

黑龙江大学硕士研究生学位论文中期审查表

学号	2181282	姓名	刘双才	指导教师	马慧
学科、专业	模式识别与智能系统		研究方向	智能检测与模式识别	
论文开题题目	基于深度学习的行人检测算法研究				
学位论文工作详细进展情况	论文答辩题目1：基于深度学习的行人检测算法研究				
	论文预计字数	25000	已完成字数	12500	
	三级标题及内容				
	一、三级标题				
	中文摘要				
	Abstract				
	第一章 绪论				
	1.1 研究的目的和意义				
	1.2 国内外研究现状				
	1.3 基于深度学习的行人检测研究				
1.4 研究内容与技术路线					
第二章 深度学习理论基础与行人检测研究					
2.1 引言					
2.2 行人检测技术理论					
2.2.1 典型的行人特征表示					
2.2.2 基于传统的机器学习行人检测					
2.2.3 行人检测评价指标					
2.2.4 行人检测数据集					
2.2.5 行人检测难点					
2.3 深度学习理论					
2.3.1 深度学习基础理论					
2.3.2 Faster-RCNN 系列算法					
2.3.3 YOLOv3 系列算法					
2.4 本章小结					

¹ 论文答辩题目如与开题题目不一致，须附修改后新的开题报告。

第三章 基于 CenterNet 的行人检测研究
3.1 引言
3.2 CenterNet 网络框架
3.3 主干网络介绍
3.3.1 Hourglass-104 特征提取网络
3.3.2 DLA-34 特征提取网络
3.4 实验
3.4.1 数据集样本划分
3.4.2 实验基本设置
3.4.3 实验结果与分析
3.5 本章小结
第四章 基于小尺度行人检测研究
4.1 引言
4.2 小尺度行人研究分析
4.3 针对精度保留的激活函数
3.4.1 relu6 激活函数
3.4.2 h-switch 激活函数
4.4 基于信息流通提出的 TRB-Net 模块
4.5 实验分析
4.6 本章小结
第五章 基于 CenterTrack 网络的行人追踪技术
5.1 引言
5.2 CenterTrack 网络研究分析
5.2.1 基于中心点追踪
5.2.2 局部关联算法
5.3 实验分析
5.4 本章小结
结论
参考文献

	<p>致谢</p> <p>攻读硕士学位期间发表的论文</p> <p>独创性声明</p> <p>二、论文研究进展情况</p> <p>通过查阅、分析和整理有关深度学习和行人检测方面的文献资料，确定了基于深度学习的行人检测算法的研究方案，并制定了工作计划。目前论文研究的进度与原工作计划基本一致。</p> <p>查阅文献，了解行人检测的研究背景、国内外研究现状和发展趋势，确定研究方案；</p> <p>设计行人检测训练的基本环境并搭建，获取数据集并对数据集进行整理和分析；</p> <p>完成对行人检测基本框架的训练，并开始对比其他试验，分析框架不足和提出改进；</p> <p>完成车载行人数据的行人追踪算法的网络训练和搭建，并进行对比分析实验结果；</p> <p>整理数据及相关资料，拟定论文初稿。</p> <p>三、阶段成果</p> <p>1、通过查阅、整理、归纳基于深度学习的行人检测和追踪算法的相关资料，针对行人检测在实际场景中应用范围和小尺度行人检测的难点确定了系统设计方案。</p> <p>2、获得行人数据并对行人检测难点进行分析</p> <p>本文中使用两个公共行人检测数据集，Caltech 行人数据集共 46806 张行人图片，ETH 行人数据集共 2820 张行人图片。如图 1 所示，训练集中数据多为大尺度行人，特征比较明显，但是测试集多为小尺度行人，如图 2 所示，小尺度行人特征点过少，并不容易被捕获，所以针对小尺度行人检测是本文研究的难点重点。</p>
--	---

