

基于DO-Conv改进的端到端车牌识别算法

赖珍向, 丁磊, 邓杰航, 顾国生

广东工业大学, 计算机学院, 广东 广州

Email: lai_zx@126.com

收稿日期: 2021年2月20日; 录用日期: 2021年3月15日; 发布日期: 2021年3月22日

摘要

提出了基于DO-Conv改进的端到端车牌识别算法, 该算法引入一种新颖的深度过参数化卷积(Depthwise Over-Parameterized Convolutional Layer, DO-Conv), 用于替换常规卷积, 提升图像特征提取的有效性, 从而提高车牌识别的准确率。研究表明: 改进后的算法在合成数据集和SYSU数据集上的识别准确率分别为97.42%和95.08%, 均优于使用传统卷积的端到端识别算法。

关键词

车牌识别, 端到端, 深度过参数化卷积, DO-Conv

An Improved End-to-End License Plate Recognition Algorithm Based on DO-Conv

Zhenxiang Lai, Lei Ding, Jiehang Deng, Guosheng Gu

School of Computer Science, Guangdong University of Technology, Guangzhou Guangdong

Email: lai_zx@126.com

Received: Feb. 20th, 2021; accepted: Mar. 15th, 2021; published: Mar. 22nd, 2021

Abstract

An improved end-to-end license plate recognition algorithm based on DO-Conv was proposed. A new Depthwise Over-Parameterized Convolutional Layer (DO-Conv) was introduced to replace the conventional convolution to improve the effectiveness of image feature extraction and the accuracy of license plate recognition. The results show that the recognition accuracy of the improved algorithm is 97.42% and 95.08% on the synthetic dataset and SYSU dataset respectively, both of which are better than the end-to-end recognition algorithm using traditional convolution.

Keywords

License Plate Recognition, End to End, Deep Over-Parameterized Convolution, DO-Conv

Copyright © 2021 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

随着我国汽车保有量的增加,城市面临的交通管理压力也越来越大,智能交通系统的建设迫在眉睫。车牌识别作为智能交通系统的重要组成部分,广泛应用于高速公路收费管理、停车场出入管理、道路交通监控等场景。

传统的车牌字符识别方法首先需要分割车牌字符,输出一组单个的字符图像,然后使用模板匹配[1]、特征提取[2]和神经网络[3]等识别算法对其进行识别,将各个字符图像的识别结果顺序组合起来便得到车牌的识别结果。这些方法很大程度上依赖字符分割的准确性,而且在识别时需要对大量的车牌特征进行提取,识别效率低。近年来,由于强大的特征学习能力,卷积神经网络(CNNs)在解决许多经典的计算机视觉问题上取得了巨大的成功,例如图像分类和语义分割。基于 CNN 的车牌识别算法[4] [5]也相继被提出来,用于解决错误的字符分割导致车牌识别率下降的问题。Li 等人[6]将车牌识别问题看作一个序列建模问题,利用 CNN 提取车牌区域特征,然后送入 LSTM [7]网络提取上下文关联信息,最后利用 CTC [8]算法进行判别得到车牌序列信息。Zherzdev 等人[9]提出一种无需使用递归神经网络(RNN)的轻型卷积神经网络,用于快速识别车牌字符。Zhang 等人[10]提出一个端到端的模型,直接通过 CNN 完成整个车牌的识别,比文献[6]中的方法更为简单。

上述基于 CNN 的车牌识别算法主要研究不同的网络结构对识别准确率的影响,他们由传统的卷积层组成。卷积层是 CNN 的核心构建模块。因此,这些核心模块的改进通常可以提高 CNN 的性能。一些用于替代传统卷积层的方法相继被提出来,这些方法可以提高 CNN 的特征学习能力或效率[11] [12]。深度过参数化卷积[13] (Depthwise Over-Parameterized Convolutional Layer, DO-Conv)由 Cao 等人提出,可以提高 CNN 在许多经典视觉任务上的性能,如图像分类、检测和分割。本文将其应用于车牌识别任务中,提出一种融合 DO-Conv 的车牌识别算法,用于提升车牌特征提取的有效性,从而提高车牌的识别准确率。

