

结合双金字塔特征融合与级联定位的车牌检测

张俊青¹,熊玉洁^{1,2},孙宪坤¹,高永彬¹

1.上海工程技术大学 电子电气工程学院,上海 201620

2.华东师范大学 上海多维度信息处理重点实验室,上海 200241

摘要:为了解决复杂环境中不同因素干扰车牌检测精确度的问题,提出了一种基于双金字塔特征融合的复杂环境下车牌检测算法。通过采用Mish激活函数的残差网络(ResNet101-M)对输入图像进行初级特征提取;在传统特征金字塔网络(feature pyramid network,FPN)的基础上,提出了一种改进的双金字塔特征融合网络(siamese feature pyramid network,SFPN)。被提取的初级特征被送入该网络进行多层特征融合。融合后的特征被送入基于形状先验的锚点设置网络来确定感兴趣区域。将所生成的感兴趣区域送入级联定位网络从而得到准确的车牌检测结果。实验结果表明,该算法在AOLP与CCPD车牌数据集上均能够有效提升检测性能。

关键词:车牌检测;深度学习;双金字塔特征融合;级联定位

文献标志码:A **中图分类号:**TP391.41 **doi:**10.3778/j.issn.1002-8331.2108-0439

License Plate Detection Using Siamese Feature Pyramid and Cascaded Positioning

ZHANG Junqing¹, XIONG Yujie^{1,2}, SUN Xiankun¹, GAO Yongbin¹

1.School of Electric and Electronic Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

2.Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China

Abstract: In order to solve the problem of license plate detection in complex environment, a novel license plate detection algorithm using siamese feature pyramid and cascaded positioning is proposed. Firstly, the original features of the images are extracted by the residual network with the Mish activation function. At the same time, a siamese feature pyramid network(SFPN) is proposed in this paper. The extracted features are sent to the siamese feature pyramid network for multi-level feature fusion. Then, a region proposal network based on the prior of shape is used to generate the regions of interest. Finally, the regions of interest are fed to the cascaded positioning network to obtain the license plate location. The experimental results show that the proposed algorithm effectively improves the performance of license plate detection on both AOLP and CCPD datasets.

Key words: license plate detection; deep learning; siamese feature pyramid; cascaded positioning

在生活质量日益提高的当今社会,车辆已经成为了出行与运输不可或缺的交通工具。因此,如何高效地进行车辆监控与管理引起了研究者的广泛关注。近些年来,自动车牌识别(automatic license plate recognition, ALPR)在现代智能交通管理中发挥了重要的作用。在智能交通系统中,自动车牌识别系统能够实时监控车辆,从而降低人工交通管理的成本与负担。此外,ALPR

在汽车违规停放、高速公路收费站、无人监管的地下停车场与追踪可疑车辆等其他场景中也具有广泛的应用。

作为现代智能交通系统的重要组成部分,自动车牌识别的应用越来越普及。一般而言,车牌识别的基本流程可划分为图像获取、图像预处理、车牌检测、字符分割与识别共四个步骤^[1]。作为自动车牌识别的核心步骤,车牌检测需要快速找到车牌位置,为后续的识别工作奠定了基础。

基金项目:国家自然科学基金(62006150);上海市青年科技英才扬帆计划(19YF1418400);上海市多维度信息处理重点实验室开放课题基金(2020MIP001);中央高校基本科研业务费专项资金。

作者简介:张俊青(1996—),男,硕士研究生,研究方向为图像处理、深度学习,E-mail:895377607@qq.com;熊玉洁(1989—),通信作者,男,博士,讲师,硕士生导师,CCF会员,研究方向为图像分析与智能系统、文档图像分析;孙宪坤(1972—),男,硕士,副教授,硕士生导师,研究方向为自动化控制、图像采集与处理;高永彬(1988—),男,博士,副教授,硕士生导师,研究方向为人工智能、机器学习。

收稿日期:2021-08-25 **修回日期:**2021-10-28 **文章编号:**1002-8331(2023)02-0240-13

1 相关工作

针对车牌检测问题,研究者在20世纪80年代便已经开始致力于相关研究工作。早期研究中采用人工设计的图像特征来判定车牌的区域,如Niblack二值化方法^[2]、基于DFT变换方法^[3]与基于水平线搜索的方法^[4]等。尽管上述方法已经能够在简单应用场景中找出车牌的大致位置,但仍不能够精准地检测车牌的位置。

在上述研究方法取得了一些成效之后,研究者们从20世纪90年代左右开始设计更多的车牌特征来进行车牌检测。根据车牌特征关注区域的不同,此类传统车牌检测方法被划分为基于边缘检测的车牌检测方法^[5]、基于字符特征的车牌检测方法^[6]、基于颜色特征的车牌检测方法^[7]和基于纹理特征的车牌检测方法^[8]四种。

在基于边缘检测的检测方法中,通常会根据车牌矩形框的尺寸在图像中搜寻车牌候选区域。Al-Ghaili等^[9]提出了垂直边缘检测的车牌检测方法,具有定位速度与实时性高的优势。但由于边缘检测对多余边缘过于敏感,并不适用于复杂背景的图像。

在基于车牌字符特征的方法中,Cho等^[10]根据车牌字符与背景存在差异性的特点利用字符间宽度来定位车牌。较基于颜色特征的方法,此类方法对于光照不敏感,因而在检测光照条件变化下的车牌图像时鲁棒性更高。而此类方法的缺点则是当图像中存在干扰性文字时会导致精确度下降。

由于车牌与字符的颜色组合与背景相比具有特异性,Azad等^[11]根据此特征设计了一种基于车牌颜色特征的检测方法。此类检测方法将RGB图像转换HSV色彩空间后对其进行分割处理后再对每部分进行车牌检测,对于倾斜车牌具有较好效果,且车牌的尺寸、大小与角度变化均不会对其产生影响。但是此类方法在输入图像存在与车牌颜色相近的物体时,定位精确度较低。

基于车牌纹理特征方法的检测原理是根据字符与底色间局部纹理的明显特征变化来检测车牌区域。Deb等^[12]利用此类特征提出了基于滑动同心窗法(sliding concentric windows, SCW)的定位算法,其能较好处理光照与视角变化等干扰问题,而缺点在于会因纹理特征计算而导致计算速度偏慢。

除上述四种传统检测方法外,传统车牌检测算法的研究开始不仅仅局限于使用单一的车牌特征来检测车牌,许多基于多种车牌特征相结合的检测方法^[13-14]也取得了不错的成效。

近年来,人工智能领域的发展十分迅速。深度学习方法在目标检测领域效果卓著,从而使得车牌检测研究取得了新突破。研究人员通过设计基于卷积神经网络(convolutional neural network, CNN)的算法来提升车牌检测的精度。和传统方法相比,基于深度学习的目标检测算法可以通过数据集自主学习目标相关特征,鲁棒性更强。

目前,主流目标检测的算法包括了单阶段检测算法

与两阶段检测算法。单阶段检测算法往往检测速度快但精度不高,而两阶段检测算法则具有较高的检测精度但需要较长的检测时间。对于单阶段检测算法如YOLO^[15]、SSD^[16]等方法在车牌检测任务上的研究,Xie等^[17]将基于YOLO框架的定位算法用于车牌检测与识别研究且达到了较高识别准确率。刘华春等^[18]设计了一种基于视觉显著性的深度学习车牌检测算法用于解决复杂环境下的车牌检测问题。Molina-Moreno等^[19]提出了一种有效的尺度自适应车牌识别系统,该系统对每个尺度目标能选择最优特征进行训练并显著提高检测时间。

对于两阶段检测算法在车牌检测任务上的研究, Ren等^[20]提出了Faster R-CNN。随后,李祥鹏等^[21]在Faster R-CNN的基础上,结合 k -means++方法选取最优车牌尺寸来提升车牌检测精度。此类方法核心思想是通过候选区域生成的算法生成含有检测目标的候选区域后,再对其进行分类与回归的处理。He等^[22]在Faster R-CNN的基础上,通过添加掩码分支,提出了Mask R-CNN。该网络在第一阶段利用残差网络(residual neural network, ResNet)与特征金字塔网络(feature pyramid network, FPN)两种网络相结合的骨干网络对图像进行特征提取。随后将特征图送入候选区域生成网络(region proposal network, RPN)生成候选区域。第二阶段,感兴趣区域匹配层(RoIAlign)的输出被送入由分类、回归与掩码三个通道组成的头网络,该网络负责输出定位的结果同时分割掩码。然而,Mask R-CNN作为一种通用目标检测模型,缺点在于无法根据车牌先验信息对图像中的车牌目标进行针对性的检测。

