

文章编号:1007-1423(2021)24-0119-05

DOI:10.3969/j.issn.1007-1423.2021.24.022

## 基于深度学习的复杂场景下车牌定位与识别

郝达慧<sup>1,3</sup>,王池社<sup>1,2</sup>,陈敏<sup>1,3</sup>

(1. 安徽理工大学,计算机科学与工程学院,淮南 232001; 2. 金陵科技学院,网络与通信工程学院,南京 211000;  
3. 江苏省人工智能交通创新应用工程研究中心,南京 211000)

**摘要:**针对光照昏暗、天气异常等复杂场景下的车牌识别效果不佳的问题,本文提出了一种基于深度学习的复杂场景下车牌定位与识别方法。方法基于 YOLOv5 的网络结构进行车牌定位,并设计了 LPCRNet 网络对定位车牌进行精确识别。将该方法应用于一个复杂场景下的标准数据集进行训练和测试,通过与其他方法的实验结果对比表明,本文方法在识别效率与准确性上都取得了更好的效果。

**关键词:**车牌识别;复杂场景;深度学习

### 0 引言

随着社会经济的飞速发展,汽车的数量也在大幅度增长。汽车在给人们带来便利的同时,也带来了很多安全隐患,如交通事故频发、车辆违规等问题。车牌定位与识别是计算机视觉与模式识别技术在智能交通领域应用的重要研究课题之一<sup>[1]</sup>,在车辆管理方面有着较为广泛的应用,如出入控制、违章监控、违法车辆跟踪等<sup>[2]</sup>。早期的车牌识别算法大多都是基于机器学习算法进行研究,使用手工选取的特征对车牌进行定位与识别。近年来,随着大数据时代的到来和计算机算力的提升,深度学习在车牌识别方向取得了重大突破,Faster R-CNN<sup>[3]</sup>、YOLO<sup>[4]</sup>等深度学习算法的提出使得车牌的定位与识别得到了新的发展。现有的车牌识别技术主要应用于特定的环境,如收费停车场出入口、高速公路 ETC 通道等。在这种检测区域固定、光照良好的环境下,车牌识别技术的准确率可以达到很高,但是在复杂的场景下识别效果较差。本文针对光照不均、天气异常等场景下的车牌识别,提出了基于无分割的车牌识别算法,基于 YOLOv5 对车牌进行定位并设计了端到端的字符识别网络(LPCRNet)对车牌字符进行识别。

### 1 相关工作

车牌识别算法可以分为车牌检测和字符识别两个部分,车牌检测是为了定位车牌在图片中的具体位置,字符识别用于对检测出来的车牌区域的字符进行识别,最终得出图片中的车牌号。

#### 1.1 车牌检测

车牌检测算法分两种,一种是基于手工选取的特征进行检测,另一种是基于深度学习方法进行检测。Yuan 等<sup>[5]</sup>提出使用线密度滤波方法提取车牌候选区域,最后基于线性 SVM 的级联分类器从候选区域中提出车牌位置。Tian 等<sup>[6]</sup>提出了一种基于边界聚类的检测方法,通过 Canny 算子获取边界图,然后基于密度的聚类方法将边界划分为不同的聚类,最后使用线性支持向量机对边界定位准确的水平候选区域进行分类。Zhang 等<sup>[7]</sup>设计了一种由卷积神经网络和递归神经网络组成的深层网络,专门针对污损、遮挡等特殊车牌的定位问题。Wang 等<sup>[8]</sup>提出了一种精度更高、计算成本更低的多任务卷积神经网络车牌检测与识别系统(MTLPR),在实际复杂的场景中具有较好的鲁棒性。

#### 1.2 字符识别

车牌字符识别方法分为两阶段和一阶段方法,两

阶段方法首先对车牌区域进行字符分割,接着对分割出来的单个字符进行识别。一阶段方法采用端到端的深度学习算法直接对车牌区域进行字符识别。余承波<sup>[9]</sup>使用垂直投影分割算法进行字符分割,提出了一种融合字符的统计特征和结构特征并结合 SVM 的字符识别方法。Kessentini 等<sup>[10]</sup>分别两个 YOLO 模型对车牌进行识别,一个用于对车牌区域进行定位,另一个模型对车牌区域的字符进行识别。Li 等<sup>[11]</sup>在车牌检测的池化层后直接接入基于 RNN 的识别网络,省去了中间字符分割的处理,整个网络采用联合损失。