不同高度行人的数量比例

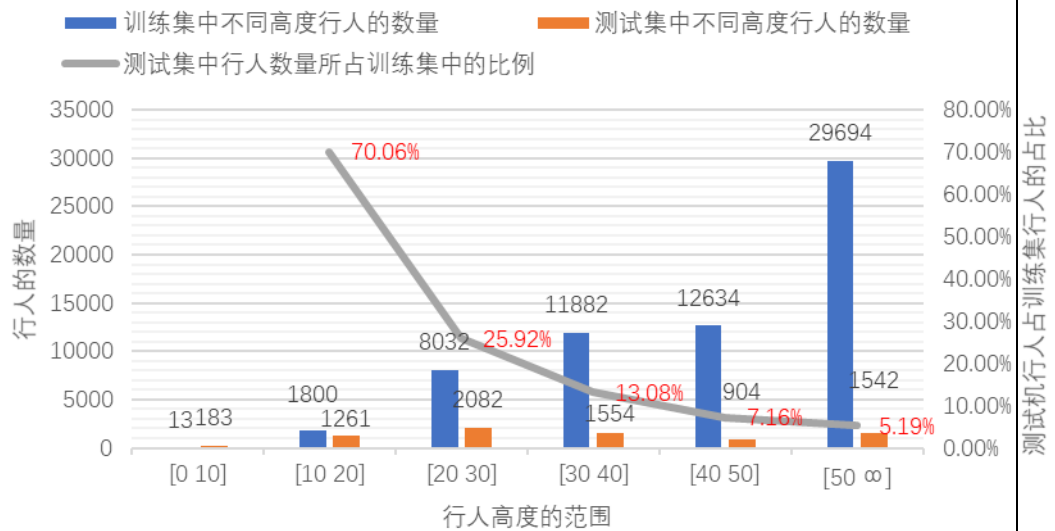


图 1 小尺度行人占比



图 2 小尺度行人数据

3、实现对 CenterNet 的行人检测算法

CenterNet 网络是 anchor-free 家族网络中的一员，是基于中心点检测的网络且端对端的简单快速检测算法。如图 3 所示，CenterNet 网络框架是 one-stage 算法，通过卷积神经网络对行人图片进行定位和分类。CenterNet 网络中采用的 backbone（Hourglass-104 和 DLA-34）进行特征提取。通过对行人数据进行特征提取得到热力图，在热力图上，通过高斯分布获得中心点，通过对中心点回归得到相应的行人尺寸。

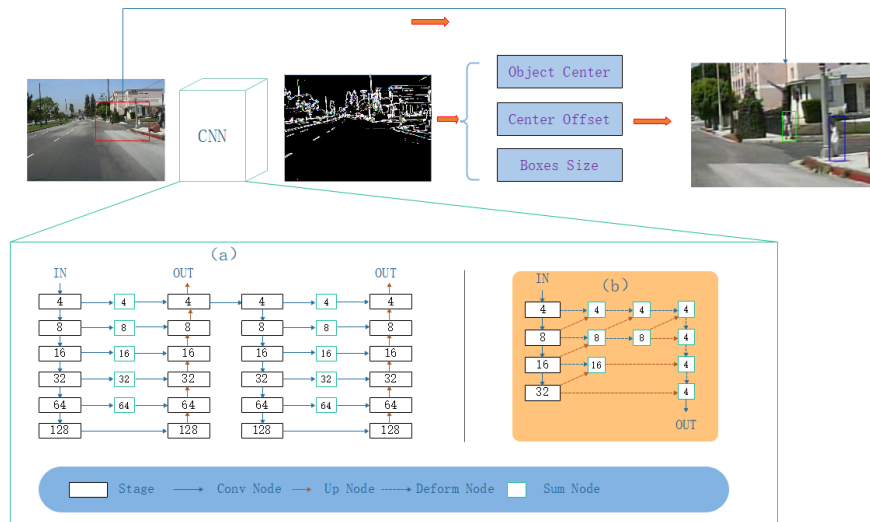


图 3.CenterNet 网络结构.(a)Hourglass-104;(b)DLA-34.