目前来讲,虽然传统方法与深度学习的方法已取得一定成果,但是复杂背景、非均匀光照与恶劣天气(例如雪天与雾天)等干扰因素仍会对车牌检测精确度造成一定影响。赵伟等^[23]为了解决复杂光线的干扰因素所导致车牌识别困难的问题,提出一种针对复杂光照的图像增强方法,以提高车牌识别准确率。

本文为了进一步提高复杂环境中车牌检测的精度,提出了一种结合双金字塔特征融合与级联定位的复杂环境下车牌检测算法。该算法通过在残差网络中使用Mish激活函数,结合提出的双金字塔特征融合网络以充分利用不同尺度的车牌特征,采取基于形状先验的锚点设置策略的设计与引入阶段化训练的级联定位网络的多策略使得本文算法能够在不同干扰情况下对车牌进行更准确与更具有针对性的检测。实验结果证明,本文算法在AOLP与CCPD车牌数据集上能进一步提升检测性能。

2 本文算法概述

本文提出算法框架如图1所示。整个网络结构分为三个部分:双金字塔特征融合网络、基于形状先验的锚点设置网络与级联车牌定位网络。车牌图像通过双金字塔特征融合网络(siamese feature pyramid network,

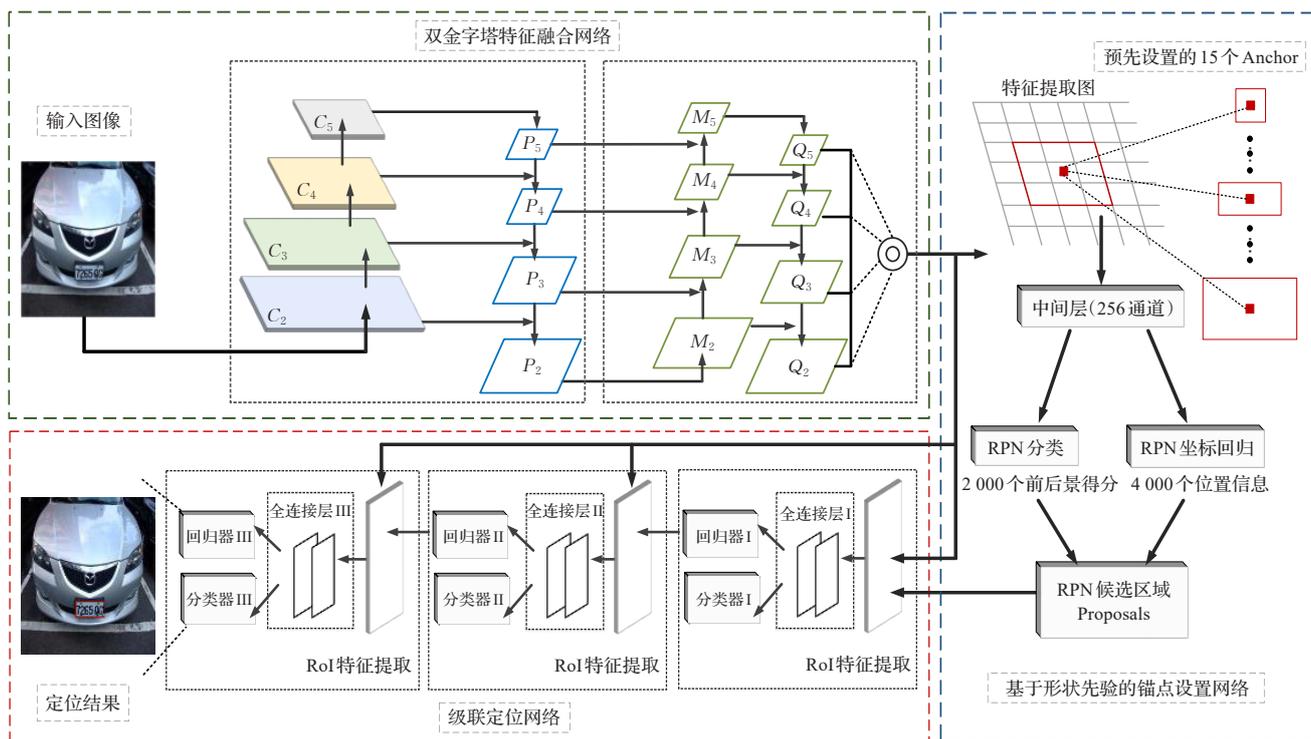


图1 本文算法的框架图

Fig.1 Framework of proposed algorithm

SFPN)提取初级图像特征并进行多尺度特征的融合,该网络的输出被送入基于形状先验的锚点设置网络获取含车牌的感兴趣区域(region of interest,RoI)的分类置信度与位置信息。最后,感兴趣区域将通过具有三个阶段的级联定位网络进行阶段化的分类和回归检测分支操作后输出最终检测信息。

2.1 双金字塔特征融合网络

在传统的FPN网络基础上,再增添了一组自下而上再自上而下的 n 型路径,所提出的具有两条 n 型路径的网络命名为双金字塔特征融合网络(SFPN)。由于本文所提出的网络与 ResNet+FPN 结构具有一定相似性,本文称原有结构为一级金字塔特征融合网络,新增的结构为二级金字塔特征融合网络。

2.1.1 一级金字塔特征融合网络

(1) ResNet101-M 网络

残差网络 ResNet 是 He 等^[24]在 2015 年提出的一种卷积神经网络,它通过其内部相当于跳跃连接的残差结构使用,同时在残差结构中加入 BN 层能够一定程度上解决深层神经网络存在的梯度弥散与爆炸的问题。

传统 Mask R-CNN 的特征提取网络由残差网络 ResNet 与特征金字塔网络 FPN 组建而成。在一级金字塔特征融合网络中,本文使用相同的方法对车牌进行特征提取。而在特征提取时将 Mish^[25]激活函数引入到传统的残差网络中替换了原先所使用的 ReLU 激活函数,并将引入 Mish 激活函数后的残差网络 ResNet101 记为 ResNet101-M。

相比于传统残差网络中使用 ReLU 作为激活函数,

Mish 激活函数在 $x=0$ 附近相对平滑。对于正的值, Mish 函数避免了由于封顶而导致的饱和。而对于负的值, Mish 函数的平滑性能更好处理梯度流。网络层数增加后, ReLU 的精度下降较快,而 Mish 函数则可以更好地保持准确性。ReLU 和 Mish 激活函数的数学表达式如(1)和式(2)所示:

$$f_1(x) = \max(0, x) \tag{1}$$

$$f_2(x) = x \times \tanh(\ln(1 + e^x)) \tag{2}$$

引入 Mish 激活函数后的残差网络 ResNet101 包括五个残差模块,分别记为 $C_1 \sim C_5$, 每个模块均包含不等的残差结构。除 C_1 外,每个残差结构均进行 3 次卷积操作,卷积核大小分别为 1×1 、 3×3 和 1×1 。其中 C_2 包含 3 个残差结构,卷积核数量分别为 64、64 和 256; C_3 包含 4 个残差结构,卷积核的数量分别为 128、128 与 512; C_4 包含 23 个残差结构,卷积核数量分别为 256、256 和 1 024; C_5 包含 3 个残差结构,卷积核数量分别为 512、512 和 2 048。整个网络共由 101 层构成,包含 100 个卷积层与 1 个池化层。

本文所使用的 ResNet101-M 网络具体的相关参数如表 1 所示,在网络结构参数设定上与传统残差网络 ResNet101 保持了一致性并在残差网络每个阶段的残差结构中使用了 Mish 作为激活函数,输入图像经过每一个阶段均得到较上阶段尺寸减半的特征图。由此完成了一级金字塔特征融合网络结构中的自底向上的过程。