2. 深度过参数化卷积

与常规卷积不同,深度过参数化卷积 DO-Conv 在常规卷积的基础上添加一个用于过参数化的组件:深度卷积。一般认为,通过在网络中添加线性层和非线性层来增加网络深度可以提高网络的表达能力,提升网络性能。但是使用更深层次的网络结构往往会增加神经节点的计算量。过参数化的一个明显优势就是过参数化使用的多层复合线性操作在训练阶段之后可以折叠为紧凑的单层形式。然后在推理阶段仅使用单层计算,从而减少了计算量。此外,DO-Conv 不仅可以提升融合模型的性能,还可以加快 CNN 的训练速度。

深度过参数化卷积的结构如图 1 所示。深度卷积算子。首先应用于深度卷积内核 D 和输入特征 P ,生成变换后的特征 $P' = D \circ P$ 。然后常规卷积运算符 $*$ 应用于常规卷积内核 W 和特征 P' ,生成变换后的特征 $O = W * P'$ 。因此,深度过参数化卷积的输出可表示为: $O = W * (D \circ P)$ 。在图 1 中, M 和 N 为输入向量的空间维度, C_{in} 为输入向量的通道数, C_{out} 为输出的通道数, $d_{mul} = M \times N$ 为深度卷积的深度乘数。

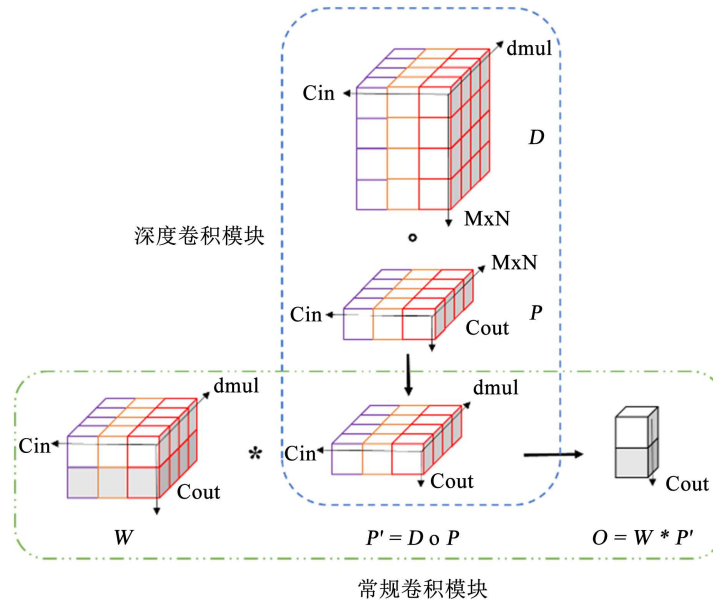


Figure 1. Deep over-parameterized convolutional structures
图 1. 深度过参数化卷积结构

3. 基于 DO-Conv 改进的端到端识别网络

本文提出的基于 DO-Conv 改进的端到端车牌识别模型包含 3 个模块：基于 DO-Conv 改进的特征提取模块、特征序列化模块和序列解码模块。图 2 给出了所提模型的流程图，图 3 给出了所提模型的网络结构，模型中 3 个模块说明如下。