本算法训练的具体步骤在 Algorithm 1 中。

Algorithm 1 本文算法流程

Input:

标记好行人检测框和中心点的图片和标签；网络中需要配置的参数。

Output:

模型预测测试集的目标检测框。

- 1: 进行网络推理，提取热力图上行人的最大的 100 个峰点值；
- 2: 通过确定行人的峰点，得到对应的行人长宽值和对应偏移量；
- 3: 将得到的数值带入到损失函数，通过网络不断迭代推理会生成更合理的模型；
- 4: 将校正集的数据放到模型上进行测试，如果欠拟合，模型需要继续增加迭代次数，如果过拟合则适当减少训练次数，最后得到合理的网络模型；
- 5: 将测试集数据放到模型上就可以准确定位识别行人。

在表 1 中，对比行人数据在 Hourglass-104 和 DLA-34 的对比结果，可以看出多尺度训练的情况下，虽然可以提高少量的精准度，但是相对速度却慢了很多，这不符合现实工业要求，所以本文针对这种情况，选择了单精度情况下的 DLA-34 作为我们选择的提取特征网络。

Single Scale	AP-1	AP-2	Time(ms/img)
Hourglass-104	33.4	79.3	77
DLA-34	38.8	79.9	28
Multi Scale			
Hourglass-104	35.1	80.1	365
DLA-34	39.5	80.2	127

表 1 单尺度和多尺度下两个 backbone 对应的两个不同数据集 (AP-1:Caltech,AP-2:ETH) 的结果

4、实现小尺度行人检测算法

如图 4 所示，小尺度行人检测框架是在 CenterNet 网络的基础上提出了 TRB-Net 模块组和针对小数敏感问题引用了 relu6 激活函数。针对小尺度行人特别是特征点过少，环境阴暗的情况下，可以进一步增强信息流通。

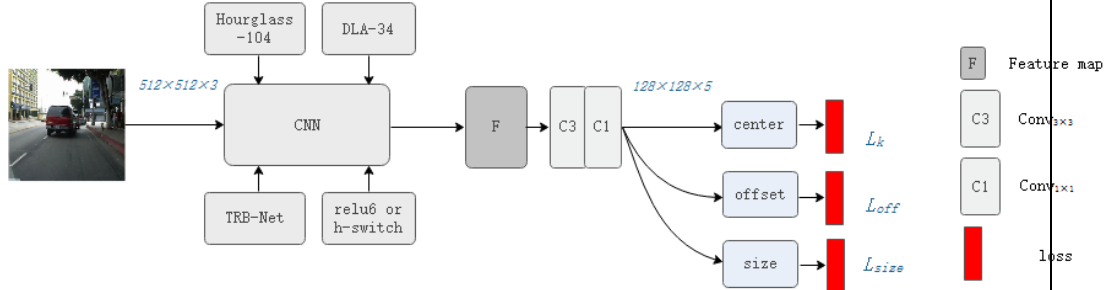


图 4. 小尺度行人检测网络框架。

本文网络模块中的 relu 在将负信号置 0 影响了网络的传播，这种情况在刚开始训练的时候尤其严重，所以本文提出了一种网络不同阶段的三种不同基础网络构建模块组 TRB-Net（如图 5 所示，一个网络阶段就是一个下采样或者上采样阶段）。在阶段初期，对比与原始的基础网络结构模块，减少经过了激活函数的次数，促进了信息在网络中的流动。如图 5 所示，将网络结构中的 TRB-Net 模块组置于不同位置，分别命名为 TRB-Net- I，TRB-Net- II，TRB-Net-III。

