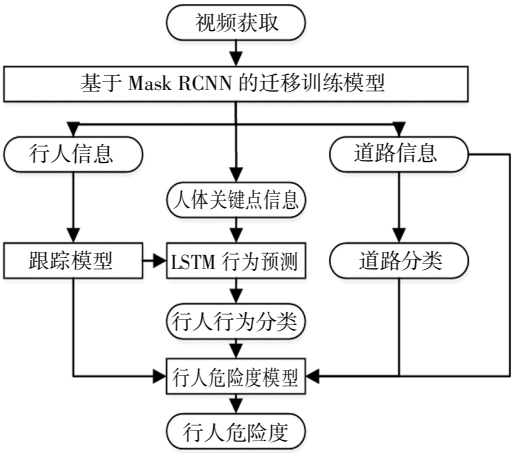


图 1 整体框架图

Fig.1 Overall flow chart

首先对交通场景中的斑马线、行车道、信号灯、辅警和行人等目标信息进行检测; 然后根据检测出的道路信息对场景进行分类,分为普通道路、人行横道、有辅警道路 3 类;再使用 deepsort[12]对行人进行跟踪, 得到行人姿态的时间序列, 通过 LSTM(Long Short Term Memory) 网络得到行人行为的分类输出,类别包括站立、行走、跑步;最后结合行人的行为、位置和道路信息,评估行人危险度.

2.1 目标检测

目标检测在智能驾驶研究中有着广泛的应用. 目标检测的准确率对后续的行人跟踪和姿态评估有重要影响, 最终影响行人危险度评估的准确率. 目前,目标检测方法发展迅速,已经从传统的基于滑动窗口策略的检测方法发展到当前的基于深度学习的检测方法. 由于 Mask RCNN 是一个实例分割架构,不仅可以检测目标,还可用于姿态估计,因此本文综合考虑实际场景,采用 Mask RCNN[13]作为目标检测方法.

Mask RCNN 是 Faster RCNN 的框架的拓展,通过在 Faster RCNN 网络中增加全连接网络, 提升算法灵活性. 如图 2 所示, 本文采用 Resnet101 网络获取主干网络图像特征,采用 RPN 生成建议框,采用 ROI Align 对建议框进行处理,最终通过 FCN 得到目标检测结果.

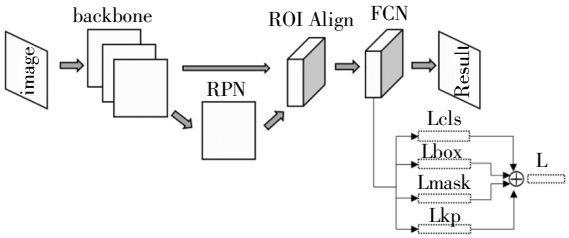


图 2 Mask RCNN 模型

Fig.2 Mask RCNN model

如图 3 所示, 输入主干网络图像后获得多图层级特征,如第一行特征图像所示,特征图像的尺寸依次减小, 第二行图像为通过上采样和特征融合得到的多层特征融合图像, 每层特征含有不同层级的图像信息和目标检测框, 经过极大值抑制选取高置信度的目标检测框,作为全连接网络的输入,得到最终目标检测结果.

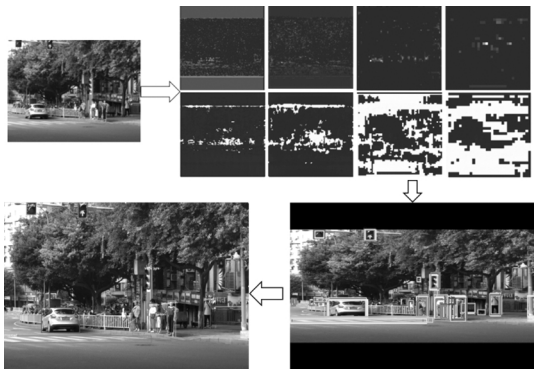


图 3 图像特征图

Fig.3 Image feature map

在全连接网络中, 添加了一个人体关键点检测分支任务,损失计算加入检测框损失、分类损失、实例损失,其中检测框损失为:

$$\text{smooth}_{L_1}(x)=\begin{cases}0.5x^2, |x|<1\\ |x|-0.5, \text{其它}\end{cases} \tag{1}$$

