

文章编号: 2095-2163(2021)11-0125-04

中图分类号: TP391.1

文献标志码: A

# 基于卷积神经网络的免字符分割车牌识别算法

陈墨林

(吉林大学 计算机科学与技术学院, 长春 130015)

**摘要:** 随着智能停车场和自动收费系统的普及,智能交通管理系统在车牌识别方面对速度和精度提出了更高的要求。为了提高车牌识别的泛化能力,本文提出了一种基于改进的 CRNN+CTC 免字符分割车牌字符识别算法,在现有的卷积神经网络的 CRNN+CTC 网络模型基础上分别对 CNN 和 RNN 进行替换后引入 CTC 函数,并有针对性地增加训练数据集,在训练过程中使用了批量归一化算法来加快学习速度。实验结果表明,该算法的鲁棒性与通用性均有一定提升,在复杂环境下车牌识别的准确性更高,达到了改进效果。

**关键词:** 车牌识别; CNN; RNN; 复杂环境

## License plate recognition algorithm without character segmentation based on convolutional neural network

CNEN Molin

(School of computer science and technology, Jilin University, Changchun 130015, China)

**[Abstract]** With the construction of intelligent parking lot and urban traffic and the popularization of automatic toll collection system, intelligent traffic management system puts forward higher requirements for speed and accuracy in license plate recognition. In order to improve the generalization ability of license plate recognition, this paper proposes a character segmentation free license plate character recognition algorithm based on improved CRNN+CTC. Based on the existing CRNN+CTC network model, CNN and RNN are replaced respectively, CTC function is introduced, and the training data set is added. Finally, batch normalization algorithm is used in the training process to speed up the learning speed. The experimental results show that the algorithm has higher accuracy in license plate recognition under complex environment, and its robustness and generality are improved.

**[Key words]** license plate recognition; convolutional neural network; recurrent neural network; complex environment

## 0 引言

车牌识别是实现交通管理系统智能化的关键。车牌识别技术主要分为两部分:一是车牌定位技术,当下车牌定位技术在国内比较成熟,尤其是以 Slimani 等人的卷积神经网络(CNN)的车牌识别为代表,其定位精度高达 99.4%,在自然环境中几乎达到 100%的准确度;二是车牌字符识别技术,目前国内大多数车牌识别方法是传统的基于字符分割的方法,这些方法的准确率只有约 86.4%,复杂环境时仍存在很大的局限性与性能低的问题<sup>[1]</sup>。因此对车牌字符识别技术的改进具有很高的价值。

车牌识别有两种方法:第一种是基于字符分割的车牌识别方法,该方法通常基于向量机、基于投影或基于模板匹配等,执行分割步骤以分离字符,执行逐个字符的识别,但基于字符分割的方法容易受到工作环境的影响,在实际应用中的结果不令人满意;第二种是基于字符无分割的方法,属于深度学习领

域,该方法识别所有字符,且不使用特定的结构分离字符,可确保图像中的总信息保持不变,不受分割结果的影响,在模糊车牌识别方面有一定的优势。无分割识别算法主要分为基于 CNN、基于 RNN 和基于两者的 CNN+RNN 网络(CRNN)3 种类型,其中 CNN 和 RNN 的组合使车牌的识别达到了一个新的水平,特别是 CRNN+CTC,在理想条件下其准确率可达 98.5%,但仍存在网络在训练期间易出现过拟合的问题,网络的结构复杂性和计算量方面有待进一步降低。因此,本文提出一种改进 CRNN+CTC 的方法。

## 1 本文算法

CRNN+CTC 的免分割车牌识别算法主要包括 CNN、RNN 和 CTC 3 个单元。首先,将车牌图像输入 CNN 以提取特征图像,并用带残差连接的深度可分离卷积取代标准卷积;其次,将得到的车牌特征图像转换为特征序列后,使用双向循环神经网络对其

作者简介:陈墨林(2002-),男,本科生,主要研究方向:人工智能。

收稿日期:2021-09-02

进行学习;最后,CTC 将获取的标签分布经过去重和整合等操作转化为最终识别结果。

### 1.1 CNN 单元

传统的 CNN 使用 VGG-16 网络模型提取车牌的多维特征,但深层网络结构的使用导致模型较为复杂庞大<sup>[2]</sup>。在这种情况下,符号识别算法嵌入到固定或移动设备中时无疑会造成难度的增加。本文采用深度可分离卷积神经网络取代原始的 CRNN 中的 8 层卷积网络,其中  $K$  为卷积核大小(kernel size), $S$  为步长(stride), $P$  为填充大小(padding size), $ReLU$  为线性整流函数,如图 1 所示。

