

一种改进卷积循环神经网络的复杂场景下的车牌识别模型

洪顺贺¹, 胡宸滔², 铁治欣^{1,3*}, 丁成富⁴

¹浙江理工大学计算机科学与技术学院, 浙江 杭州

²浙江理工大学信息科学与工程学院, 浙江 杭州

³浙江理工大学科技与艺术学院, 浙江 绍兴

⁴聚光科技(杭州)股份有限公司, 浙江 杭州

收稿日期: 2023年3月24日; 录用日期: 2023年5月18日; 发布日期: 2023年5月25日

摘要

识别自然场景图像中的汽车牌照是一项重要而又具有挑战性的任务。许多现有方法对于在固定场景下收集的牌照表现良好, 但它们的性能在诸如车牌角度倾斜、光照强度过亮或过暗、图片模糊等复杂的环境中显著下降。本文提出了一种改进的卷积循环神经网络车牌识别模型, 在网络中加入幻影模块(Ghost Block)和卷积块注意模块(Convolutional Block Attention Module, CBAM), 能够提高车牌字符特征提取的丰富程度的同时在通道和空间方向上对车牌字符特征进行加权, 提高模型对车牌字符识别的准确率。最后通过实验验证了本文提出的模型的有效性。

关键词

车牌识别, Ghost Block, CBAM, LSTM

An Improved Convolutional Recurrent Neural Network for License Plate Recognition Model in Complex Scenes

Shunhe Hong¹, Chentao Hu², Zhixin Tie^{1,3*}, Chengfu Ding⁴

¹School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou Zhejiang

²School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou Zhejiang

³Keyi College, Zhejiang Sci-Tech University, Shaoxing Zhejiang

⁴Focused Photonics (Hangzhou) Inc., Hangzhou Zhejiang

Received: Mar. 24th, 2023; accepted: May 18th, 2023; published: May 25th, 2023

*通讯作者。

文章引用: 洪顺贺, 胡宸滔, 铁治欣, 丁成富. 一种改进卷积循环神经网络的复杂场景下的车牌识别模型[J]. 建模与仿真, 2023, 12(3): 2498-2504. DOI: 10.12677/mos.2023.123229

Abstract

Recognizing car license plates in images of natural scenes is an important and challenging task. Many existing methods perform well for license plates collected in fixed scenes, but their performance degrades significantly in complex environments such as tilted license plate angles, too bright or too dark light intensities, and blurred images. In this paper, we propose an improved convolutional recurrent neural network license plate recognition model by adding Ghost Block and Convolutional Block Attention Module (CBAM) module to the network, which can improve the richness of license plate character feature extraction while weighting the license plate character features in channel and spatial direction to improve the accuracy of the model for license plate character recognition. Finally, the effectiveness of the model proposed in this paper is verified by experimental evidence.

Keywords

License Plate Recognition, Ghost Block, CBAM, LSTM

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

车牌是机动车的一个重要的身份标识，对其的精准识别具有较高的应用价值，并已经成为最近的研究热点[1]。车牌识别技术已经被广泛应用于车辆身份认证、车辆密度统计、智能驾驶和交通场景理解等方面。

传统的车牌字符识别方法有模板匹配和垂直投影等方法。郑琳[2]采用模板匹配法对单个车牌字符图像进行识别，能有效降低字符图像的噪声与字符特征描述符冗余问题。曾夏明[3]利用 Hough 变换与像素点投影的方法来完成车牌的倾斜校正与分割，最后采用支持向量机来分类汉字字符与数字字母字符。但是传统方法非常容易受车牌图像中光照不均、噪声等因素的影响，导致车牌识别准确率低。深度学习技术的出现为复杂场景中的车牌图片识别任务提供新的途径。Zherzdev 和 Gruzdev 提出的 LPRNet [4]采用空间变换网络将车牌图像进行仿射变换，对车牌形状进行矫正，使得车牌图像更好地被识别。Gong 等人[5]提出的方法将车牌检测网络提取的特征与车牌识别网络共享，减小计算量，提高车牌识别速度。