(2) 特征金字塔网络

特征金字塔网络 FPN 由 Lin 等^[26]提出,被研究者们用于检测不同尺度目标。FPN 网络通过构造一条自下

表1 ResNet101-M网络结构
Table 1 Network structure of ResNet101-M

结构	对应参数
Input	—
Convolution1+Mish	卷积层: 卷积核:7×7×64,步长:2
	池化层: 卷积核:3×3×64,步长:2
Convolution2+Mish	残差模块:3
	残差结构操作: 卷积核:1×1×64,步长:1
	卷积核:3×3×64,步长:1
	卷积核:1×1×256,步长:1
Convolution3+Mish	残差模块:4
	残差结构操作: 卷积核:1×1×128,步长:2
	卷积核:3×3×128,步长:1
	卷积核:1×1×512,步长:1
Convolution4+Mish	残差模块:23
	残差结构操作: 卷积核:1×1×256,步长:2
	卷积核:3×3×256,步长:1
	卷积核:1×1×1024,步长:1
Convolution5+Mish	残差模块:3
	残差结构操作: 卷积核:1×1×512,步长:2
	卷积核:3×3×512,步长:1
	卷积核:1×1×2048,步长:1

而上和自上而下的 n 型网络路径,再利用侧向连接使各层特性相融合,从而能够在多尺度目标检测任务中带来显著性能提升。

ResNet101-M网络采用自下而上的路径得到的不同尺度的特征图 $C_2 \sim C_5$ 。FPN网络结构将ResNet101-M网络最高层特征图 C_5 使用大小为 1×1 的卷积核改变通道数为256后得到新特征图 P_5 。随后对 P_5 进行2倍上采样处理,将其与经过 1×1 卷积层处理后的通道数为256的 C_4 进行侧向连接得到新特征图 P_4 ,从而将高层特征与低层特征相融合使高层特征得到增强。

基于上述思想,将 $P_i (3 \leq i \leq 5)$ 进行2倍上采样后,再对 $C_{i-1} (3 \leq i \leq 5)$ 进行 1×1 卷积层处理,最后对 P_i 与 C_{i-1} 进行横向连接融合得到 $P_{i-1} (3 \leq i \leq 5)$,由此完成了特征提取骨干网络结构中自上向下路径的过程(对特征图进行 1×1 卷积操作是为了通过元素相加进行侧向连接操作来获得相同特征图的尺寸)。

2.1.2 二级金字塔特征融合网络

对于二级网络结构中的上升沿路径, P_2 先通过 1×1 的卷积层操作得到新特征图 M_2 ,然后 M_2 利用 3×3 大小且步长2的卷积核下采样处理,再与 1×1 卷积层处理后的 P_3 进行侧向连接得到新特征图 M_3 。以此类推,将 $M_i (2 \leq i \leq 4)$ 进行下采样提取特征后与进行同通道数的 $P_{i+1} (2 \leq i \leq 4)$ 进行侧向连接后得到了 $M_{i+1} (2 \leq i \leq 4)$ 。

通过构建上升沿路径所得新特征图 $M_2 \sim M_5$ 已经具有了高层和底层相融合的特征信息,在此基础上进一步融合不同尺度的特征图得到更多信息。对于下降沿路径,具体操作与原始FPN网络进行自上而下的过程一致,对 M_5 用 1×1 的卷积层过滤得到新的特征图 Q_5 ,将 $Q_i (3 \leq i \leq 5)$ 进行2倍上采样之后,再与进行 1×1 卷积层处理后 $M_{i-1} (3 \leq i \leq 5)$ 进行横向连接得到 $Q_{i-1} (3 \leq i \leq 5)$ 。最后对 $Q_2 \sim Q_5$ 做 3×3 卷积操作消除上采样混叠效应后将其送入基于形状先验的锚点设置网络。特征图融合路径的具体融合公式如式(3)~(6)所示:

$$P_i = \begin{cases} conv_{1 \times 1}(C_i), i = 5 \\ Upsampling(P_{i+1}) \odot conv_{1 \times 1}(C_i), 2 \leq i \leq 4 \end{cases} \quad (3)$$

$$M_i = \begin{cases} conv_{1 \times 1}(P_i), i = 2 \\ conv_{3 \times 3}(M_{i-1}) \odot conv_{1 \times 1}(P_i), 3 \leq i \leq 5 \end{cases} \quad (4)$$

$$Q'_i = \begin{cases} conv_{1 \times 1}(M_i), i = 5 \\ Upsampling(M_{i+1}) \odot conv_{1 \times 1}(M_i), 2 \leq i \leq 4 \end{cases} \quad (5)$$

$$Q_i = conv_{3 \times 3}(Q'_i), 2 \leq i \leq 5 \quad (6)$$

其中, P_i 表示原始FPN网络中与残差网络融合后的特征图, M_i 表示双金字塔特征融合网络中新增自下而上的路径中的特征图, Q'_i 表示为双金字塔特征融合网络中与 M_i 进行融合后的特征图, Q_i 表示对 Q'_i 卷积操作消除上采样混叠效应后的特征图。 $conv(\cdot)$ 表示对特征图进行卷积操作。 $Upsampling(\cdot)$ 表示对特征图进行2倍上采样卷积操作。 \odot 则表示对同通道数的特征图进行通道方向上的横向拼接操作。

本文中双金字塔特征融合网络结构如图2中所示,它采用了双金字塔特征融合网络结构来替代原先网络所用ResNet与FPN所组成的特征提取骨干网络。

传统特征金字塔网络在进行特征融合时往往存在两个弊端。其一,低层级的特征与高层级的特征融合路径过长所导致准确定位难度较高、低层级的特征在整体融合过程中产生的作用较弱;其二,网络输出到下一阶段的特征图中仅包含本层级与高一层级的特征,对低于自身层级的特征利用不足。

针对上述问题,本文提出的双金字塔特征融合网络,通过在特征提取与融合阶段加入了二级特征金字塔融合网络,使得特征融合时网络结构中不仅仅具有自底向上的侧向连接路径,还具有自上而下的侧向连接路径。这种改进缩小了较低层级特征与顶部层级特征间的融合路径,提升了低层级的特征在整个特征层级中所发挥的作用。因而,最终双金字塔特征融合网络所输出的特征图能同时包含自身层级、更高层级和更低层级三种不同的特征信息。如此设计能高效率地利用不同尺度特征中的信息,让底层的位置信息与高层的语义信息得以充分利用,从而进一步提高了网络检测的准确率。