## 2 基于深度学习的车牌识别算法

车牌检测是车牌字符识别的基础步骤,检测结果直接影响后续字符识别的准确率。本文的检测模块基于 YOLOv5 对图像中的车牌进行定位,由于只需要检测车牌一个类别,选用 YOLOv5 中深度最浅的 YOLOv5s 网络结构。

传统的车牌字符识别大都是基于字符分割的识别方法,识别效果受字符分割的影响很大,当存在车牌图像倾斜角度过大、光照昏暗等干扰时,字符分割效果并不理想。本文的识别模块基于 CRNN 网络<sup>[12]</sup>设计了一个端到端识别车牌字符的 LPCRNet 网络。

### 2.1 检测模块

检测模块使用 CSPDarkNet 作为主干特征提取网

络,从输入车辆图像中获取丰富的特征信息,有效缓解了大型卷积神经网络的梯度消失问题。Neck 部分选用 PANET(FPN+PAN),用于生成特征金字塔,会增强模型对于不同尺度的车牌的检测。检测模块的网络结构如图 1 所示。

Bounding box 损失函数采用的是 GIoU\_Loss, GIoU 的计算公式如下:

$$GIoU = IoU - \frac{A^c - U}{A^c} \quad (1)$$

其中 IoU 可以用下式进行表示:

$$IoU = \frac{I}{U} = \frac{I}{A^p + A^g - I} \quad (2)$$

其中  $I$  表示预测框与目标框的相交面积,  $U$  为预测框的面积  $A^p$  与目标框的面积  $A^g$  之和减去  $I$ ,  $A^c$  为预测框与目标框的最小外接矩形的面积。

### 2.2 识别模块(LPCRNet)

CRNN 网络由卷积层、循环层、转录层 3 部分组成,本文基于 CRNN 网络设计了 LPCRNet。由于车牌图像的特征较为简单, LPCRNet 的卷积层采用网络较浅的 ResNet18 中的卷积层来提取车牌图像的特征;循环层采用 BiLSTM (即双向 LSTM), BiLSTM 对车牌的上下文信息有较好的表示能力,合并正向 LSTM 和反向 LSTM 的序列信息完成对序列标签的预测;最后的转录层使用 CTC 方法对 BiLSTM 输出的预测序列进行对齐映射,输出预测的车牌号码。识别模块结构如图 2 所示。

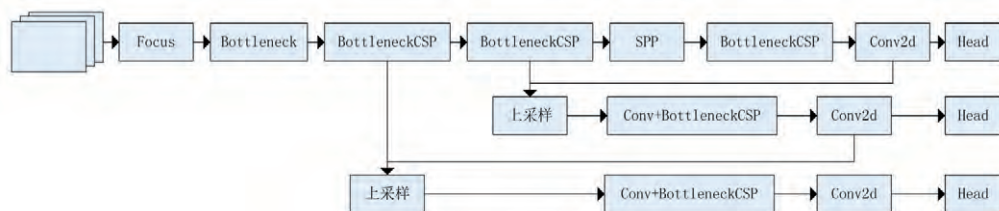


图1 检测模块网络结构

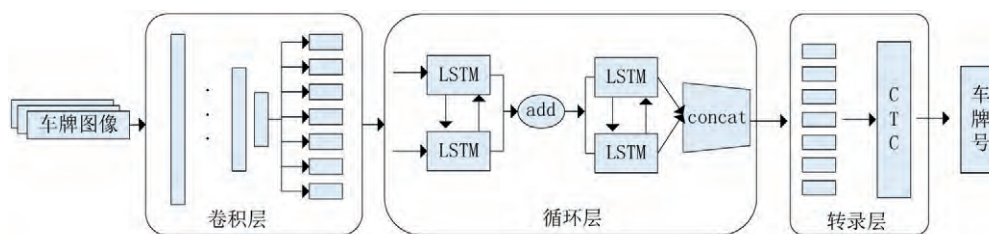


图2 识别模块网络结构

循环层通过叠加两个 BiLSTM 网络将输入的序列特征进行解码,第二层 LSTM 的处理过程可以用以下公式表示:

$$o_i^{(2)} = \text{LSTM}^{(2)}(\text{LSTM}^{(1)}(x_i)) \quad (3)$$

循环层的输入序列  $x_i$ , 经过双层 LSTM 网络后输出  $o_i^{(2)}$ , 然后经过 Softmax 层得到该序列的预测标签  $\hat{y}_i = i, i \in 1, 2, \dots, T$  的概率:

$$P(\hat{y}_i = i | x_i) = \frac{\exp(w_i(o_i^{(2)}))}{\sum_{l=1}^T \exp(w_l(o_i^{(2)}))} \quad (4)$$

