

文章编号: 2095-2163(2021)06-0031-05

中图分类号: TP391.4

文献标志码: A

# 停车管理中车牌识别技术的研究

姚 砺, 马 睿, 周同辉, 万 燕

(东华大学 计算机科学与技术, 上海 201600)

**摘要:** 本文研究了基于视频的路边停车管理中车牌识别问题, 提出了动态车牌识别算法。在视频场景下, 基于打分机制选取最佳结果, 充分利用视频的特性, 提高车牌识别率, 有效地解决了车牌遮挡导致无法识别的问题。对于单帧车牌的识别, 使用场景文本识别算法 EAST 定位车牌, 再通过投影变换进行矫正, 以提高在倾斜状态下的识别率。实验证明该方法对倾斜扭曲车牌的识别率有一定提升。

**关键词:** 路边停车; 多目标跟踪; 车牌识别; DeepSORT

## Research on key methods of roadside parking based on deep learning

YAO Li, MA Rui, ZHOU Tonghui, WAN Yan

(Computer Science and Technology Department, Donghua University, Shanghai 201600, China)

**[Abstract]** This paper studies the problem of license plate recognition based on video in roadside parking management and proposes a dynamic license plate recognition algorithm. In the video scene, the best result is selected based on the scoring mechanism, and the license plate recognition rate is improved by making full use of the characteristics of video, which effectively solves the problem of license plate occlusion leading to unrecognition. For the recognition of single frame license plate, the scene text recognition algorithm EAST is used to locate the license plate, and then the projection transformation is used to correct the license plate, so as to improve the recognition rate in the tilt state. Experimental results show that this method can improve the recognition rate of skewed license plate to a certain extent.

**[Key words]** roadside parking; multi-target tracking; license plate recognition; DeepSORT

## 0 引言

近年来, 中国的汽车保有量呈现出飞速增长。仅 2020 年, 全国新注册登记机动车约 3 328 万辆, 全国机动车保有量达 3.72 亿辆。随着机动车数量的增长, 市民对停车位的需求日益迫切, 原有的一些停车场已难以满足不断增长的停车需求, 重要的交通枢纽、医院、商场等开始出现停车难的问题。

为解决停车难的问题, 交管部门在路边划分了一些临时停车位缓解停车位紧张的问题, 但是这也给相关部门带来了管理上的压力。如停车收费问题、违规停车的监测以及处罚问题等, 管理不当会造成道路拥堵, 增加发生交通事故的可能性。国务院通过的《“互联网+”行动的指导意见》提出将互联网与各行业深度融合, 因此路边停车管理手段与“互联网+”思维的结合势必成为符合国情民意发展的新领域。

目前路边停车管理系统多需要进行道路改造, 引进地磁检测器、传感器等硬件设备来检测车辆的停放<sup>[1-2]</sup>, 不仅增加了实施成本, 还可能会影响其它车辆的正常行驶。针对于此, 本文提出了基于视频的路边停车管理方案。该方案利用路边监控系统进行停车管理, 即方便实施又节省成本, 可以在路边复杂的环境下完成车辆的定位, 车牌识别以及停车行为识别。

## 1 总体设计

本文方法是在路边固定位置放置摄像头用于监控和管理路对面的停车位。在道路对面放置摄像头进行监控的好处是容易判定停车, 但由于路边停车多为侧向停车, 车辆进入前后车牌可能会倾斜角度较大甚至被自身遮挡, 这也对车牌识别提出了更大的挑战。为此, 本文使用动态车牌识别算法, 可以很好的应用于该场景。

**作者简介:** 姚 砺(1967-), 男, 博士, 副教授, 主要研究方向: 图像处理和深度学习; 马 睿(1997-), 男, 硕士研究生, 主要研究方向: 计算机视觉和深度学习; 周同辉(1994-), 男, 硕士研究生, 主要研究方向: 计算机视觉和深度学习; 万 燕(1970-), 女, 博士, 教授, 主要研究方向: 图像处理和深度学习。