式中:x 为检测框的真实值和预测值之间的差值. 分类损失和关键点损失为交叉熵损失 $H_{y'}(y)$的均值,表示如下:

$$H_{y'}(y)=-\sum y_i' \log(y_i) \tag{2}$$

式中:y'和 y 分别为真实标签和该标签输出值. 目标检测对 4 类进行检测,分别是道路场景的人行横道、信号灯、辅警和行人. 本文对单目车载视频,每帧应用目标检测模型,得到图像中目标的位置信息,以及车道信息[14]中的行人关键点的信息. 道路场景的检测是对场景分类的基础. 行人检测在后续多目标跟踪、行人和道路位置分析、行人行为分析中有重要作用.

2.2 道路分类

在已有的行人分析方法中,主要研究行人意图,未对道路场景进行区分. 本文针对不同道路场景对行人危险度的影响, 以重庆为例对国内的城市道路进行分析, 分别在双休日和工作日的下午 14:00–18:00,记录沙坪坝区学校附近的 3 个路口(包含指挥情况) 的过街行人数量; 工作日非高峰期 (下午 14:00–16:00)的普通路段过街行人情况,后者由记录仪在实际行车场景记录, 并纪录了违规行人的数量,具体数量见表 2.

表 2　不同情况行人数量

Tab.2　Number of pedestrians in different situations

	过街人数/个	违规人数/个	时长/h
高峰期(指挥)	845	16	2
高峰期(人行道)	547	48	2
普通(道路)	3	3	1.5
普通(路口)	65	6	0.5

结果表明,在有辅警道路场景,过街行人数量最多,同时违规人数比例较低;无辅警道路场景,人行横道处存在较多行人过马路不遵守红绿灯, 大幅超出斑马线区域的情形,而在普通道路,则极少有行人横穿马路,故将交通场景分为 3 类:

1)普通道路.

2)人行横道.人行横道是人车交互频繁的场景,所以本文将有人行横道的场景与一般行车时场景区分开.

3)有辅警道路.有辅警指挥的地方,行人、车辆更为密集,但双方也有更大的约束和注意力.

3 种道路场景如图 4 所示.

(a)普通道路　　(b)人行横道　　(c)辅警道路

图 4　不同道路场景

Fig.4　Different road scenarios

2.3 行人行为检测

通过上文的行人识别,本文得到了行人检测框,以及检测框中的行人骨架节点. 检测行人之后对行人进行追踪,使用获得的 2D 姿态数据,累积行人骨架随时间变化的骨架特征.

本文使用一种多目标跟踪方法 Deepsort[12],算法主要使用两种关联关系来衡量是否匹配.

1)运动关联 $d^{(1)}$:使用马氏距离描述.

$$d^{(1)}(i,j)=(d_j-y_i)^{\mathrm{T}}\boldsymbol{S}_i^{-1}(d_j-y_i) \tag{3}$$

式中:d_j 表示第 j 个检测框的位置;y_i 表示第 i 个跟踪器对目标的预测位置;$\boldsymbol{S}_i$ 表示检测位置与平均跟踪位置之间的协方差矩阵.

2)外观关联 $d^{(2)}$:使用余弦距离度量.

$$d^{(2)}(i,j)=\min\{1-r_j^{\mathrm{T}}r_k^{(i)}|r_k^{(i)}\in R_i\} \tag{4}$$

对每一个 d_j 得到一个特征向量 r_j(128 维,基于文献[12]的方法),T 为矩阵的转置,对每一个跟踪目标构建序列 R_i, 存储每一个跟踪目标成功关联的最近 k 帧的特征向量. 对两种匹配加权融合,得到关联结果 $C_{i,j}$ 为:

$$C_{i,j}=\lambda d^{(1)}(i,j)+(1-\lambda)d^{(2)}(i,j) \tag{5}$$

由于车载视频的运行性,运动关联有较差效果,但运动关联的阈值依然有效,如果不满足关联阈值,就不能进入综合度量,这里 $\lambda=0.1$. 运动关联对短期预测特别有用;外观关联对长时间遮挡有效.

对于跟踪匹配参数, 本文设置预匹配和匹配失败的阈值,如果连续 9 帧(0.3 s)之内没有匹配,则跟踪生成新的行人. 如果在 30 帧(1 s)内,行人没有高匹配分数和新的检测结果则结束跟踪.