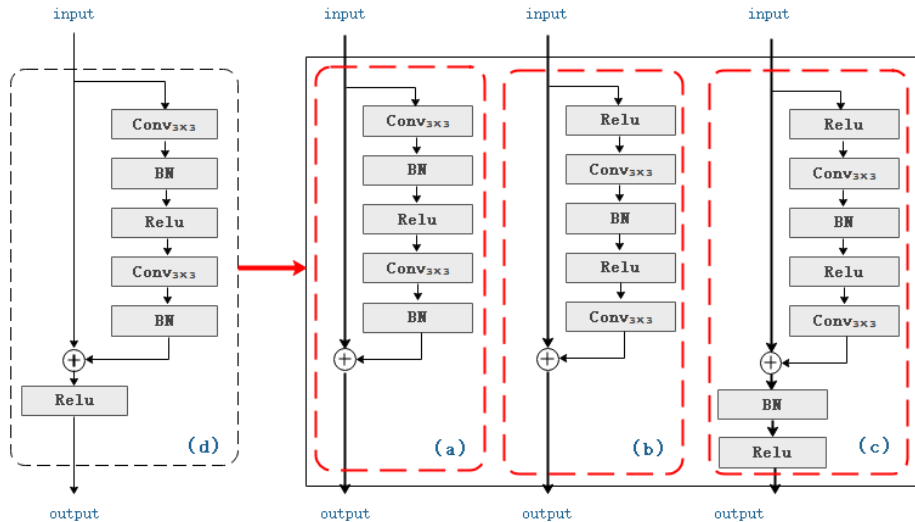


图 5. Three ResNet Block(TRB-Net)。网络在每一个上采样或者下采样阶段分别使用三个基础模块组(a)阶段初期使用的残差模块 TRB-Net- I；(b)阶段中期时期使用的残差模块 TRB-Net- II；(c)阶段最后使用的残差模块 TRB-Net-III；(d)原网络中使用的基础模块。

CenterNet 网络中的 Backbone 中的 Hourglass-104 网络和 DLA-34 网络中的激活函数都使用的是 relu，随着激活函数近几年的不断改进，relu 虽然

解决了梯度消失和梯度爆炸等问题，但是已经不能对网络精度有较大大的改进，所以本研究将网络结构中使用的激活函数 relu 进行了最大值的限制，并通过对比其他方法来进行实验评估。

relu6 是在原来的激活函数 relu 上进行了改进，限制了它的最大输出值为 6，如式（1）：

$$\text{relu6}(x) = \begin{cases} \text{relu}(x) & \text{if } x < 6 \\ 6 & \text{if } x \geq 6 \end{cases} \quad (1)$$

如果对 relu 的激活范围不加限制，会产生数值爆炸，分布在一个很大的范围内，则会带来一定的精度损失。同时 relu6 激活函数可以让网络尽早的学习到稀疏特征，而且，因为限制了最大的数值，将会增加小数位表达能力，这对原本有遮挡或者小样本的数据集有很大的好处。

h-switch 是在 relu6 上进行了一定的改进从而达到限制，使其在网络层数加深的同时，非线性激活函数的成本也会随之减少，如下式（2）：

$$\text{h-switch}(x) = x \frac{\text{relu6}(x+3)}{6} \quad (2)$$

如图 6 中可以看到（b）中在所有情况下 FPPI 的结果，对比其他的方法，本文提出的方法优势在于小尺度行人检测结果的提升，（b）中针对行人的高度为像素大于 20 的所有行人，所以除了 $[50 \infty]$ 部分行人的检测， $[20 \ 50]$ 中小尺度的行人检测就成了 FPPI 值浮动的关键，本文提出的方法较 AdaptFaster RCNN 的结果高出 10.87%，表明本文提出的方法是真实有效的。