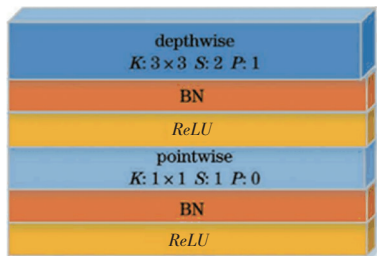


图 1 深度可分离卷积

Fig. 1 Depth wise separable convolution block

这不仅压缩了模型的大小,减少了卷积层的计算量,还提高了模型的计算速度。其具体流程是先向 CNN 输入  $160 \times 48$  pixel 像素的 RGB 图像,经过  $3 \times 3$  卷积层后将其送入深度可分离卷积层,提取图像特征后送入下一层可分离卷积层,并且在逐层之间加入一次极大值池化操作,最后输出图像特征图。

### 1.2 RNN 单元

由于 RNN 单元存在梯度消失的问题,因此其可以存储的范围受到了限制,这也增加了训练的工作量。本文采用长短期存储模型(LSTM)构建 RNN 模块,因其适合于时间的处理和预测具有较长延迟的事件。本文在 RNN 模块使用 Bi-LSTM 以解决 LSTM 和 RNN 只能预测某一个方向的问题,有助于捕获更多的信息。同时 Bi-LSTM 结构比 LSTM 易于提取抽象特征,也更有利于分类。

CRNN 的处理流程:通过 CNN 将 3 个通道大小为  $160 \times 48$  pixel 提取的车牌图像特征图按图 2 的方式从左到右提取特征向量,进而转换成大小为  $1 \times 25 \times 512$  的卷积特征矩阵;向 Bi-LSTM 输入由每一个通道转换成的特征长度  $D = 512$  和时间步长为 25 的时间序列,最后输出由归一化指数函数 Softmax 计算得到的后验概率矩阵。

### 1.3 CTC 损失函数的引入

RNN 的训练常使用交叉熵损失函数。但在国内

车牌识别的实际应用中,由于车牌字符的长度和数量存在差异,其识别过程中不能预测输出文字的长度,因此使用常见的交叉熵函数训练网络的效果不理想。为解决这个问题,本文引入 CTC (connectionist temporal classification) 损失函数训练 RNN。CTC 损失函数引入了空字符,给出了一个固定的输出长度,且可以进行递归运算以快速计算梯度,有效地解决了该问题。

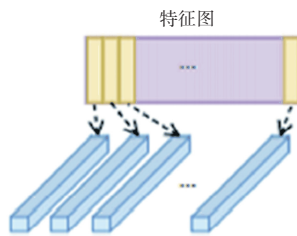


图 2 提取特征向量

Fig. 2 Extract feature vector

对于给定的输入  $x$ , 使用 CTC 中的最佳路径解码算法进行解码,将其转换为对应于最高概率的字符串<sup>[3]</sup>,其最佳路径  $\pi$  的概率定义为式(1):

$$p(\pi | x) = \prod_{t=1}^T y_{\pi_t}^t \quad (1)$$

其中,  $y_{\pi_t}^t$  表示第  $t$  帧为  $\pi_t$  的概率。在此基础上定义多对一映射函数  $B$ , 从而对于一个给定的标签其输出概率可以用式(2)描述成所有可能路径概率之和的形式,即:

$$p(Z | x) = \sum_{\pi_i \in B^{-1}(Z)} p(\pi_i | x) \quad (2)$$

在式(1)和式(2)得到的损失函数的基础上引入  $L2$  范数正则项以减少过拟合现象的出现,表达式为式(3):

$$J(S) = \alpha \sum_S S^2 - \sum_{(x,Z) \in S} \ln \left( \sum_{\pi_i \in B^{-1}(Z)} \prod_{t=1}^T y_{\pi_t}^t \right) \quad (3)$$

其中,  $\alpha$  为  $L2$  范数正则项系数,设为 0.8,即可有效减少过拟合现象的出现。

## 2 训练方式

### 2.1 数据集

为了进一步提高本文算法的检测性能和鲁棒性,本文样本分为 3 个数据集:GAN\_LP、Reldv1、CCPDv1。GAN\_LP 是由对抗生成的仿真数据集,随机从中挑选出 15 000 张作为本实验样本;Reldv1 是从面向应用程序 LP(AOLP)基准 RELD 数据集中随机挑选出 4 000 张模糊车牌图像组成的本实验数据集;CCPDv1 是从中国科学技术大学建立的 CCPD

数据集中挑选出的复杂环境下拍摄的 4 500 张模糊车牌图像组成的本实验数据集。

## 2.2 训练细节

本实验训练的硬件环境:主频 4 GHz,物理内存 16 GB 的 Intel 酷睿 i7-6700K,GPU 为 NVIDIA GTX1080,操作系统选用 ubuntu 14.01。