将获得的行人的骨架节点序列, 经过 LSTM 网络学习,获得行人行为的分类输出,包括站立、行走、跑步等. LSTM 是一种特殊的循环神经网络, 该网络能解决长依赖问题,适合于对序列信息进行预测.所以本文选择 LSTM 网络来得到行人行为分类,LSTM 模型的建立如表 3 所示.

表 3 中的输入值有 3 维, 第一维为批处理数量

150,决定了每次更新网络参数的参考数据量;第二维为行为姿态序列的时间步长 9, 这跟本文跟踪方法中,出现新的行人确定帧数是相同的;第三维就是本文网络使用的主要关键点特征数. 输出层为全连接和 softmax 结合,得到 3 种行为的分类. LSTM 网络分两层,每层 64 个神经元.

表 3 LSTM 模型参数设置

Tab.3 LSTM model parameter setting

LSTM 参数	数值
输入	[150,9,26]
输出	3
学习率	0.001
神经元 / 层数	64/2
全连接层	1×3
训练次数	100

2.4 行人危险评估

最后通过行人危险度评估模型得到行人危险值. 行人危险评估模型的分场景评估算法流程如图 5 所示.

1 Pedestrian Risk Assessment Model

```
1: Pedestrian P, Crossroad C, Assist Police A
2: if P is exist then
3:    if A is exist then
4:       Calculate Risk = Risk(p = c)
5:    else
6:       if C is exist then
7:          Calculate Risk = Risk(p = b)
8:       end if
9:    end if
10:   Calculate Risk = Risk(p = a)
11: end if
```

图 5 行人危险评估算法

Fig.5 Pedestrian hazard assessment algorithm

模型综合了目标检测的结果,对行人信息、道路信息进行结合. 考虑城市道路环境特色和行人在多种道路场景的不同反映,得到的行人危险度函数为:

$$\mathrm{Risk}(p)=\min(A(p)+B(p),1) \tag{6}$$

式中:$A(p)$ 是不同场景下行人行为的危险度;$B(p)$ 是在不同场景下行人与道路相对位置的危险度;p 代表场景分类. 危险度设计为[0,1]之间的值,危险函数的值越接近 1,表示行人具有越高的危险度,各危险度情况如表 4 所示.

表 4 危险度

Tab.4 Degree of risk

所处场景	p(场景类别)	$A(p)$	$B(p)$
一般道路	a	(0,20%,40%)	10%,50%
人行横道	b	(0,30%,60%)	0,40%
辅警指挥道路	c	(0,20%,30%)	20%,60%

在表 4 中,a、b 和 c 代表了道路的分类, 在普通道路,行为和位置有相同的影响;在人行横道,运动行人有明显的过马路意图,危险度更高;在有指挥的情况下,人车行为更规范,并且行人十分密集,检测难度高,这时位置的影响更大. $A(p)$代表在场景 p 的情况下行人站立、行走、跑步 3 种行为的危险度. 通过数据分析,$A(\mathrm{a})=(0,20\%,40\%)$表示在一般道路交通场景下行人站立、行走、跑步 3 种行为对应的危险度分别为 0、20%、40%. $B(p)$为道路位置信息和行人检测框位置的相对关系,公式如下:

$$B(p)=(B_{p\max}-B_{p\min})*\left|((S-P)/(S-R))\right|+B_{p\min} \tag{7}$$

式中:$B_{p\max}$ 和 $B_{p\min}$ 为在 p 场景下行人的最大危险度和最小危险度. S 为图像宽度的一半;P 为行人下边框中心点距侧边框的距离;R 为道路位置(单侧),位置都是基于图像尺寸而得, 公式中 3 个参数区分在图像的左右半区. 综合 $A(p)$和 $B(p)$得到行人在多种道路场景的危险度.

3 实验结果与分析

3.1 数据集

为了测试验证本文提出方法的有效性, 采用在计算机视觉领域广泛应用的 COCO2017 数据集、BDD100K 数据集和重庆市街景数据集,用于提取图像特征、训练目标检测模型,以及测试目标检测模型和行为危险评估模型的性能.

本文选取 500 张图像进行标注, 标注图像中的行为、信号灯、人行横道及其类别、辅警检测框等,其中 400 张已标注的图像用于训练, 剩余 100 张图像用于测试验证. 测试数据涉及 80 个普通道路场景, 30 个人行横道场景以及 10 个有辅警场景.