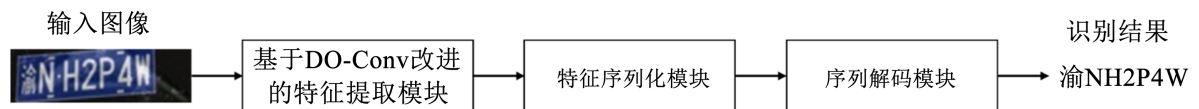


Figure 2. The flow chart of the proposed model
图 2. 所提模型流程图

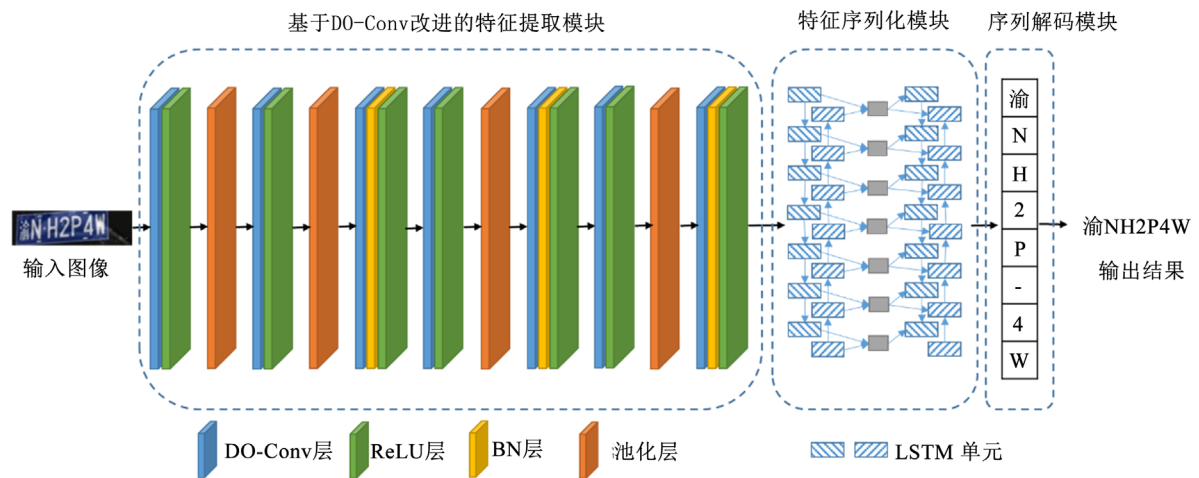


Figure 3. The proposed model network structure
图 3. 所提模型网络结构

3.1. 基于 DO-Conv 改进的特征提取模块

随着卷积神经网络的发展,其网络结构使用的卷积层数也在不断增加,在增强网络的表达能力时也带来了更大的计算量。针对这种情况,本文设计了一种基于 DO-Conv 改进的卷积神经网络,通过引入 DO-Conv 替换传统的卷积方式,可以在不增加网络层数的情况下提升网络的特征提取能力。

本文网络结构以 VGG 网络作为基础网络框架,其网络结构如图 3 所示。其中,输入图像统一处理为 110×32 大小的灰度图像。就网络卷积过程而言,整个卷积网络由 7 个 DO-Conv 层及最大池化层构成。此外, Batch Normalization 和 ReLU 函数被用来加速训练过程,提高网络的泛化能力。该网络通过对样本进行卷积、池化操作进行特征提取。

DO-Conv 层: 卷积网络中的特征提取层。卷积过程如第 1 章所述,卷积操作之后传递给激活函数,得到的结果是一组特征图。不同的特征图提取不同的特征。

池化层: 也叫下采样层,本文采用最大池化,主要操作是取区域上的一个值做下采样;其目的是进一步减少网络的参数,保持某种不变性(旋转、平移等)。

网络的具体参数设置如表 1 所示,其中 filters 表示卷积核的个数, k 表示卷积核的尺寸, s 表示步长, p 表示填充。

Table 1. Fusion of CNN parameter settings for DO-Conv

表 1. 融合 DO-Conv 的 CNN 参数设置

Type	Configuration
Input	$110 \times 32 \times 1$ gray-scale image
DO-Conv, ReLU	filters: 64, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
Max pooling	$k: 2 \times 2$, $s: 2 \times 2$
DO-Conv, ReLU	filters: 128, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
Max pooling	$k: 2 \times 2$, $s: 2 \times 2$
DO-Conv, BatchNorm, ReLU	filters: 256, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
DO-Conv, ReLU	filters: 256, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
Max pooling	$k: 2 \times 2$, $s: 2 \times 2$
DO-Conv, BatchNorm, ReLU	filters: 512, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
DO-Conv, ReLU	filters: 512, $k: 3 \times 3$, $s: 1 \times 1$, $p: 1$
Max pooling	$k: 2 \times 2$, $s: 2 \times 2$
DO-Conv, BatchNorm, ReLU	filters: 512, $k: 3 \times 3$, $s: 1 \times 1$, $p: 0$