**通讯作者:** 姚 砺 Email: Yaoli@dhu.edu.cn

**收稿日期:** 2021-04-14

路边停车管理的关键是对停车车辆的车牌进行识别,基于视频的路边停车场景下的车牌识别相比传统的车牌识别主要存在两点不同:一是车辆在停靠时,需不停地移动变换角度,所以大部分截取到的车牌不是正对摄像头;二是由于车位安排较密,车辆停入后可能会遮挡住车牌,导致无法识别。基于这两点,本文提出了动态车牌识别算法:基于视频,在车辆进入镜头内即逐帧开始进行识别,最后通过投票算法选举出最佳结果,可以防止车辆自身遮挡或者相互遮挡导致的无法识别。针对路边停车场景车牌倾斜扭曲的情况,本文算法在定位到车牌后进行倾斜矫正,极大地提高了倾斜状态下车牌的识别率。

动态车牌识别算法需要对同一车辆进行跟踪,持续对其车牌进行识别。对于车辆跟踪,主要难点是路边车辆密集并且会有干扰,需要在短暂的遮挡后能进行重识别。本文提出了基于 DeepSORT<sup>[3]</sup> 的车辆跟踪算法,可在短暂遮挡后重识别。算法在跟踪时,对车辆的停车行为进行判断,并识别其车牌。

路边停车管理总体流程为:车辆跟踪算法和动态车牌识别算法并行执行。跟踪算法在跟踪到车辆后,将图片存储,并通过消息中间件通知车牌识别算法识别对应车辆的车牌,车牌识别算法识别后将识别结果存储到数据库中。如果跟踪过程中判断到车辆停入车位,为防止识别的延迟等情况,先延迟等待,跟踪算法再去数据库中读取相应的车牌识别结果,通过投票算法选取最优结果,记录车辆停车信息。

## 2 动态车牌识别算法

传统车牌识别分为图像预处理、车牌定位、字符分割和字符识别几个步骤。对车牌进行灰度化处理 and 去噪后,进行边缘检测得到定位车牌,再进行二值化、字符分割,最后通过模版匹配得到结果。但这种方法准确率较低,目前多使用基于卷积神经网络(CNN)的车牌识别算法,主要分为两步:定位车牌和识别车牌。

上述算法是基于图片进行车牌识别,对视频场景下的车牌识别,上述算法不能很好的利用视频的特性,由于视频角度问题,单帧图片的识别结果很可能产生误差。针对此问题,本文提出了动态车牌识别算法,可以充分利用视频多帧之间的关联,提高视频状态下的车牌识别率。总体流程为:首先对同一车辆每一帧进行识别,并对图片质量进行打分。图片质量主要包括车牌大小、车牌倾斜程度以及图

片清晰度。以打分为权重,通过投票算法选取最优识别结果。

### 2.1 车牌识别

一般基于深度学习的车牌识别算法主要使用目标检测算法来定位车。如 R-CNN 系列<sup>[4-7]</sup>、YOLO 系列<sup>[8-11]</sup>,结果多为矩形框,以四维向量 $(x, y, w, h)$ 表示。但在路边停车场景下,视频拍摄到的车牌多数会有不同程度的倾斜,如图 1 所示(来自 CCPD 数据集<sup>[12]</sup>)。由目标检测算法定位的车牌位置不够精确,识别效果会因此受到影响。



图 1 目标检测算法定位倾斜角度较大的车牌

Fig. 1 The target detection algorithm locates the license plate with large tilt Angle

针对上述问题,本文使用了场景文本检测算法 EAST<sup>[13]</sup>进行车牌定位。本文基于 EAST 完成了车牌定位,精确定位车牌边缘的 4 个点,便于后续的矫正和识别。

EAST 主要分为两个阶段:首先基于 FCN(全卷积网络)模型生成文本框的预测,再用 NMS(非极大值抑制)产生最终的结果。输出可以有两种形式,一是 RBOX(旋转矩形),包含 4 个通道的水平边界框(AABB)和 1 个通道的旋转角度 $\theta$ ;另一种是 QUAD(任意四边形),包含 8 个通道,表示从矩形的 4 个顶点到像素位置的坐标偏移。由于拍摄角度问题,路边停车场景下拍摄的车牌多不仅有旋转倾斜,还存在一定程度的变形,定位到的结果可能不是矩形,而是经过扭曲后的普通四边形,仅进行旋转无法将其矫正为正对的车牌,故本文选择了 QUAD 输出。图 2 展示了 RBOX 和 QUAD 输出定位的车牌,其中红线是 RBOX 输出的结果,绿线是 QUAD 输出的结果。通过 EAST 定位到车牌的 4 个角点后,通过投影变换对车牌进行倾斜矫正,使车牌正对镜头。