3.2 超参数设置

超参数的设置会影响 Mask RCNN 的效果和性能, 本文设置超参数的原则是保证准确性的同时尽可能提升训练和检测速度. 本文分别针对不同的数

据集设置不同的超参数,具体参数设置如表 5 所示.

表 5　检测模型参数设置

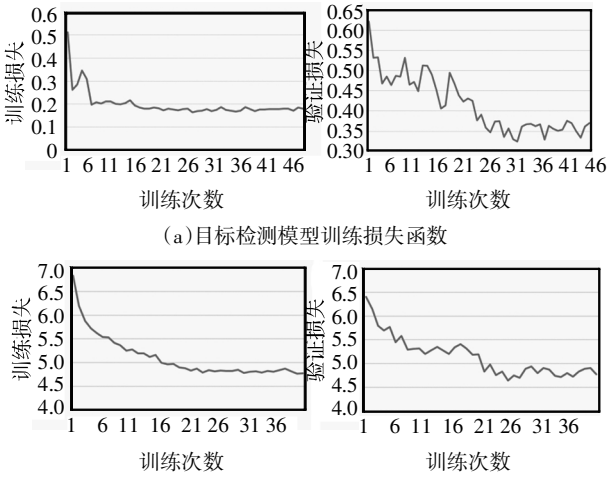
Tab.5　Detection model parameter setting

训练次数	使用数据	训练层数	训练参数	学习率	训练步长
第 1 次	COCO2017-instance	50	Head/全部	0.002/0.000 2	1 000
第 2 次	BDD100K	50	部分	0.002	100
第 3 次	COCO2017-keypoint	40	Head	0.002	1 000

3.3　目标检测结果分析

目标检测是本文整体方案的基础，对本文的行人信息识别和道路分类有重要作用. 因此,本文针对车载环境中的目标检测和关键点检测进行实验评估,分析 Mask RCNN 模型性能.

图 6 为目标检测模型在收敛过程中的损失,包括行人检测框和行人关键点识别过程中的损失,结果表明,经过一定次数的迭代训练,模型收敛并取得较好的收敛效果. 其中,行人检测框和行人关键点检测的精度 AP(Average Precison)和召回率 AR(Average Recall)结果如表 6 所示.



(a)目标检测模型训练损失函数

(b)关键点检测模型训练损失函数

图 6　损失函数曲线

Fig.6　Loss function curve

表 6　AP 和 AR

Tab.6　AP and AR

类别	AP(0.5)	AP(0.75)	AP(0.5:0.9)	AR(0.5:0.9)
检测框	0.427	0.694	0.490	0.485
关键点	0.50	0.789	0.556	0.574

本文利用收敛后的目标检测模型对行人行为分类模型进行训练,基于 JAAD 行人关键点序列,采用 LSTM 对行人关键点序列进行训练,得到行人行为分类模型. 行人行为分类准确率如图 7 所示. 实验结果表明，行人行为分类模型在 25 次迭代训练后收敛,准确率稳定在 90%左右，表明该模型能有效进行行人行为分类.

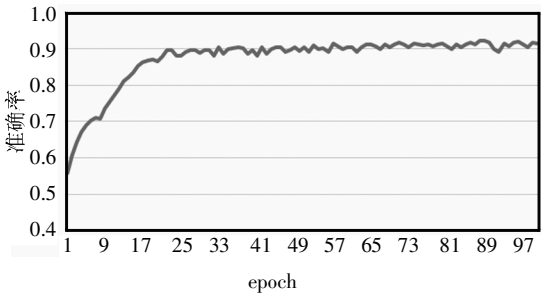


图 7　行人行为分类准确率

Fig.7　Classification accuracy of pedestrian behavior

3.4　行人危险评估结果分析

本文基于上述目标检测模型和行人行为分类模型进行行人危险评估,根据行人危险程度预设阈值,区分行人危险程度，并设置经验阈值来对行人危险情况进行分类，判断场景中的行人行为是否属于危险情况. 评估准确性如图 8 所示,结果表明当阈值设置为 0.7 时准确率最高,平均准确率为 91.7%,其中普通道路场景的准确率为 91.3%，人行横道场景的准确率为 93.3%，有辅警道路的准确率为 90.0%. 因此,本文将危险程度大于 0.7 的行人行为归类为危险行为,驾驶系统需要采取避险措施. 另外,针对部分无道路的场景,对行人危险程度进行评估,准确率为 83.3%,相对于有道路的场景而言较低,是文本技术局限性之一.