2.2 基于形状先验的锚点设置网络

2.2.1 候选区域生成网络

Faster R-CNN网络中采用候选区域生成网络替换

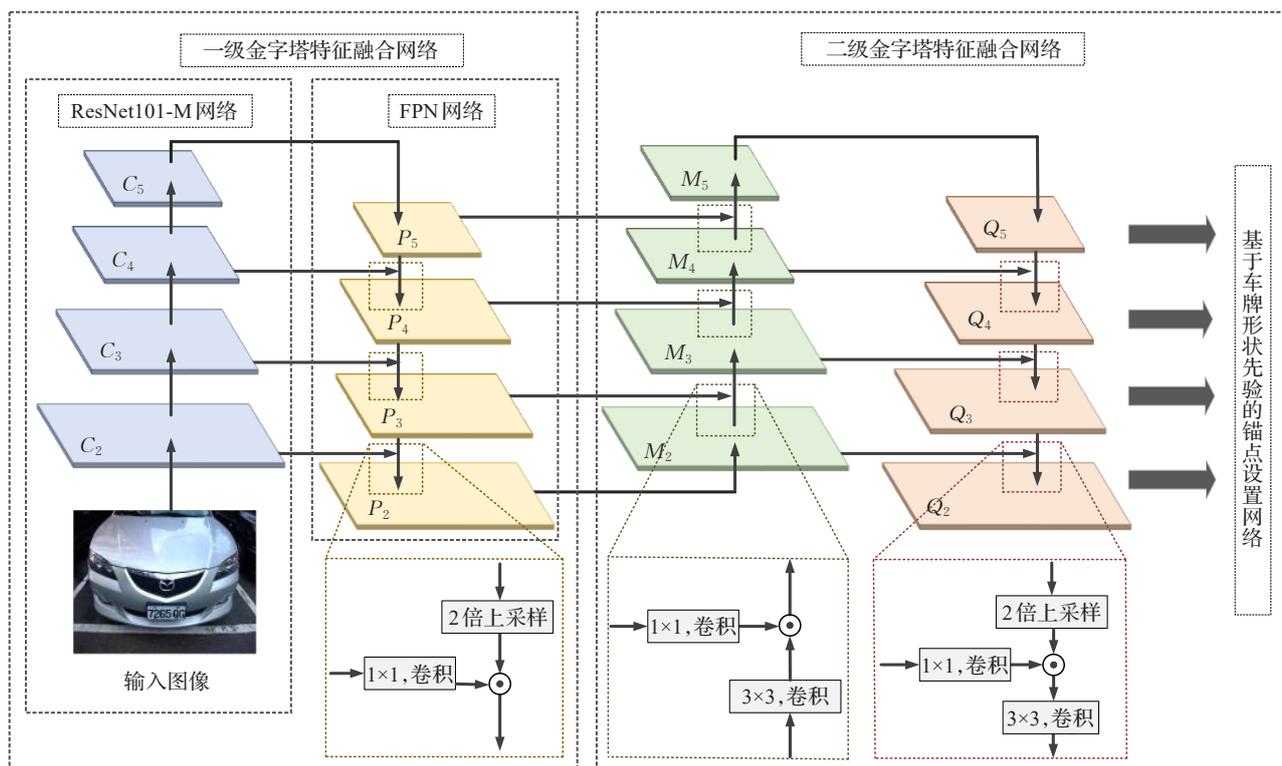


图2 双金字塔特征融合网络框架

Fig.2 Framework of SFPN

了Fast R-CNN网络所使用的候选区域生成方法,这种设计使得Faster R-CNN在生成候选区域的计算效率提升了一个数量级,在本文算法中同样采用结合锚框机制的候选区域生成网络来生成候选区域。其中锚框机制采用预先设定好的宽高比例与锚点尺寸来确定初始候选区域。当骨干网络不同尺度的特征图被作为输入送入RPN网络后,对于该特征图中的每一个像素均会生成 k 个候选区域。而网络会根据非极大值抑制算法(non-maximum suppression, NMS)对候选区域进行筛选之后,保留含有前景的候选区域。候选区域生成网络原理如图3所示。

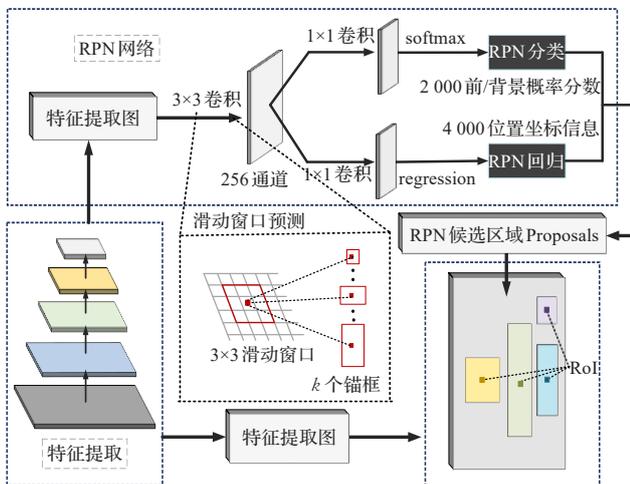


图3 候选区域生成网络结构

Fig.3 Structure of RPN

2.2.2 基于形状先验的锚点设置

在传统Mask R-CNN所用的候选区域生成网络RPN中,锚框的尺寸与比例设定直接影响着候选区域的生成。原始的锚框设定是根据Pascal VOC2007数据集所设定的,为能够检测出更多形状的目标,Mask R-CNN将锚框尺寸设定为 $\{32^2, 64^2, 128^2, 256^2, 512^2\}$ 共5种尺寸与 $\{1:2, 1:1, 2:1\}$ 共3种比例,具有通用性。如图4所示,本文算法的应用场景主要应用于复杂场景下的车牌检测。由于车牌的候选框大多为长条形状(倾斜角度较大时约1:1),因而原先的具普适性锚框比例设定对于车牌检测来说存在弊端。

由于原先锚框尺寸已经可以覆盖不同尺寸目标的车牌,仅宽高比例并不具有针对性。因此,基于车牌的形状学先验性,本文中算法在保持对传统Mask R-CNN的RPN网络中5种锚框尺寸设定一致的情况下,将原先3种 $\{1:2, 1:1, 2:1\}$ 比例优化为 $\{1:1, 2:1, 5:2\}$ 。在不改变原有锚框比例数量设定的基础上,通过增添一组横向的锚框比例替代竖向的锚框比例,使得最后RPN网络所生成的候选区域更为准确,并且可以降低计算量同时节约模型训练时间。图4所示为改进前后示意图。

2.3 级联定位网络

2.3.1 定位网络与级联结构

感兴趣区域匹配层用来提取候选区域的特征。它能够根据候选区域的位置信息从特征图上截取其相对应的位置,并通过双线性插值算法与最大值池化操作后将其池化为统一的 7×7 尺寸的特征图送入全连接层。



(a)原始锚点设置 (b)改进后锚点设置

图4 基于形状先验的锚点设置

Fig.4 Anchors based on shape prior

而在全连接层阶段,网络通过两个通道数为1 024的全连接层完成候选框的分类与回归两个分支的操作。整个定位网络利用提取的车牌特征为输入,输出定位到的车牌位置信息。

为得到更加高质量的检测结果,本文在原有网络结构中引入级联定位机制。新引入的级联定位网络共由三个阶段性的定位网络组成。在级联定位网络中,每个定位网络具有相同的结构。而对于所有定位网络按阶段依次设置0.5、0.6与0.7共3种不同的IoU阈值来划分正负样本,将每个阶段定位网络的输出作为下个阶段定位网络的输入从而使得级联后定位网络输出结果的精度获得逐步的提升。基于上述思想,不同于原始网络仅用0.5IoU阈值作为正负样本界定标准,引入后的新定位网络会每次对回归边框都提高一定的精度,通过调整回归边框为下一阶段找到一个IoU更高的正样本,以实现提高检测框精确度的效果。

2.3.2 损失函数

在模型训练阶段,本文方法共使用了四个损失函数来评估模型预测值与真值间的误差。总损失函数如式(7)所示:

$$L = L_{cls_r} + L_{bbr_r} + \sum_{i=1}^3 (L_{cls_i} + L_{bbr_i}) \quad (7)$$

其中, L_{cls_r} 与 L_{bbr_r} 分别表示基于形状先验的锚点设置网络训练时感兴趣区域RoI前/后景分类与位置坐标损失, L_{cls_i} 与 L_{bbr_i} 分别表示各阶段定位网络输出的分类及矩形框位置坐标的损失。

在基于形状先验的锚点设置网络的训练误差中, L_{cls_r} 使用二分类交叉熵损失,而 L_{bbr_r} 使用 Smooth-L1 损失。具体公式如式(8)与式(9)所示:

$$L_{cls_r} = -\frac{1}{N_{cls}} \sum_{i=1}^{N_{cls}} \ln[p_i p_i^* + (1-p_i)(1-p_i^*)] \quad (8)$$

$$L_{bbr_r} = \frac{1}{N_{reg}} \sum_{i=1}^{N_{reg}} p_i^* R(t_i, t_i^*) \quad (9)$$

式(8)中,网络会将产生的锚框划分为前景(包含车牌目标)与后景(不包含车牌目标)共两类。 N_{cls} 表示训练选择的锚框数量, p_i 表示训练时锚框被预测为前景的概率, p_i^* 为真值标签,锚框为正样本标签值为1,反之则标签值为0。

式(9)中, N_{reg} 表示特征图中的像素数量, t_i 表示网

络训练锚点预测的坐标偏移量, t_i^* 表示网络训练锚点较真值的实际坐标偏移量, $R(\cdot)$ 为 Smooth-L1 函数。

感兴趣区域匹配层会将统一尺寸的感兴趣区域输出到全连接层的最终分类与矩形框回归的分支。该阶段的损失与网络训练阶段相似,具体相关计算公式如式(10)与式(11)所示:

$$L_{cls_i} = -\frac{1}{N_{cls}} \sum_{k=1}^{N_{cls}} \ln[p_k p_k^* + (1-p_k)(1-p_k^*)] \quad (10)$$

$$L_{bbr_i} = \frac{1}{N_{reg}} \sum_{k=1}^{N_{reg}} p_k^* R(t_k, t_k^*) \quad (11)$$

其中, N_{cls} 表示预测目标数量, N_{reg} 表示矩形框数量, p_k 表示训练时目标预测为正样本的概率, p_k^* 为真值标签,目标预测区域为正样本标签值为1,反之则标签值为0。