基于深度学习技术，本文提出一种改进卷积循环神经网络车牌识别模型 LSTM-Ghost-CBAM，将传统卷积循环神经网络中的循环神经网络(Recurrent Neural Network, RNN)替换为长短期记忆网络(Long short-term memory, LSTM)，并且在网络中加入 Ghost 模块提高车牌字符特征提取的丰富程度，同时加入 CBAM 模块在通道方向和空间方向上对车牌字符特征进行加权，提高模型对车牌中单个字符识别的准确率。最后在 CCPD 数据集的最具有挑战的三个子数据集上对所提出的模型进行实际测试，实验结果表明本文提出的车牌识别模型能够在车牌角度倾斜、光照强度过亮或过暗、图片模糊等复杂环境中进行正确识别。

2. 模型结构

本文提出的车牌识别模型 LSTM-Ghost-CBAM 的总体框架如图 1 所示，先将输入的车牌图片进行预处理，将图像的大小调整为 128*64*3。在卷积特征提取模块，对图片进行特征提取，并进行下采样操作，

图片被转化为 $32*4*512$ 的特征向量。然后将得到的特征图使用幻影模块处理,提高模型的表达能力。再使用卷积块注意模块来从通道和空间方向上关注车牌中的字符特征,提高车牌识别的精度。最后使用长短期记忆网络解码器对提取到的车牌特征进行解码得到大小为 $32*66$ 的特征向量,并使用连接时间分类损失函数计算当前训练得到的序列与目标序列的误差值,最终得到车牌字符。

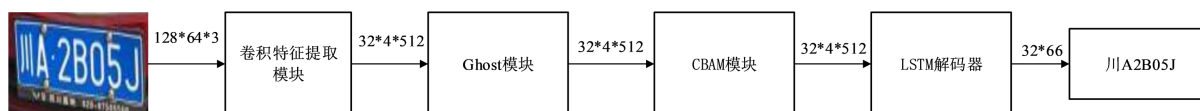


Figure 1. The general framework of license plate recognition model in this paper

图 1. 本文车牌识别模型总体框架图

2.1. 卷积特征提取模块

VGGNet [6]是一种经典的卷积神经网络,它在 ImageNet 图像识别挑战赛中取得了较好的成绩,被广泛应用于计算机视觉领域。所以在车牌特征提取部分,我们参考 VGGNet 设计了一个卷积特征提取器,利用其中的卷积层来提取输入数据的不同特征。卷积层的参数主要有卷积核大小、填充方式和步长大小,这三者共同决定了输出特征图的尺寸。

卷积核(filter)的宽度和高度都比较小,但是深度和输入数据一致,其作用可类比视觉皮层细胞的感受野。卷积核中的参数就是权重值,初始值可以通过随机得到,也可通过预训练得到,然后再通过误差反向传播算法不断更新。另外需注意卷积核的个数和输入图像通道数必须一致。

步长(stride)是卷积核在输入特征图中移动的位置间隔,卷积步长只对输入矩阵的长和宽这两个维度有效。

填充(padding)就是在图像周围使用指定的参数进行填充,常用 zero padding,即用 0 来填充。之所以需要填充,是因为随着卷积过程的进行,每次卷积都会导致图像尺寸的变小,如果图像很小、进行卷积的次数很多,最后可能只会剩下一个像素。

2.2. Ghost 模块

Ghost 模块[7]是一种在卷积神经网络中用于增加模型的通道数的模块,可以在不增加计算量的情况下提高模型的表达能力,同时也能够减少过拟合的风险。其核心思想是将输入特征图拆分成两部分,其中一部分保留原始的卷积特征,而另一部分则使用少量计算(cheap operations)的线性变换对特征进行提取,此处通常使用深度分离卷积。最后将两部分特征图进行拼接,通过这种方式,Ghost 模块可以在增加模型的宽度时不增加计算量,提高模型的表达能力,减少过拟合的风险。