其中,  $w_i$  和  $w_l$  表示预测标签为  $i$  和  $l$  时 Softmax 层的参数变换。

随后, CTC 对循环层的输出  $\hat{y}_i$  进行对齐映射, 得到序列路径  $\pi$  的概率为:

$$P(\pi | \hat{y}) = \prod_i P(\hat{y}_{\pi_i}^i, \theta) \quad (5)$$

其中  $\hat{y}_{\pi_i}^i$  表示在  $i$  位置处  $\hat{y}_i$  的字符与路径  $\pi$  在该位置上的字符相同的概率。  $P(l | \hat{y})$  为所有可映射为  $l$  的路径的概率之和。损失函数的最终目标是最大化  $P(l | \hat{y})$ , 也就是最大化  $P(\hat{y}_{\pi_i}^i, \theta)$ :

$$L_{CTC} = \max(P(l | \hat{y})) = - \sum_i \sum_l \ln P(\hat{y}_{\pi_i}^i, \theta) \quad (6)$$

识别模块的训练过程中, 目标是通过循环层的输出序列找到概率最大的近似最优路径。

### 3 实验结果与分析

#### 3.1 数据集与评价指标

(1) 数据集。实验数据集来源于中科大的研究人员构建的 CCPD 数据集<sup>[13]</sup>, 包含 25 万多张图片, 其中包括 9 种不同的场景: 模糊、异常天气、倾斜等。具体内容如表 1 所示。

表 1 CCPD 数据集内容

场景	数量/k	描述
Base	200	正常车牌 (一张图像只有一个车牌)
Fn	20	车牌过大或过小
Db	20	光线过亮或过暗
Rotate	10	水平倾斜 20° 至 25°, 垂直倾斜 -10° 至 10°
Tilt	10	水平倾斜 15° 至 45°, 垂直倾斜 15° 至 45°
Weather	10	异常天气 (雨天、雪天、雾天)
Blur	5	相机抖动造成的图像模糊
Challenge	10	其他场景
Np	5	无车牌车辆

实验随机选取除无车牌外的其他 8 种场景下的 10000 张图作为车牌检测的数据集, 根据图片中车牌的坐标截取出的 10000 张车牌图像作为识别模块的数据集, 其中 8500 张作为训练集, 1500 张作为测试集。

(2) 评价指标。检测模块使用目标检测中较常使用的准确率 (Precision) 和召回率 (Recall) 进行评估性能。车牌字符识别有两种常用评价方法, 一种是整体评价, 另一种是字符级评价。整体评价是指车牌所有字符都识别正确的数量在测试集总数中所占的比例, 公式如下:

$$RA = \frac{C_p}{N_p} \quad (7)$$

其中,  $C_p$  表示字符全部识别正确的车牌数量,  $N_p$  表示测试集的样本总数。还有较为常见的 5 字识别率和 6 字识别率, 分别表示车牌中正确识别 5 个字符和 6 个字符的数量与测试集总数的比例。

字符级评价指的是测试集中识别正确的字符数占所有字符的比例, 公式如下:

$$CRA = \frac{C_c}{N_c} \quad (8)$$

其中,  $C_c$  表示所有识别正确的字符数,  $N_c$  是指测试集中所有字符数。

本文基于以上两种评价方法对识别模块进行评估。

#### 3.2 实验结果

各场景下车牌检测效果如图 3 所示, 由检测结果可知, 检测模块对于多种复杂场景下车牌定位具有较好的效果。



图 3 车牌检测实例

本文检测模块与其他目标检测算法在 CCPD 数据集上的实验对比结果如表 2 所示, 由对比结果可知, 本文方法的准确率和召回率明显高于 SSD 算法, 且速度都较快于其他算法。



表2 不同车牌检测算法的检测结果对比

算法	precision	recall	处理时间/ms
SSD <sup>[14]</sup>	96.45%	74.42%	91
Faster R-CNN	98.19%	90.13%	333
本文	98.6%	91.7%	11

识别模块与其他字符识别算法的对比结果如表3所示,结果表明 LPCRNet 在识别率和速度上都优于对比算法。

表3 不同车牌识别算法的识别结果对比

算法	5 字识别率	6 字识别率	RA	CRA	处理时间/ms
EasyPR	89.3%	78.6%	68.5%	80.5%	324
OpenALPR	96.4%	84.1%	79.1%	83.6%	68
LPRNet	96.6%	83.6%	80.3%	88.6%	81
本文	98.6%	91.7%	89.9%	98.4%	6