3 车牌检测实验及结果分析

3.1 实验数据集

为了对本文算法进行性能方面的评估,本文共使用了两种车牌数据集来测试算法在车牌检测中的性能。为更好地验证本车牌检测算法在复杂背景、非均匀光照条件与恶劣天气等干扰情况下的精确性与鲁棒性,实验使用的 AOLP 数据集中包括了水平角度与不同的倾斜角度车牌,同时还涉及到城市交通复杂背景下的车牌。CCPD 数据集中样本则用于测试在雨天、雪天与雾天恶劣天气、非均匀光照条件与拍摄抖动模糊的干扰情况下车牌检测的精确度。车牌数据集的详细介绍如下。

(1) AOLP 数据集

AOLP^[27]数据集是一个被广泛使用的开源公共车牌检测数据集。该数据集由2 049个台湾车牌图像样本构成。AOLP数据集一共分为3个子数据集:AC(access control)数据集、LE(law enforcement)数据集与RP(road patrol)数据集。上述数据集分别含681、757和611个样本。由于AC数据集样本车辆多是以受限行驶速度通过道口时被拍摄捕捉到,因而样本均为水平方向的车牌图像;LE数据集样本来自城市交通出行的车辆,包含行人、路灯与路标的复杂道路背景等的干扰因素;RP数据集样本大多数是倾斜的车牌图像。

本文实验将 AOLP 数据集的样本随机划分为训练集(85%)与测试集(15%)。为保证训练集与测试集中三个子数据集样本的分布一致性,对子数据集都按同样比例随机进行数据分割。AOLP数据集实验训练集为1 740个车牌图像,测试集为309个车牌图像。

(2) CCPD 数据集

CCPD(Chinese City Parking Dataset)数据集是一个公共中国车牌数据集。该数据集是由 Xu 等^[28]在 ECCV2018 上提出的一个全新的大型综合多样性国内车牌数据集,样本数量超过25万张。样本数据采集自国内各大停车场,所有图像样本分辨率均为720×1 160,且图像均拍摄于不同时间与光照条件(如白天、傍晚与

表2 车牌检测实验中不同数据集的详细介绍

Table 2 Detailed introduction of different datasets in license plate detection experiments

数据集	数据集描述	图像数量	分辨率	训练集数量	测试集数量	车牌类型
AOLP-AC	仅含有水平方向的车牌	681	240×352	578	103	台湾车牌
AOLP-LE	城市复杂道路背景下的车牌	757	480×640	643	114	台湾车牌
AOLP-RP	不同角度水平/垂直方向的倾斜车牌	611	240×320	519	92	台湾车牌
CCPD-Blur	拍摄时抖动模糊及清晰度差的车牌	20 611	720×1 160	16 611	4 000	大陆车牌
CCPD-Weather	雨天、雪天及雾天情况下的车牌	9 999	720×1 160	7 999	2 000	大陆车牌

夜晚),不同的气候环境(如雨、雪天)等挑战性的复杂环境。

出于验证算法对不同干扰情况下的表现,本文实验在CCPD车牌数据集上选择了CCPD中的CCPD-Blur与CCPD-Weather两个子数据集来分别验证本文方法复杂环境下车牌检测的性能与鲁棒性。其中CCPD-Blur数据集样本中存在抖动模糊与不均匀光照的干扰因素;CCPD-Weather数据集样本中存在极端天气条件的干扰因素。

实验共使用了五种具有不同特色的车牌数据集(如表2所示),由于不同的车牌数据集中的样本具有明显的差异性与挑战性,从而可以更好地验证本文方法在复杂环境下不同干扰时的车牌检测效果,进一步验证本文改进算法的可行性。

3.2 实验评价指标

基于车牌目标检测与自然场景下通用文本检测应用场景的相似性,本实验使用与遵循文献[29]中所用的评价指标用作衡量本文算法的验证性能指标。本文实验中共使用了4种评价指标来评估算法在车牌数据集上的检测性能,包括了精确率(Precision),召回率(Recall),F值(F_1 -Measure)与图像平均处理时间 T 。

精确率(Precision)定义为检测算法所预测为车牌样本且预测准确的样本数量占算法预测的样本数量的比率;召回率(Recall)定义为检测算法所预测为车牌样本且预测准确的样本数量占测试集总样本数量的比率;F值(F_1 -Measure)是综合检测性能评价指标,通过对精确率与召回率综合计算来衡量算法总体性能,该数值越大则表明算法性能越优异;图像平均处理时间 T 定义为本文中算法处理单幅车牌图像的平均时间。上述评价指标具体计算方式如式(12)、(13)所示:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F_1 - Measure = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (13)$$

其中, TP 表示车牌样本被检测为正样本且实际也为正样本的样本数量; FP 表示车牌样本被检测为正样本但实际为负样本的样本数量; FN 表示车牌样本被检测为负样本但实际为正样本的样本数量。

同时,在实验中定义交并比(intersection over union, IoU)作为界定车牌正/负样本的阈值标准。交并比IoU的具体计算公式如式(14)所示:

$$IoU = \frac{R_{pred} \cap R_{gt}}{R_{pred} \cup R_{gt}} \quad (14)$$

其中, R_{pred} 表示为训练模型所检测到的边界框面积, R_{gt} 表示数据标注真值的边界框面积。当IoU数值在实验中被设置0.5时,对于一个车牌区域,模型检测到的边界框与真值边界框的IoU大于50%,则判定为正样本,反之为负样本。在本实验中,采用0.5、0.6与0.7共3种不同IoU阈值来分别测定训练本文提出的网络在两种车牌数据集上的车牌检测性能。

3.3 实验细节

车牌检测实验环境为: Intel i7-9700k 3.6 GHz CPU, NVIDIA GeForce RTX2080 GPU,实验所用的操作系统为Windows10 x64,所用的深度学习框架为TensorFlow 1.9.0,所用开发语言为Python 3.6.9,编译环境为PyCharm 2019。

本文实验中所构建网络模型采用MS COCO预训练模型,学习率参数设置为0.000 75,批量参数设置为2,优化策略为随机梯度下降法,动量参数设置为0.9, L_2 正则化权重衰减系数设置为0.000 1,训练最大迭代次数共90 000次。基于车牌形状先验的锚点设置网络训练时的锚框筛选阈值为0.7。车牌检测实验中所涉及的相关参数具体设置如表3所示。

表3 实验所涉及的相关参数具体设置

Table 3 Relevant parameters involved in experiments

参数	相关设置
网络预训练模型	MS COCO
学习率	0.000 75
批量	2
动量参数	0.9
权重衰减	L_2 正则化
权重衰减系数	0.000 1

3.4 不同改进策略消融实验分析

本文提出的模型由3个部分组成,包括了双金字塔特征融合网络、基于形状先验的锚点设置网络和级联定位网络。

3.4.1 实验训练步骤

本文方法在模型训练阶段的具体步骤如下:

输入:复杂环境下的车牌图像。

步骤1 将输入图像统一为512×512尺寸。

步骤2 将缩放后的车牌图像送入双金字塔特征融

合网络进行特征提取。输入图像通过一级特征融合网络提取初级特征 $P_2 \sim P_5$ 。

步骤3 将步骤2中的特征图 $P_2 \sim P_5$ 送入二级特征融合网络进行特征增强得到特征图 $Q_2 \sim Q_5$ 。

步骤4 根据预先设置的基于车牌形状的锚框尺寸与比例,基于形状先验的锚点设置网络会对步骤3得到的特征图中的每个点均生成相对于原图像的3种不同比例的锚框。

步骤5 基于形状先验的锚点设置网络对每个锚框预测前/后景得分与位置偏移后对锚框进行修正并筛选出部分锚框作为候选区域。

步骤6 根据损失函数基于形状先验的锚点设置网络,计算锚框与真值的IoU,并按IoU阈值分别筛选等量的正负样本进行训练。

步骤7 根据损失函数训练级联定位网络,计算步骤5中候选区域与真值的IoU,并按IoU阈值分别为0.5,0.6与0.7界定筛选200个正负训练样本后送入感兴趣区域匹配层。该层将候选区域的坐标映射回对应的特征图后进行特征图尺寸统一后送入回归器与分类器,分别得到其回归位置偏移与分类得分。