2.3. CBAM 模块

CBAM 模块[8]是一种用于卷积神经网络中的注意力模块,主要包含两个部分:通道注意力和空间注意力。在通道注意力部分中,CBAM 模块利用全局平均池化操作和两个全连接层来生成通道注意力向量,该向量根据每个特征通道的重要性对特征进行加权,以确定哪些通道应该被强调。在空间注意力部分中,CBAM 模块利用不同的卷积核大小和池化操作来学习空间注意力图,该图根据每个空间位置的重要性对特征进行加权,以确定哪些位置应该被强调。最终,通道注意力向量和空间注意力图被相乘并与原始特征相加,形成最终的 CBAM 模块输出。

2.4. 长短期记忆网络解码器

LSTM 是一种特殊的 RNN,基本思想是引入记忆单元和门控机制,使网络能够选择性地存储和检索

信息,从而防止较早期的信号在处理过程中逐渐消失。一个 LSTM 主要包括三种类型的门结构:输入门,遗忘门和输出门,其模型框架如图 2 所示。

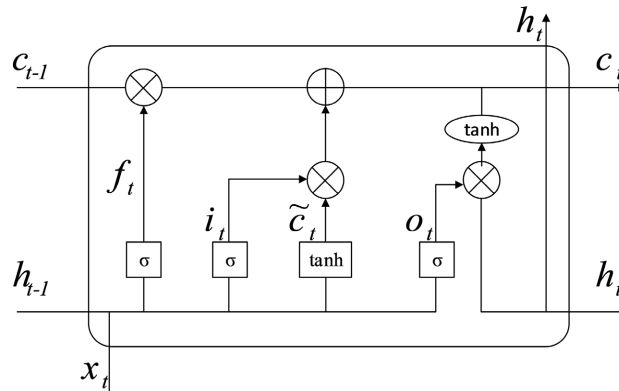


Figure 2. LSTM model framework structure
图 2. LSTM 模型框架结构图

在图 2 中, h_{t-1} 和 c_{t-1} 分别表示上一时刻的输出和细胞状态, x_t 、 h_t 和 c_t 分别表示当前时刻的输入、输出和细胞状态, \tilde{c}_t 表示当前暂时的细胞状态。 σ 表示一种门控结构,一般选用 Sigmoid 作为激活函数,其输出范围为 0 到 1,代表门限的开关,0 代表不许任何量通过,1 代表允许任意量通过。每个 LSTM 接收 c_{t-1} 、 h_{t-1} 、 x_t 三种输入信息,这些信息首先被送入遗忘门 f_t 和输入门 i_t 处理,其中遗忘门用于控制删除哪些信息,输入门用于筛选需要保留的信息,其公式如下:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5)$$

经过遗忘门和输入门处理后, t 时刻的状态也会更新:

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (6)$$

最后输出门将信息整合,生成输出信号为 o_t :

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \times \tanh(c_t) \quad (8)$$

3. 实验与结果

3.1. 实验数据集

CCPD [9]是一个大型的、多样化的中国城市车牌图像开源数据集。每张图像中仅包含一个车牌,每个车牌由 7 个字符组成,第一个字符为省份(共 31 个类别,不包含台湾省)、第二个字符为字母,剩下的

五个字符中的每一字符为字母或数字(所有出现的字母不包含“1”和“0”，数字和字母共34个类别)。按识别难度、车牌区域的照明、捕获时与车牌的距离、水平和垂直倾斜程度以及天气(雨、雪、雾)，该数据集被分为9个子数据集，如表1所示(详见文献[9]第三节表3)。我们在做模型对比实验时，将子数据集CCPD-base作为训练验证集，因为该子数据集包含的车牌图片最多，有20万个带有详细注释的独特车牌图像，其中第一个字符为“皖”的图片有19.2万张，占数据集总量的96%。在训练过程中我们随机选取CCPD-base中的一半数据作为训练集，另外一半用作为验证集。测试数据集是CCPD-DB、CCPD-Tilt和CCPD-Challenge这三个子数据集，因为这三个子数据集最能体现光照强度、车牌倾斜和车牌模糊等自然场景对车牌识别网络性能的影响。