本文使用 TensorFlow 进行 20 000 次 64 批迭代训练,超参数设定值如下:初始学习率  $Base\_lr$  为 0.001,学习率更新因子  $\gamma$  为 0.1,学习率更新策略为 Multistep,激活函数  $Active\_fun$  为 ReLU,损失函数优化策略  $type$  为 SGD,权重衰减因子  $weight\_decay$  为 0.000 5,动量因子  $momentum$  为 0.8,一批包含的样本数目为 64。

## 3 实验结果分析

### 3.1 实验结果

本文主要从 3 个维度对算法检测性能和鲁棒性进行评估:车牌识别准确度 ( $RA$ )、字符识别准确度 ( $CRA$ ) 和单张车牌平均识别时间  $T$ 。对大量传统方法和主流方法逐一复现,在本文中的 3 个数据集上进行实验,并准确记录相应数据,为进一步改进算

法奠定基础。本文的 CRNN+CTC 改进算法是重复进行算法复现、实验结果分析、算法改进、实验对比这 4 个步骤得到的。本文列出改进后的算法在 3 个数据集下车牌识别准确度 ( $RA$ ) 和字符识别准确度 ( $CRA$ ) 的实验结果。

改进后算法的  $RA$  指标在 GAN-LP 数据集和 CCPDv1 数据集上分别为 96.7% 和 96.9%,在 RELD 数据集上达到了 97.8%; $CRA$  指标在 GAN-LP 数据集和 CCPDv1 数据集上分别为 97.1% 和 97.2%;在 RELD 数据集上达到了 98.4%。通过实验数据不难看出, $RA$  和  $CRA$  这两个指标在 GAN-LP 数据集和 CCPDv1 数据集上的结果几乎是一致的,在 RELD 数据集上的结果要好于在 GAN-LP 数据集和 CCPDv1 数据集上的结果。通过对 3 个数据集的测试,本文提出的改进算法的车牌识别准确率为 96.7%~98.4%,识别性能较好,整体准确率较高。

### 3.2 结果对比与分析

为了比较和分析改进后算法的检测性能,将本文改进的 CRNN+CTC 分别与传统的字符分割识别方法、LPR-Net<sup>[4]</sup> 和原 CRNN+CTC 这 4 种方法在 3 个数据集上的测试结果进行对比,结果见表 1。

表 1 实验对比结果

Tab. 1 Experimental comparison results

Method	GAN_LP			Reldv1			CCPDv1		
	$RA$ %	$CRA$ %	$T$ /ms	$RA$ %	$CRA$ %	$T$ /ms	$RA$ %	$CRA$ %	$T$ /ms
Traditional	84.1	77.3	25	85.5	75.8	32	83.5	74.2	31
LPR-Net	94.3	92.7	63	88.5	86.2	46	87.4	86.7	72
CRNN+CTC	96.9	96.1	46	96.6	97.0	44	97.1	96.4	48
Ours	96.7	97.1	29	97.8	98.4	36	96.9	97.2	32

从表 1 中可以看出,传统的基于字符分割的车牌识别方法和 LPR-Net 算法在含有模糊车牌的 Reldv1、CCPDv1 数据集下的识别效果并不理想, $RA$  和  $CRA$  均小于 90%。而原 CRNN+CTC 算法在 Reldv1、CCPDv1 数据集下识别准确率在 96% 以上,鲁棒性优于传统的基于字符分割的车牌识别方法和 LPR-Net 算法,这也是本文选择在 CRNN+CTC 网络模型基础上改进的原因。通过对比实验得出,本文算法在 3 个数据集下的  $RA$  与  $CRA$  均明显好于传统的检测方法和 LPR-Net 算法,尤其是在 CCPDv1 数据集下表现得更加突出,准确度分别比前两个算法增加了 13.4 和 9.5 个百分点,鲁棒性与泛化能力强。与改进前的 CRNN+CTC 相比,经过替换 CNN 和 RNN 层并引入加入了正则项的 CTC 的改进算法,对字符识别的准确性进一步提高,在一定程度上解决

了复杂环境下车牌识别的问题,特别是在 Reldv1 数据集下, $RA$  与  $CRA$  更是分别提高了 1.2 和 1.4 个百分点,鲁棒性与通用性得到进一步增强。另外,通过比较表中单张车牌的平均识别时间,不难发现改进后的方法较改进前所用时间大幅减少,在 GAN\_LP 和 CCPDv1 数据集下减少了约 33%,单张车牌平均识别时间降约 31 ms,识别速率得到明显提升。综上所述,该算法不仅提高了车牌字符识别准确率,同时减少了车牌识别的时间,复杂环境下车牌识别具有较好的识别效果,具有更强的泛化能力。

## 4 结束语

本文提出了一种基于 CRNN+CTC 深度神经网络智能车牌识别算法,使用深度可分离卷积和 Bi-LSTM (下转第 130 页)