3.4.2 实验重要过程与可视化分析

对于3.4.1节中实验训练阶段步骤中重要过程结果,本文在本小节进行更深入的实验过程与可视化分析。

对于本文方法在特征提取阶段所使用的双金字塔特征融合网络,表4中详细地列出了双金字塔特征融合网络结构的相关参数。

对于本文所提出的级联定位网络,图5给出了不同阶段下的级联定位网络在车牌定位的质量上的视觉比较。为了更直观地比较级联定位网络对于车牌检测结果在精细程度上的修正情况,图5对于每一幅车牌检测的结果均给出了检测边界框与真实标注框的IoU数值(用于评估检测边界框与真实标注框的接近程度,IoU

表4 双金字塔特征融合网络的相关参数

Table 4 Relevant parameters in SFPN

单元结构	特征图	输出尺寸
一级特征融合网络(Input)	输入	512×512×3
一级特征融合网络(Convolution1+Mish)	C_1	128×128×64
一级特征融合网络(Convolution2+Mish)	C_2	128×128×256
一级特征融合网络(Convolution3+Mish)	C_3	64×64×512
一级特征融合网络(Convolution4+Mish)	C_4	32×32×1 024
一级特征融合网络(Convolution5+Mish)	C_5	16×16×2 048
一级特征融合网络(FPN网络)	P_5	16×16×256
	P_4	32×32×256
	P_3	64×64×256
	P_2	128×128×256
	M_2	128×128×256
二级特征融合网络(上升沿路径)	M_3	64×64×256
	M_4	32×32×256
	M_5	16×16×256
	Q_5	16×16×256
	Q_4	32×32×256
二级特征融合网络(下降沿路径)	Q_3	64×64×256
	Q_2	128×128×256

数值越大表示检测结果越接近真值,即车牌检测结果越精细)。

在网络训练阶段,原有定位网络在训练阶段会将3.4.1小节步骤5中生成的候选区域先计算与真值的IoU后按0.5的IoU阈值来界定正负训练样本。较之原有定位网络的单阶段训练,定位网络在引入级联定位的改进策略后会将定位网络的训练分为三个阶段来训练网络的分类器与回归器。这种级联训练能够以阶段化提高界定正负样本的IoU阈值的方式来进一步筛选高质量的候选区域,从而使得最终检测结果更精细。

由图5给出的检测结果比较可以看出,三级定位网络的检测结果在不同情况下均比前两级定位网络的检测结果更精细。可以发现,图5(a)、(b)与(c)中检测结果的IoU值呈现依次递增的趋势,说明级联网络不同阶



图5 不同阶段性级联定位网络的定位效果对比

Fig.5 Comparison of detection results of cascaded positioning network in different stages

段能够对检测结果在精细程度上进行修正。在第2行的样例中,前两级定位网络的检测结果边缘均不够细致,而在第三级定位网络中得到了接近真值的修正。对于第4行的样例,前两级定位网络均未能完整检测出整个车牌区域,会影响后续对数字4的识别,得益于在第三级定位网络的修正,最终检测结果能够完整地检测车牌区域。

结合上述分析,原有定位网络在引入级联训练的改进策略后能够对车牌检测的精细程度进行修正,每经过一个阶段都可以得到一个比上阶段精度更高的结果,最终获得高质量检测回归框,从而证明本文算法引入级联定位策略具有一定的成效性。

3.4.3 消融实验结果和分析

本文算法包括了4种不同改进策略,包括双金字塔特征融合网络、Mish激活函数、基于车牌形状先验的锚

点设置网络与级联定位网络。为了验证本文所提出的各种策略的有效性,在AOLP数据集上进行了训练并在测试集上进行了5组消融实验。

使用不同改进策略的消融实验性能统计与车牌检测示例对比如表5与图6所示,以下分别对消融实验进行定量与定性的评估与分析。

通过表5对实验中不同改进策略后算法的精确率与召回率结果观察对比可以发现,在不同IoU阈值下实验所用的4种改进策略对算法的检测性能均有一定提升,其中精确率提升较明显。消融实验中不同结果均在低IoU阈值下保持较高的召回率。

第一组实验中,使用传统Mask R-CNN作为基准方法(Baseline)在AOLP数据集上进行测试,不同IoU阈值下的精确率均为97.47%,而F值均为98.56%。基于

表5 不同改进策略在AOLP数据集的消融实验

Table 5 Ablation experiments using different improvement strategies on AOLP

所用改进策略	精确率			召回率			F值			用时/ms	模型/MB
	0.5IoU	0.6IoU	0.7IoU	0.5IoU	0.6IoU	0.7IoU	0.5IoU	0.6IoU	0.7IoU		
Baseline	0.974 7	0.974 7	0.974 7	0.996 7	0.996 7	0.996 7	0.9856	0.985 6	0.985 6	270	241
Baseline+ ①	0.980 9	0.980 9	0.980 9	0.996 7	0.996 7	0.996 7	0.988 8	0.988 8	0.988 8	275	265
Baseline+ ①②	0.984 0	0.984 0	0.984 0	0.996 7	0.996 7	0.996 7	0.9903	0.990 3	0.990 3	278	265
Baseline+ ①②③	0.990 3	0.990 3	0.990 3	0.996 7	0.996 7	0.993 5	0.993 5	0.993 5	0.990 3	271	265
Baseline+ ①②③④	1.000 0	1.000 0	1.000 0	0.996 7	0.996 7	0.996 7	0.9983	0.998 3	0.998 3	377	371

注:Baseline代表传统Mask R-CNN; ①代表双金字塔特征融合网络; ②代表特征提取中Mish激活函数; ③代表基于车牌形状先验的锚点设置网络; ④代表引入级联定位网络。

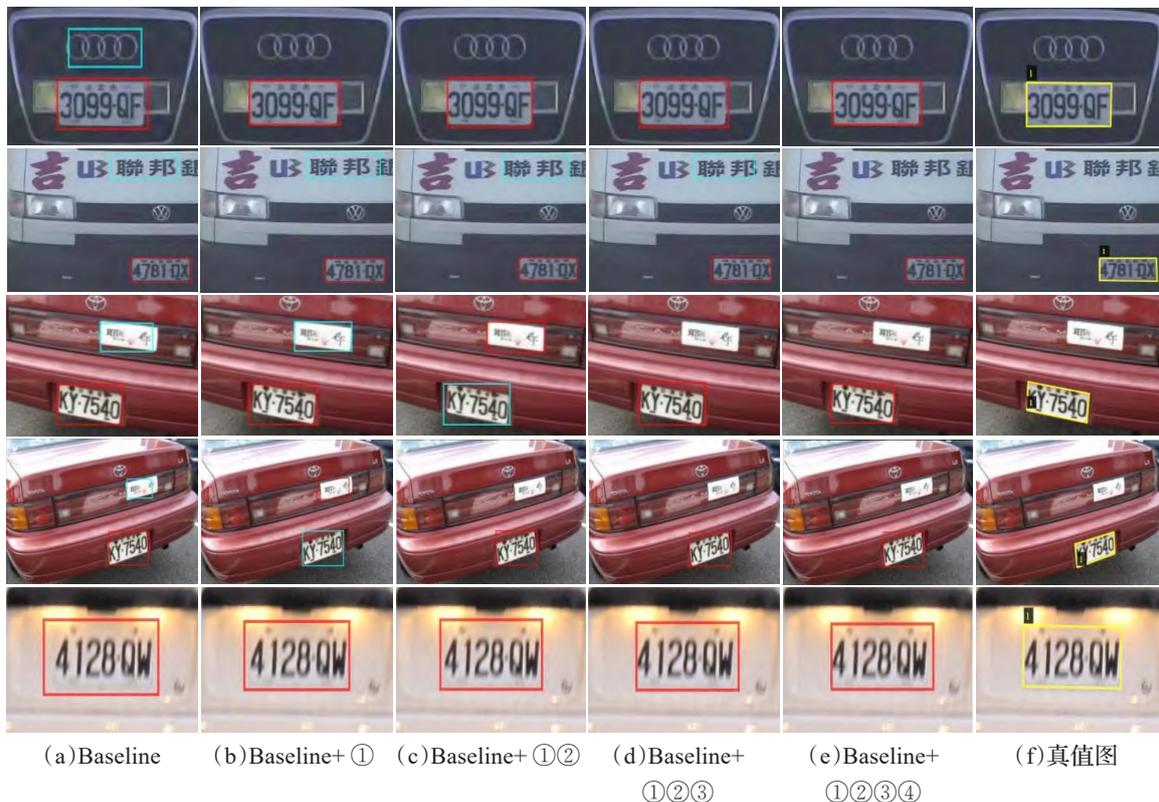


图6 使用不同改进策略消融实验的检测结果对比

Fig.6 Comparison of detection results inablation experiments using different improvement strategies