Table 1. The situation of each sub-dataset in the dataset CCPD

表 1. 数据集 CCPD 中各子数据集情况表

子数据集名称	描述	图片数量
CCPD-Base	普通车牌图片	200 k
CCPD-FN	车牌离摄像头拍摄位置相对较近或较远	20 k
CCPD-DB	车牌区域亮度较亮、较暗或者不均匀	20 k
CCPD-Rotate	车牌水平倾斜 20 到 50 度，竖直倾斜-10 到 10 度	10 k
CCPD-Tilt	车牌水平倾斜 15 到 45 度，竖直倾斜 15 到 45 度	10 k
CCPD-Weather	车牌在雨雪雾天气拍摄得到	10 k
CCPD-Challenge	在车牌检测识别任务中较有挑战性的图片	10 k
CCPD-Blur	由于摄像机镜头抖动导致的模糊车牌图片	5 k
CCPD-NP	没有安装车牌的新车图片	5 k

3.2. 实验环境

车牌识别模型的实验环境如表2所示。

Table 2. Software and hardware parameters of the experimental platform

表 2. 实验平台的软硬件参数表

项目	参数
操作系统	Ubuntu 18.04.6 Linux
GPU	NAVIDA GTX 2080Ti
RAM	16 G
深度学习框架	Keras + Cuda11.0
编程语言	Python3.6

3.3. 实验细节

本文所使用的车牌识别算法训练过程中使用CTC损失函数和Adam优化器，batch size设置为512。初始学习率为0.4，epoch不设固定值，当连续5个epoch损失没有减少时，将学习率乘以衰减系数0.9，当连续30个epoch没有得到更低的损失时，就结束训练。

3.4. 模型训练

首先, 将车牌图片输入到 LSTM-Ghost-CBAM 组合模型中进行训练, 每进行一轮训练, 就使用 CTC 来计算当前序列与目标序列的误差值, 然后通过神经网络的误差反向传播机制不断更新模型的权重, 直到训练结束, 模型的误差更新过程如图 3 所示。可以看出经过 96 轮的迭代后, 模型的损失值不再下降, 训练结束, 在此过程中, 模型的损失值从 4.76 下降到 0.87, 收敛性良好。

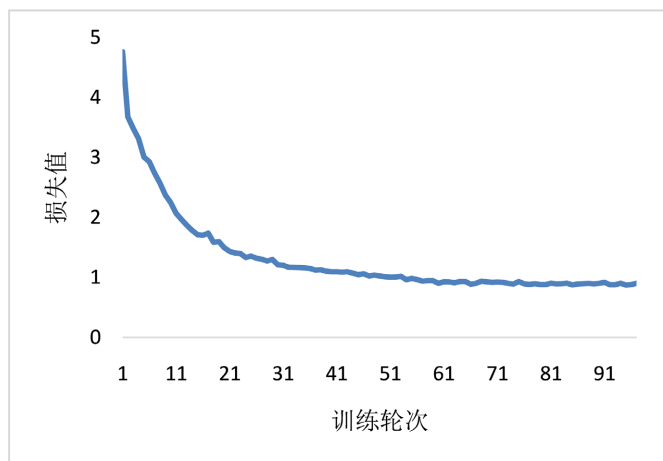


Figure 3. Loss curve during model training process
图 3. 模型训练过程中的损失曲线图

3.5. 模型识别效果验证

为了验证本文提出的 LSTM-Ghost-CBAM 模型的性能, 我们使用 3.1 节中所述的数据集对模型进行测试, 并与 LSTM、LSTM-Ghost、LSTM-CBAM 三种算法模型进行比较, 只有当车牌上的所有字符(包括汉字)都被正确识别时, 车牌识别结果才被认为是正确的, 实验测试结果如表 3 所示。

Table 3. Experimental results of different license plate recognition models on a subset of CCPD data
表 3. 不同车牌识别模型在 CCPD 数据子集上的实验结果表