设:给定一张图像,对其进行投影变换步骤如下:首先将二维平面中 $(x, y)$ 处的点,通过变换矩阵投影到一个三维空间 $(X, Y, Z)$ ,如式(1):

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} & \hat{a}_{13} \\ \hat{a}_{21} & \hat{a}_{22} & \hat{a}_{23} \\ \hat{a}_{31} & \hat{a}_{32} & \hat{a}_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \quad (1)$$



图 2 RBOX 输出和 QUAD 输出对比

Fig. 2 RBOX output and QUAD output contrast

再由三维空间  $(X, Y, Z)$  变回二维平面  $(x', y')$ , 如式(2) 和式(3):

$$x' = \frac{X}{Z}, \quad (2)$$

$$y' = \frac{Y}{Z}. \quad (3)$$

即通过上述变化, 原图片  $(x, y)$  处的点将变化为  $(x', y')$  处的点。

通过投影变换, 可以将任意四边形的车牌矫正为矩形, 方便后续的认识。图 3 展示了倾斜车牌经过 EAST 算法定位及投影变换矫正的效果。通过精确定位及倾斜矫正, 可以显著提高识别准确率。

经过矫正后的车牌, 可使用 CRNN<sup>[14]</sup> 网络来进行识别。



(a) 精确车牌定位效果  
(a) Accurate license plate positioning effect

(b) 矫正后  
(b) After correction

图 3 倾斜车牌精确定位及矫正后的效果

Fig. 3 Accurate positioning and correction effect of slanting plate

CRNN 是一种卷积神经网络结构, 用于解决基于图像的序列识别问题。将文字识别看作对序列的预测, 采用了对序列预测的 RNN 网络。通过 CNN 将图片的特征提取出来后, 采用 RNN 对序列进行预测, 最后通过一个 CTC 的翻译层得到最终结

果。使用 CRNN 网络不用对字符进行分割, 减少了分割带来的误差。

## 2.2 投票算法

由于动态车牌识别算法是基于视频的识别, 对每一帧或者每隔几帧, 则对镜头内所有车辆车牌进行识别, 将造成对同一辆车产生多个识别, 最终导致对多个识别结果的最终选取问题。因此, 本文引入投票算法, 为每一帧识别的车牌图像进行质量打分, 以此打分为权重, 在车辆停入后, 根据识别结果进行投票, 选取得分最高的结果作为最终结果。

在路边停车场景下, 车牌的图像质量主要由 3 部分决定:

(1) 车牌大小: 某一摄像头会监控一片区域, 在车辆移动过程中, 车辆离摄像头近、车牌面积大时识别效果较好;

(2) 车牌变形程度: 在车辆反复移动停车的过程中, 车牌会以不同的角度呈现在摄像头下, 车牌会有不同程度的变形, 变形程度小的车牌会比较容易识别;

(3) 图像清晰度: 由于车辆可能处于运动之中, 会存在运动模糊, 模糊程度小的图像, 识别准确率较高。

基于上述分析, 本文提出车牌图像质量公式如下:

$$Q = S + \frac{\lambda}{D} + \mu C. \quad (4)$$

其中,  $S$  为车牌面积;  $D$  为变形程度;  $C$  为图片清晰度;  $\lambda$ 、 $\mu$  为权重。

变形程度可以使用车牌的最小外接矩形面积与车牌面积的比值来表示, 计算公式如下:

$$D = \frac{S'}{S}. \quad (5)$$

其中,  $S'$  为车牌最小外接矩形的面积。

图像清晰度可以使用拉普拉斯梯度法来计算, 其公式如下:

$$C = \sum_y \sum_x |G(x, y)|, \quad G(x, y) > T. \quad (6)$$

其中,  $T$  为阈值,  $G(x, y)$  是图像与拉普拉斯算子的卷积, 拉普拉斯算子如下:

$$L = \begin{bmatrix} \hat{e}^0 & 1 & 0 \\ \hat{e}^1 & -4 & 1 \\ \hat{e}^0 & 1 & 0 \end{bmatrix}. \quad (7)$$