Baseline的车牌检测示例如图6(a)所示,可以发现Baseline模型的检测结果存在大量误检的问题,模型会将车标与其他干扰性文字等非车牌区域误判为车牌区域(例如图6第1~4行)。此外,该方法中车牌位置的检测也不够准确。

第二组实验中,在Baseline的基础上加入了双金字塔特征融合网络并进行测试,不同IoU判定下的精确率均为98.09%, F 值为98.88%。可以发现在特征提取阶段添加双金字塔融合网络后,多层车牌特征能够充分融合从而有效提升模型精度(0.62个百分点),且不会增加过多检测时间。基于Baseline+①模型的车牌检测示例如图6(b)所示。该组示例仍存在不少误检的情况,但较之Baseline,该模型对一些干扰情况的误判已经减少。与此同时,该方法的定位准确度也有一定提升。由此可见,引入双金字塔特征融合网络具有不错的效果。

第三组实验中,本文在Baseline+①模型中加入了Mish激活函数。在AOLP测试集上不同IoU阈值判定下取得的精确率均为98.40%, F 值均为99.03%。由此可见,通过在特征提取中引入Mish激活函数可以提升网络的分类精度,使得模型较Baseline+①在精确率上提升了0.31个百分点, F 值上提升了0.15个百分点。基于Baseline+①②模型的车牌检测示例如图6(c)所示。该组示例中对于非车牌区域的误检率相比第二组实验进一步减少。实验结果表明,引入Mish激活函数能够有效提高分类精度。

第四组实验中,本文在Baseline+①②模型的基础上增加了基于形状先验的锚点设置网络并在AOLP数据集上进行测试。表5的定量评估显示,通过将通用锚框优化为符合车牌检测任务的锚框后模型较Baseline+①②在低IoU上提升了0.32个百分点的精确率。基于Baseline+①②③模型的车牌检测示例如图6(d)所示。该方法的检测结果在分类精确度上的表现比之前模型更优异。

第五组实验中,本文在Baseline+①②③模型中引入级联定位网络在AOLP数据集上进行测试,在0.5、0.6与0.7IoU下的精确率均为100.00%, F 值均为99.83%。其车牌检测示例如图6(e)所示,不难看出网络在引入级联定位策略后的分类精度与定位精细度均有较大的提高,且车牌定位更贴合真值。基于上述分析,本文方法能够在不同IoU阈值下进一步提升车牌检测性能,所用改进策略均具有成效。

同时,本文结合表5中数据对所提出算法结合模型参数与所用时间进行如下分析:对于检测速度而言,传统Mask R-CNN测试单幅图像平均耗时270 ms。引入双金字塔融合网络后,在特征提取阶段运算量有所增加,模型参数由241 MB提升到了265 MB,平均耗时增加5 ms;在引入Mish激活函数后,模型参数保持不变,

平均耗时增加3 ms。可以看出,策略①②在提升模型检测精度的同时,由于运算量增加导致模型参数与检测时间有所增加,但增幅不大。在优化RPN中的锚框比例后,模型参数保持不变,平均检测时间由278 ms降低至271 ms,进一步验证优化锚框比例后由于锚框比例对于车牌形状的实用性从而使得模型计算量有所降低。最后,在之前模型基础上引入了级联定位策略,改进后的模型(即本文算法)在总体精度性能上提升幅度较之前策略都要明显,不同IoU下精确度提升高达0.97个百分点, F 值上提升高达0.48个百分点。但在单幅检测时间上增加了106 ms,在模型参数上增加了106 MB。由此可见,级联定位策略在被引入后模型在检测时会按三个不同阶段性迭代检测优化精度,从而导致整体检测时间与模型参数的增加。对于改进后的算法虽然提升了较多精度,但也额外提升了计算量,对于单幅车牌的平均检测时间保持在377 ms,可满足实际应用需求。

3.5 与相关方法的性能对比分析

除消融实验外,还将本文算法与近年来其他使用AOLP数据集的文献方法进行比较。表6~8分别给出了0.5IoU阈值下不同算法在AC、LE与RP三个AOLP子数据集上的检测结果比较情况。

表6 不同算法在AOLP-AC数据集结果比较

Table 6 Comparison of different algorithms results on AOLP-AC

算法	精确率	召回率	F 值
文献[27]	0.910 0	0.960 0	0.934 3
文献[30]	0.985 3	0.983 8	0.984 5
文献[31]	0.926 0	0.968 0	0.946 5
文献[32]	0.993 0	0.994 0	0.993 5
本文算法	1.000 0	0.994 1	0.997 0

表7 不同算法在AOLP-LE数据集结果比较

Table 7 Comparison of different algorithms results on AOLP-LE

算法	精确率	召回率	F 值
文献[27]	0.910 0	0.950 0	0.929 6
文献[30]	0.977 5	0.976 2	0.976 8
文献[31]	0.935 0	0.933 0	0.933 9
文献[32]	0.992 0	0.992 0	0.992 0
本文算法	0.994 7	0.988 1	0.991 4

表8 不同算法在AOLP-RP数据集结果比较

Table 8 Comparison of different algorithms results on AOLP-RP

算法	精确率	召回率	F 值
文献[21]	0.988 9	—	—
文献[27]	0.910 0	0.940 0	0.924 8
文献[30]	0.952 8	0.955 8	0.954 3
文献[31]	0.929 0	0.962 0	0.945 2
文献[32]	0.989 0	0.988 0	0.988 5
本文算法	0.995 1	0.993 4	0.994 2

对于该部分实验的数据集划分,本文遵循文献[30]中原则,即在表6中使用LE与RP数据集样本作为训练来评估算法在AC数据集上的检测性能,表7与表8以此类推。

通过表6~8中的检测结果比较可以看出,以0.5IoU阈值为统一评估标准,根据表6对于AC数据集,本文算法在精确率与召回率为100.00%和99.41%,本文方法明显优于其他算法的检测结果。

如表7所示,在LE数据集中本文算法的精确率为99.47%,召回率为98.81%, F 值为99.14%。精确率比文献[32]算法提升了0.27个百分点,召回率略低于文献[32]算法,实验结果表明本文方法为确保检出结果的准确性,选择牺牲了一定召回率。

而在表8的RP数据集的检测结果比较中,本文算法在精确率与召回率均表现优异,精确率和召回率均优于其他算法,整体的优越性相对明显。

除了在AOLP数据集上的3个子数据集与其他文献中不同算法进行对比实验外,还将本文提出的算法与SSD(文献[16]算法)、文献[17]算法、文献[19]算法、Faster R-CNN(文献[20]算法)、Mask R-CNN(文献[22]算法)、YOLOv3(文献[33]算法)在多个车牌数据集上进行了车牌检测实验,来评估本文算法在不同类型车牌检测场景下的优越性和有效性。

表9中给出了本文改进算法以及其他算法在不同类型车牌数据集上的对比结果。由表可见,本文算法与原Mask R-CNN(文献[22]算法)相比, F 值在AOLP数

据集上提升了1.27个百分点,在CCPD-Weather数据集上提升了0.83个百分点,在CCPD-Blur上提升了1.71个百分点。可以看出,改进后的算法在不同类型车牌数据集上比未改进前均有性能提高,其中对于清晰度较差的车牌数据有显著的提升。综合分析来看,本文算法与其他算法相比在精度上都具有优势,因而改进的算法模型具有一定的鲁棒性与泛化性。

表9 不同算法在不同数据集上的 F 值结果

Table 9 F -measure of different algorithms on different datasets

算法	数据集		
	AOLP	CCPD-Weather	CCPD-Blur
文献[16]	0.987 2	0.989 4	0.979 1
文献[17]	0.994 7	—	—
文献[19]	0.989 8	—	—
文献[20]	0.988 7	0.978 9	0.968 8
文献[22]	0.9856	0.988 4	0.972 0
文献[33]	0.995 0	0.995 6	0.986 5
本文算法	0.998 3	0.9967	0.989 1