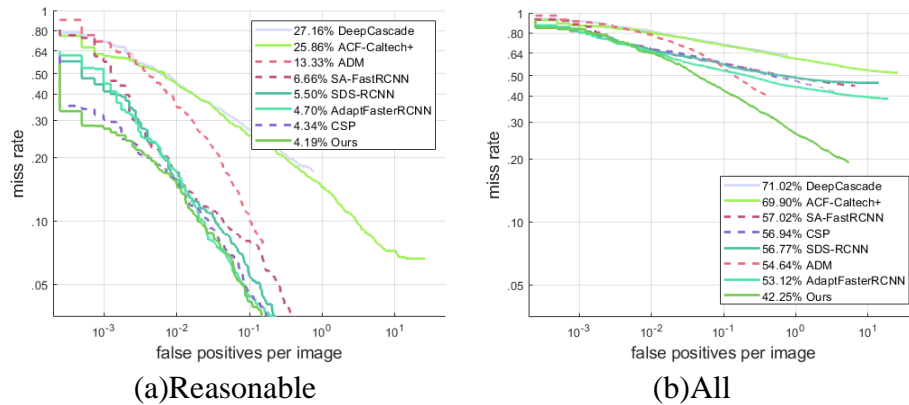


图 6 在 Caltech 数据集上和最新的方法进行对比

在表 1 中，观察 A 部分得到经典的 CSP、SDS-RCNN、AdaptFaster RCNN

算法在行人高度为[40 50]像素的结果,相当于对于本文提出的方法在 MR-2 低了 6.31%、4.31%、0.2%,但是针对低于 40 像素点的检测结果,本文提出的方法比其他方法要高出很多。在表 2 中的 B 部分,本文将对行人在 $\text{Occ} \in [2 \dots 65]$ 遮挡部分的小尺度行人进行结果对比,本文提出的方法在[40 50]范围内可以捕获到 43.41%的行人,对比 AdaptFaster RCNN 高出 10%,而置于 40 到 30 个像素点的遮挡行人,其他经典算法基本并不具有很好的捕获能力,本文提出的算法 MR-2 可以达到 67.44%。C 部分将 IOU 值从 0.5 变成了 0.75,对行人的准确性要求更高,在这种情况下,本文的方法依然可以达到 48.79%,说明针对小尺度的行人,本文再准确性和召回率上达到了双赢。

Method	Occ ∈ [.2 .65]	Test pedestrians height range MR-2(%)				IoU
		[10 20]	[20 30]	[30 40]	[40 50]	
A						
CenterNet+ relu6	-	82.60	55.89	39.41	16.70	0.5
CenterNet+ h-switch	-	91.76	57.64	47.13	18.49	0.5
CenterNet+ TRB-Net	-	91.20	56.33	39.42	18.02	0.5
Ours	-	80.98	54.47	38.71	15.71	0.5
B						
CenterNet+ relu6	✓	-	-	69.05	54.91	0.5
CenterNet+ h-switch	✓	-	-	73.35	58.64	0.5
CenterNet+ TRB-Net	✓	-	-	70.01	54.48	0.5
Ours	✓	-	-	67.44	43.41	0.5
C						
CenterNet+ relu6	-	-	-	-	51.96	0.75
CenterNet+ h-switch	-	-	-	-	52.04	0.75
CenterNet+ TRB-Net	-	-	-	-	51.47	0.75
Ours	-	-	-	-	48.79	0.75

表 2 针对 occ, height, IoU 进行的对比实验数据结果; A 部分针对小尺度行人检测结果; B 部分针对遮挡的小尺度行人检测结果; C 部分针对 IoU=0.75 的小尺度行人的检测结。

	<p>四. 学位论文工作完成比例: 50%</p> <p>五. 后期工作安排:</p> <ul style="list-style-type: none">1、完成行人追踪的实验对比工作。2、增加更多模型之间的对比。3、撰写论文。 <p>本人签名: _____</p> <p>年 月 日</p>
--	---

导师意见：（1.对论文进展状况的评价及是否能够继续进行学位论文工作；2.如论文题目变更，须说明原因。）

导师签字：

学位评定分委会意见：（1.是否通过中期审查；2.如论文题目变更，是否同意）

学位评定分委会主席签名：（公章）

年 月 日

研究生学院意见：

（公章）

年 月 日