## 3 基于 DeepSORT 的车辆跟踪算法

### 3.1 技术路线

动态车牌识别算法需要对车辆进行持续跟踪,



以连续地对其车牌进行识别。在路边停车环境下,路况复杂,有很多行人和过往车辆干扰,且会存在车辆之间互相遮挡,短暂出镜的情况,可能会导致跟踪丢失,影响车牌识别结果。经研究,本文基于DeepSORT完成了车辆跟踪算法。DeepSORT是一个多目标跟踪算法,与传统目标跟踪算法相比,其融合了对跟踪对象的外观特征进行跟踪,可以对短暂离镜的目标进行重识别,提高了跟踪的稳定性。由于DeepSORT算法只能对车辆进行跟踪,而无法对车辆行为进行识别,因此本文算法在DeepSORT基础上进行改进,加入了停车判定模块。

算法以DeepSORT为基础,以YOLO-v3为检测器,先通过检测器检测目标,生成检测(detection),通过卡尔曼滤波生成预测,再使用匈牙利算法,将检测的检测框和卡尔曼滤波的预测进行匹配。对于匹配成功的detection预测生成新的跟踪(track)或更新原有track的状态;每次更新状态,会判断其停车状态,如果判定为驶入,则读取车牌识别情况,开始停车计时;若判定为驶出,则停止停车计时。

对于检测框和预测框的匹配,使用 $d^{(1)}(i,j)$ 和 $d^{(2)}(i,j)$ 两个特征融合作为匹配度量。 $d^{(1)}(i,j)$ 为第 $i$ 个track的预测位置到第 $j$ 个detection的检测位置的马氏距离,公式如下

$$d^{(1)}(i,j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i). \quad (8)$$

其中, $d_j$ 表示第 $j$ 个detection的检测位置; $y_i$ 表示第 $i$ 个track的预测位置; $S_i$ 为两者的协方差矩阵。

对于每一个detection,都会通过深度学习提取外观特征,得到一个特征向量,此外还会创建一个集合 $R_i$ 来保存第 $i$ 个track近100个匹配成功的detection的特征向量。通过加入外观特征的匹配,可以完成对短暂离镜目标的重识别。 $d^{(2)}(i,j)$ 为第 $i$ 个track的最近100个成功匹配的特征向量集与当前帧第 $j$ 个detection的特征向量间的最小余弦距离,公式如下:

$$d^{(2)}(i,j) = \min \{ 1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i \}. \quad (9)$$

其中, $r_j$ 表示第 $j$ 个detection的外观特征向量, $r_k^{(i)}$ 为 $R_i$ 中第 $k$ 个特征向量。

两个特征通过线性加权结合,作为最终的匹配度量:

$$c(i,j) = \lambda d^{(1)}(i,j) + (1 - \lambda) d^{(2)}(i,j). \quad (10)$$

其中, $\lambda$ 为线性加权的权值。本文使用了DeepSORT论文中给出的经验值0,即仅使用外观特征进行匹配。

### 3.2 停车判定

为判定停车状态,本文使用车位占用比occupation ratio来量化车位被占用的概率。当occupation ratio大于设定阈值时即认为车位被占用。计算公式如下:

$$occupation\ ratio = \frac{C \cap P}{P}. \quad (11)$$

其中, $C$ 表示车辆的面积; $P$ 表示停车位的面积; $C \cap P$ 表示两者相交的面积。

在实际使用时,由于摄像头拍摄角度问题,可能会有车辆暂时路过遮挡停车位,除此之外,算法可能存在某一帧的误判,如果仅依据occupation ratio判断,可能会造成误判,本文加入不匹配帧数来解决此问题。在track和detection匹配之后,将检测目标记录与当前判断的状态是否一致。当状态不一致时,记录不匹配帧数,当此帧数大于阈值时,再判定车辆停入或驶出车位。