除了表9给出的与其他算法的定量对比分析以外,图7中本文将所提出算法与其他算法进行了定性对比分析。其中,第1~5行分别代表了正常情况下的水平车牌、城市道路场景的车牌、停车场倾斜角度的车牌、雨雪极端天气下的车牌与光照不足而且抖动模糊的车牌共5种不同类型车牌检测场景的结果对比图。对于所提出的算法在不同类型车牌检测场景中与其他算法的有效性与优势的定性分析如下述:



图7 不同算法在不同车牌检测场景下的结果对比

Fig.7 Comparison of different algorithms detection effects in various license plate detection scenarios

由图7第1行可以看出,对于正常的水平车牌检测,本文算法与其他算法都可以正确检测到车牌的位置。但在准确性与细致性上,文献[16]算法、文献[20]算法和文献[33]算法表现较差,本文算法稍优于文献[22]算法。由图7第2行可以看出,对于城市复杂道路场景的车牌检测,文献[20]算法的回归框质量最低,文献[16]算法与文献[33]算法也不细致。文献[22]算法左侧部分未能完整定位整个车牌字符,因此该结果可能会影响后续识别的效果。而本文算法的回归框与真值框最贴合。由图7第3行可以看出,对于停车场倾斜角度的车牌检测,文献[16]算法、文献[20]算法与文献[33]算法的检测框与真值标注的切合程度均较差,本文算法与文献[22]算法在回归框质量上稍占据优势。由图7第4行可以看出,对于雨雪极端天气下的车牌检测,文献[20]算法的回归框两侧都检测到了多余背景区域,过于粗糙。文献[33]算法的回归框在左侧部分也检测了多余的背景。而本文算法与除上述两种算法以外的其他算法相比,在细腻程度上的优势更为显著。由图7第5行可以看出,对于光照不足而且抖动模糊的车牌检测,各个检测算法之间检测结果差距不大,本文算法稍优于其他4种算法。

综合上述定量与定性分析,实验结果表明,本文算法得益于双金字塔特征融合、车牌形状先验与级联定位等策略改进后较传统Mask R-CNN检测性能更优,回归质量更好。与其他不同算法的相比也有一定的鲁棒性,同时在不同类型的车牌检测场景下检测结果表现稳定,具有一定的泛化性。

4 结束语

实验证明,本文算法能够对车牌目标进行针对性的检测,同时结合双金字塔特征融合网络与使用级联结构定位的策略实现了高质量定位,使得检测更具有效性与精确性。今后,将尝试使用多边形来对车牌目标进行形状更贴合的检测以期达到更优检测效果。同时,如何在保证计算速度前提下实现高精度车牌检测是未来的研究方向。

参考文献:

- [1] ANAGNOSTOPOULOS C N E, ANAGNOSTOPOULOS I E, PSOROULAS I D, et al. License plate recognition from still images and video sequences: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2008, 9(3): 377-391.
- [2] COETZEE C, BOTHA C, WEBER D. PC based number plate recognition system[C]//Proceedings of the IEEE International Symposium on Industrial Electronics, 1998: 605-610.
- [3] PARISI R, CLAUDIO E D D, LUCARELLI G, et al. Car plate recognition by neural networks and image processing[C]//Proceedings of the IEEE International Symposium on Circuits and Systems, 1998: 195-198.
- [4] BARROSO J, DAGLESS E L, RAFAEL A, et al. Number plate reading using computer vision[C]//Proceedings of the IEEE International Symposium on Industrial Electronics, 1997: 761-766.
- [5] ZHENG D N, ZHAO Y N, WANG J X. An efficient method of license plate location[J]. Pattern Recognition Letters, 2005, 26(15): 2431-2438.
- [6] ZHOU W G, LI H Q, LU Y J, et al. Principal visual word discovery for automatic license plate detection[J]. IEEE Transactions on Image Processing, 2012, 21(9): 4269-4279.
- [7] ABOLGHASEMI V, AHMADYFARD A. An edge-based color-aided method for license plate detection[J]. Image and Vision Computing, 2009, 27(8): 1134-1142.
- [8] YU S, LI B, ZHANG Q, et al. A novel license plate location method based on wavelet transform and EMD analysis[J]. Pattern Recognition, 2015, 48(1): 114-125.
- [9] AL-GHAILI A M, MASHOHOR C, RAMLI A R, et al. Vertical edge-based car license plate detection method[J]. IEEE Transactions on Vehicular Technology, 2013, 62(1): 26-38.
- [10] CHO B K, RYU S H, SHIN D R, et al. License plate extraction method for identification of vehicle at a railway level crossing[J]. International Journal of Automotive Technology, 2011, 12(2): 281-289.
- [11] AZAD R, DAVAMI F, AZAD B. A novel and robust method for automatic license plate recognition system based on pattern recognition[J]. Advances in Computer Science, 2013, 2(3): 64-70.
- [12] DEB K, KHAN I, SAHA A, et al. An efficient method of vehicle license plate recognition based on sliding concentric windows and artificial neural network[J]. Procedia Technology, 2012, 4: 812-819.
- [13] 魏亨, 邱实, 李晨, 等. 计算机多尺度辅助定位车牌算法[J]. 电子学报, 2018, 46(9): 2188-2193.
WEI T, QIU S, LI C, et al. License plate detection algorithm based on computer multi scale assist[J]. Acta Electronica Sinica, 2018, 46(9): 2188-2193.
- [14] 葛艳, 陈晨. 基于文字定位和颜色检测的多车牌定位方法研究[J]. 计算机与数字工程, 2018, 46(3): 575-579.
GE Y, CHEN C. Research on multi-license plate location based on multi-algorithms[J]. Computer & Digital Engineering, 2018, 46(3): 575-579.
- [15] REDMON J, DIVVALA S, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of European Con-

- ference on Computer Vision. Heidelberg: Springer, 2016: 21-37.
- [17] XIE L L, AHMAD T. A new CNN-based method for multi-directional car license plate detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(2): 507-517.
- [18] 刘华春, 侯向宁. 采用视觉显著性和卷积网络的车牌定位算法[J]. 计算机技术与发展, 2019, 29(6): 74-78.
- LIU H C, HOU X N. A license plate location algorithm using visual saliency and convolutional network[J]. Computer Technology and Development, 2019, 29(6): 74-78.
- [19] MOLINA-MORENO M, GONZALEZ-DIAZ I, DIAZ-DE-MARIA F. Efficient scale-adaptive license plate detection system[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2109-2121.
- [20] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of Neural Information Processing System. Cambridge: MIT Press, 2015: 91-99.
- [21] 李祥鹏, 闵卫东, 韩清, 等. 基于深度学习的车牌检测和识别方法[J]. 计算机辅助设计与图形学学报, 2019, 31(6): 979-987.
- LI X P, MIN W D, HAN Q, et al. License plate location and recognition based on deep learning[J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31(6): 979-987.
- [22] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 42(2): 386-397.
- [23] 赵伟, 鞠美玉, 李子双. 复杂光线下的车牌图像增强算法[J]. 计算机辅助设计与图形学学报, 2016, 28(6): 976-985.
- ZHAO W, JU M Y, LI Z S. A license plate image enhancement algorithm under complicated illumination[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(6): 976-985.
- [24] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [25] MISRA D. Mish: a self regularized non-monotonic neural activation function[J]. arXiv:1908.08681, 2019.
- [26] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision, 2017: 2117-2125.
- [27] HSU G S, CHEN J C, CHUNG Y Z. Application-oriented license plate recognition[J]. IEEE Transactions on Vehicular Technology, 2013, 62(2): 552-561.
- [28] XU Z B, YANG W, MENG A J, et al. Towards end-to-end license plate detection and recognition: a large dataset and baseline[C]//Proceedings of European Conference on Computer Vision. Heidelberg: Springer, 2018: 261-277.
- [29] LI H, WANG P. Toward end-to-end car license plate detection and recognition with deep neural networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(3): 1126-1136.
- [30] LI H, SHEN C H. Reading car license plates using deep convolutional neural networks and LSTMs[J]. arXiv:1601.05610, 2016.
- [31] SELMI Z, ALIMI A M. Deep learning system for automatic license plate detection and recognition[C]//Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition, 2017: 1132-1138.
- [32] SELMI Z, HALIMA M B, PAL U, et al. DELP-DAR system for license plate detection and recognition[J]. Pattern Recognition Letters, 2020, 129: 213-223.
- [33] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv:1804.02767, 2018.