4 结语

由于车牌识别在光照过亮或过暗、车牌倾斜、异常天气等场景下的识别效果不尽如人意,本文提出了基于深度学习的车牌识别算法,检测模块基于 YOLOv5 网络结构定位出图像中的车牌位置,识别模块设计无分割的字符识别算法(LPCRNet),首先将图片输入卷积层对车牌图像进行特征提取,接着将车牌的序列特征送入双层的 BiLSTM 网络对序列标签进行预测,最后使用 CTC 对预测标签转录得到最终的预测车牌号。本文提出的算法在 CCPD 数据集中各复杂场景下的车牌检测准确率达到了 98.6%,字符识别的准确率达到 88.9%,实验结果表明,本文提出的算法对于复杂场景下车牌的检测和识别都有较好的鲁棒性。

参考文献:

[1]GOU C, WANG K F, YAO Y J, et al. Vehicle license plate recognition based on extremal regions and restricted boltzmann machines [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(4): 1096–1107.

[2]YUAN Y, ZOU W, YONG Z, et al. A robust and efficient approach to license plate detection[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2017, 26(3).

[3]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, 2015, 39(6): 1137–1149.

[4]REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of CVPR, 2015: 779–788.

[5]YUAN Y, ZOU W, ZHAO Y, et al. A robust and efficient approach to license plate detection[J]. IEEE Transactions on Image Processing, 2017, 26(3): 1102–1114.

[6]TIAN J, WANG G, LIU J, et al. License plate detection in an open environment by density-based boundary clustering[J]. Journal of Electronic Imaging, 2017, 26(3): 33017–33017.

[7]ZHANG J, LI Y, LI T, et al. License plate localization in unconstrained scenes using a two-stage CNN-RNN[J]. IEEE Sensors Journal, 2019: 1–1.

[8]WANG W, YANG J, CHEN M, et al. A light CNN for end-to-end car license plates detection and recognition[J]. IEEE Access, 2019, 7: 173875–173883.

[9]余承波. 基于支持向量机(SVM)的车牌识别[D]. 安徽理工大学, 2018.

[10]KESSENTINI Y, BESBES M D, AMMAR S, et al. A two-stage deep neural network for multi-norm license plate detection and recognition[J]. Expert Systems with Applications, 2019, 136.

[11]LI H, WANG P, SHEN C. Toward end-to-end car license plate detection and recognition with deep neural networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(3): 1126–1136.

[12]SHI B, BAI X, YAO C. 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 and Machine Intelligence, 2016, 39(11): 2298–230.

**作者简介:**

郝达慧(1997—),女,江苏盐城人,硕士研究生,研究方向为人工智能、计算机视觉

王池社(1974—),男,安徽安庆人,博士,教授,研究方向为大数据技术、人工智能、机器学习

陈敏(1995—),男,江苏淮安人,硕士研究生,研究方向为人工智能、计算机视觉

收稿日期:2021-03-23

修稿日期:2021-06-02

## License Plate Location and Recognition in Complex Scenes Based on Deep Learning

Hao Dahui<sup>1,3</sup>, Wang Chishe<sup>1,2</sup>, Chen Min<sup>1,3</sup>

(1. College of Computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001;

2. College of Network and Communication Engineering, Jinling Institute of Technology, Nanjing 211000;

3. Jiangsu Research Center for Innovative Application Engineering of Artificial Intelligent Transportation, Nanjing 211000)

**Abstract:** Aiming at the problem of poor license plate recognition effect in complex scenes such as dim light and abnormal weather, this paper proposes a license plate location and recognition method based on deep learning in complex scenes. License plate location was carried out based on the network structure of YOLOv5, and LPCRNNet was designed to recognize the license plate accurately. This method is applied to a standard data set in a complex scene for training and testing. The experimental results show that this method achieves better results in recognition efficiency and accuracy compared with other methods.

**Keywords:** license plate recognition; complex scene; deep learning

(上接第 118 页)

## Prediction of Lymph Node Metastasis Status Based on Multi-Source Heterogeneous Data

Huang Kemin

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006)

**Abstract:** A deep learning classification model based on multi-source heterogeneous data is proposed to evaluate the status of axillary lymph node metastasis in patients with early-stage breast cancer. The model extracts the features of traditional radiomics features, B-mode ultrasound images and shear wave elastic ultrasound image data, constructs a feature space and performs classification prediction. The classification accuracy rate of this model for lymph node metastasis status reached 86.2%. Experimental results show that multi-source heterogeneous data fusion can effectively improve the diagnostic performance of axillary lymph node metastasis status.

**Keywords:** deep learning; multi-source heterogeneous data; radiomics; axillary lymph node metastasis