## 4 实验结果及分析

实验环境:Ubuntu 18.04 操作系统,Intel(R) Xeon(R) Bronze 3204 CPU @ 1.90GHz 处理器, GeForce RTX 2080ti 11GB 显卡。

本文使用CCPD数据集进行训练和评估。CCPD数据集是一个大型的中文车牌数据集,总共有30万张图片,将其分为普通图片和较难识别图片。较难识别的又分为模糊、光线亮或暗、倾斜等情况,标注了车牌的4个角点和车牌号码等信息。本文将这些图片按照训练:验证:测试=8:1:1划分进行实验。

使用CCPD数据集和车牌位置标注信息训练EAST模型。利用CCPD的标注截取其中车牌,形成车牌数据集。由于车牌数据集分布不均匀,大部分车牌是皖,即安徽的车牌,其他省份车牌较少,故本文对车牌数据集进行了增强,合成了一些车牌图像,并加入了噪声以模拟真实环境下的效果,如图4所示。通过扩充后的数据集训练CRNN模型,对两个模型进行串联评估。对是否进行矫正分别做了测试,其中不加入矫正的识别中,CRNN的训练集也未经矫正,测试结果见表1。

表1 是否加入矫正对比结果

Tab. 1 Comparison results of whether to add corrective

算法	正向车牌	倾斜车牌
不加入矫正	92.2%	70.3%
加入矫正	99.4%	96.2%



图4 合成的车牌图像

Fig. 4 The generated license plate image

由表1中结果可以发现,无论是正向车牌还是倾斜车牌,加入矫正后的测试准确率都有较大的提升。

将其与CCPD数据集论文中的RPNet识别率对比结果见表2。

表2 车牌识别测试结果

Tab. 2 License plate recognition test results

算法	正向车牌	倾斜车牌
RPNet	98.5%	92.5%
本文算法	99.4%	96.2%

分析表2中结果可以发现,对正向车牌的识别,两者识别率差距较小,而对倾斜车牌的识别,RPNet的识别率只有92.5%,而本文算法比RPNet的效果有显著提高。

## 5 结束语

本文对路边停车管理中的车牌识别进行了研究,提出了动态车牌识别算法,基于视频完成车牌识别,利用视频多帧的关联,在车辆入镜后即开始进行识别,对每一帧识别到的车牌图片质量打分,以该分值为比重,通过投票机制选出最佳识别,完成了对视频中车辆的车牌识别。通过场景文本检测算法EAST定位车牌的4个角点,经投影变换矫正后通过CRNN进行车牌识别。在CCPD数据集上的实验表明,该算法对大角度倾斜车牌的识别率较高。

动态车牌识别算法需要对车辆进行持续跟踪,并识别其停车行为。本文采用基于DeepSORT的车辆跟踪算法对车辆进行跟踪。以YOLO-v3为检测

器检测车辆,对DeepSORT算法改进,加入停车判定来对车辆行为进行判定,在跟踪的同时对其车牌进行识别,该算法可以较好地与动态车牌识别算法配合,防止车辆停入车位后车牌被遮挡导致无法识别。

## 参考文献

- [1] 陈红英. 基于LoRa和Nb-IoT的智慧路边无人值守停车云系统[J]. 微型电脑应用, 2020, 36(1): 48-51.
- [2] 熊礼平. 基于物联网技术的路边停车系统的设计[J]. 电子世界, 2018(1): 186, 188.
- [3] WOJKE N, BEWLEY A, PAULUS D. Simple Online and Realtime Tracking with a Deep Association Metric[J]. IEEE, 2017: 3645-3649.
- [4] GIRSHICK R, DONAHUE T, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [5] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904-1916.
- [6] GIRSHICK R. Fast r-cnn[C]//2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2016.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[J]. IEEE, 2016.
- [9] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2017: 6517-6525.
- [10] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[J]. arXiv e-prints, 2018.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. 2020.
- [12] XU Z, YANG W, MENG A, et al. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline[C]// European Conference on Computer Vision. Springer, Cham, 2018.
- [13] ZHOU X, YAO C, WEN H, et al. EAST: An Efficient and Accurate Scene Text Detector[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.
- [14] SHI B, XIANG B, CONG Y. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(11): 2298-2304.