模型	数据集	DB	Tilt	Challenge	平均精度
	LSTM		65.5	64.7	45.0
LSTM-Ghost		74.8	68.7	58.4	69.2
LSTM-CBAM		80.8	77.5	65.0	76.0
LSTM-Ghost-CBAM		82.5	79.0	65.3	77.3

表 3 中的实验结果表明, 使用 LSTM 模型在 CCPD 数据集上的测试精度低于 LSTM-Ghost 模型和 LSTM-CBAM 模型的测试精度, 这是因为 Ghost 模块可以提高特征提取的完整性, CBAM 模块可以在通道方向和空间方向上将字符相关的特征进行加权, 因此能够提高模型的单个字符识别准确性, 从而提高车牌识别整体精度。使用 LSTM-Ghost-CBAM 模型在 CCPD 数据集上的测试精度要高于其他三个模型的测试精度, 这说明 Ghost 模块和 CBAM 模块叠加使用能够显著性的提高模型对车牌图片的识别精度。

4. 总结

本文提出一种改进的卷积循环神经网络车牌识别模型 LSTM-Ghost-CBAM, 用于识别处于复杂场景中的车牌图片。该模型改进传统的卷积循环神经网络, 在卷积特征提取网络与 LSTM 解码器之间加入了 Ghost 模块和 CBAM 模块, 提高模型对车牌字符识别的准确率。并且选取 CCPD 数据集中最能体现光照强度、车牌倾斜和车牌模糊等场景的三个子数据集 CCPD-DB、CCPD-Tilt 和 CCPD-Challenge 进行测试实验, 测试的平均精度分别为 82.5%、79.0%和 65.3%。实验结果表明本文所提出的模型鲁棒性好, 能够完成在这些复杂场景中的车牌识别任务。但是由于训练使用的数据集中的中文分布不均匀, “皖” 占到 96%, 而测试数据集中还有训练数据集中没出现过的中文字符, 导致模型不能充分学习到中文信息, 影响测试精度。因此, 未来可以通过采集更多的车牌数据而得到中文字符分布均衡的数据集, 来提高车牌识别的准确率。

基金项目

浙江省公益技术研究项目(2015C31024)。

参考文献

- [1] Wang, W.H. and Tu, J.Y. (2020) Research on License Plate Recognition Algorithms Based on Deep Learning in Complex Environment. *IEEE Access*, **8**, 91661-91675. <https://doi.org/10.1109/ACCESS.2020.2994287>
- [2] 郑琳, 王福龙. 改进 HOG 特征的车牌识别算法[J]. 软件导刊, 2022, 21(5): 193-197.
- [3] 曾夏明. 基于 Hough 变换与 SVM 的高速公路车牌识别系统[J]. 西部交通科技, 2022(3): 175-180. <https://doi.org/10.13282/j.cnki.wcst.2022.03.046>
- [4] Zherzdev, S. and Gruzdev, A. (2018) Lprnet: License Plate Recognition via Deep Neural Networks. arXiv preprint arXiv:1806.10447.
- [5] Gong, Y., Deng, L., Tao, S., et al. (2022) Unified Chinese License Plate Detection and Recognition with High Efficiency. *Journal of Visual Communication and Image Representation*, **86**, Article ID: 103541. <https://doi.org/10.1016/j.jvcir.2022.103541>
- [6] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.
- [7] Han, K., Wang, Y., Tian, Q., et al. (2020) GhostNet: More Features from Cheap Operations. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 1577-1586. <https://doi.org/10.1109/CVPR42600.2020.00165>
- [8] Zhang, Z.X. and Wang, M.W. (2022) Convolutional Neural Network with Convolutional Block Attention Module for Finger Vein Recognition. arXiv preprint arXiv:2202.06673.
- [9] Xu, H., Zhou, X.D., Li, Z., et al. (2021) EILPR: Toward End-to-End Irregular License Plate Recognition Based on Automatic Perspective Alignment. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 2586-2595. <https://doi.org/10.1109/TITS.2021.3130898>