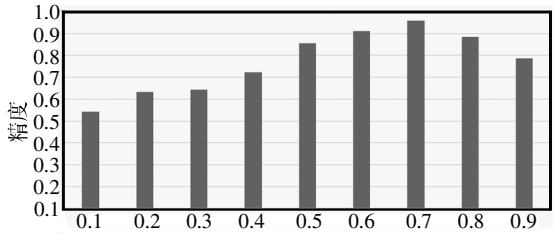


图 8　阈值选择

Fig.8　Threshold selection

4　结　论

本文提出了一种城市道路场景下行人危险评估方法,应用于智能辅助驾驶系统,辅助驾驶员安全行

车. 通过检测模型,本文成功获取了行人信息和道路场景信息,分析行人行为危险程度. 该方法首先建立目标检测模型,提取行车环境和人体姿态关键点,然后通过对行人行为进行分类, 最后建立行人行为危险评估模型,并实验分析模型的最佳阈值,验证模型在不同城市场景下的准确率. 该模型可用于辅助智能驾驶系统进行行人避让. 实验结果验证本文提出的行人危险度评估方法的有效性.

同时,本文提出的方法存在一定局限性,有待进一步研究,主要包括:场景可变性和颜色相似性会影响目标检测结果, 在多重遮挡场景下目标检测存在较大误差.

参考文献

[1] GHORI O,MACKOWIAK R,BAUTISTA M,*et al*. Learning to forecast pedestrian intention from pose dynamics [C]// Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu:IEEE,2018:1277—1284.

[2] SCHNEIDER N,GAVRILA D M. Pedestrian path prediction with recursive Bayesian filters:A comparative study [C]// Proceedings of the Pattern Recognition. Berlin,Heidelberg:Springer Berlin Heidelberg,2013:174—183.

[3] KELLER C G,GAVRILA D M. Will the pedestrian cross? A study on pedestrian path prediction [J]. IEEE Transactions on Intelligent Transportation Systems,2014,15(2):494—506.

[4] SCHMIDT S,FRBER B. Pedestrians at the kerb – recognising the action intentions of humans [J]. Transportation Research Part F Psychology & Behaviour,2009,12(4):300—310.

[5] FANG Z,L PEZ A M. Is the pedestrian going to cross? Answering by 2D pose estimation [C]// Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu:IEEE,2018:1271—1276.

[6] LI Y,LAN C,XING J,*et al*. Online human action detection using joint classification-regression recurrent neural networks [C]// Proceedings of the Computer Vision – ECCV 2016. Cham:Springer International Publishing,2016:203—220.

[7] DENG Q,TIAN R,CHEN Y,*et al*. Skeleton model based behavior recognition for pedestrians and cyclists from vehicle sce ne camera [C]// Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu:IEEE,2018:1293—1298.

[8] HARIYONO J,JO K H. Detection of pedestrian crossing road:A study on pedestrian pose recognition [J]. Neurocomputing,2017,234:144—153.

[9] LIN T Y,MAIRE M,BELONGIE S,*et al*. Microsoft COCO:Common objects in context [C]// Proceedings of he Computer Vision – ECCV 2014. Cham:Springer International Publishing,2014:740—755.

[10] YU F,XIAN W,CHEN Y,*et al*. BDD100K:A diverse driving video database with scalable annotation tooling [J/OL]. https://arxiv.org/abs/1805.04687,2018-05-12.

[11] KOTSERUBA Y,RASOULI A,TSOTSOS J. Joint attention in autonomous driving (JAAD) [J/OL]. https://arxiv.org/abs/1609.04741v5,2017-04-27.

[12] VEERAMANI B,RAYMOND J W,CHANDA P. DeepSort:deep convolutional networks for sorting haploid maize seeds [J]. BMC Bioinformatics,2018,19(9):289.

[13] HE K,GKIOXARI G,DOLL R P,*et al*. Mask R-CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(2):386—397.

[14] BRABANDERE B,GANSBEKE W,NEVEN D,*et al*. End-to-end lane detection through differentiable least-squares fitting [J/OL]. https://arxiv.org/abs/1902.00293v1,2019-02-01.