3.2. 特征序列化模块

特征序列化模块是将提取出来的图像特征进行序列化,并实现与输入标签的序列到序列映射。本文采用长短时记忆网络(Long Short-Term Memory, LSTM)来进行序列化。整个模块由两层双向 LSTM 和全连接层组成,如图 2 所示。其中,每个双向 LSTM 层均由两个相对的 LSTM 构成,每个 LSTM 都有 128 个隐藏节点。全连接层有 71 个由 Relu 激活的神经元,对应 71 个类别(即 37 个中文字符,10 个数字,24 个字母,不含“ I ”、“ O ”)。当输入特征序列 X 时,全连接层的输出可表示为 $C = \{c_1, c_2, \dots, c_T\}$,其中 $c_t \in R^{71}$ 为第 t 个特征序列对应的序列标签。

3.3. 序列解码模块

本文采用链接时序主义分类(Connectionist Temporal Classification, CTC)来对特征序列化模块的输出进行解码, 预测最终的识别结果。CTC 通过使用 Softmax 层将循环层的输出 $\mathbf{C} = \{c_1, \dots, c_T\}$ 转换成概率序列 $\mathbf{Y} = \{y_1, y_2, \dots, y_T\}$, 其中 T 是特征序列的长度且

$$y_t = \text{softmax}(c_t) \tag{1}$$

首先, 用 L 表示任务中的所有标签。定义一组长度为 T 的序列 L' , $L' = L \cup \emptyset$ 。 L' 包含所有预定义的标签, 包括一个“blank” 标签 \emptyset 。我们将在全部时间步长里依次观察到的标签作为路径 π 。

为了处理路径 π 和目标序列 l 之间的关系, CTC 定义了多对一映射 β 。 β 首先删除重复的标签, 然后删除所有的空白标签。给定标签序列 l , 将可行路径定义为能够从 l 映射到 β 的所有 π 。每个目标的条件概率被定为所有可行路径概率的和, 即:

$$p(l|Y) = \sum_{\pi \in \beta(\pi)=l} p(\pi|Y) \tag{2}$$

其中

$$p(\pi|Y) = \prod_{t=1}^T y_t^{\pi_t}, \forall \pi \in L' \tag{3}$$

其中 π_t 是路径 π 中的第 t 个元素, $y_t^{\pi_t}$ 是标签 π_t 的评估概率。

4. 实验结果分析

4.1. 数据集

由于没有公开的大型中文数据集, 本文使用 opencv 根据国内车牌规则生成各种类型的合共 12 万张车牌图像的合成数据集, 图像的分辨率均为 110×32 。部分生成图像如图 4 所示。训练集和测试集之比为 8:2, 即训练集 104,000 张, 测试集 26,000 张。



Figure 4. Composite license plate image
图 4. 合成车牌图像

为了验证所提模型在真是环境中的性能, 本文还使用了开源中文车牌数据集 SYSU [14]。SYSU 数据集采集了超过 3000 张来自道路交通监控探头的车牌图像样本。由于该数据集未标注车牌位置, 采用开源软件 HyperLPR 对车牌区域进行定位, 成功定位 3671 张图像, 其中 2939 张被用于训练, 剩余 732 张被用于测试。

4.2. 实验设置

本文实验算法采用编程语言 Python、C++编写, 使用 pytorch 框架实现。其他软硬件配置如下: Intel Core i7-8700k 3.70 GHZ 12 核 8G 内存, NVIDIA 1080ti 显卡, 显存 11 GB, ubuntu16.04LTS 系统。

网络训练的学习率为 $1e-5$ ，模型的 batch size 为 128，模型使用 RMS 优化器进行优化。

4.3. 评价指标

本文采用识别准确率(RA) [15]来评估车牌识别模型的性能，计算公式如下：

$$RA = \frac{R}{N} \quad (4)$$

其中 R 是正确识别所有字符的车牌数量， N 是用于测试的车牌数量。

4.4. 实验结果

为了探讨 DO-Conv 卷积层对车牌识别准确率的影响，本文将模型中的 DO-Conv 层替换为常规卷积，然后分别在 SYSU 数据集和合成车牌数据集上进行消融实验。实验结果如表 2 所示。

Table 2. Comparison of structural performance of different models

表 2. 不同模型结构性能对比

模型结构	SYSU 数据集识别准确率	合成数据集识别准确率
使用常规卷积的模型	94.81%	96.90%
本文模型	95.08%	97.42%

由表 2 可知，本文所提模型在 SYSU 数据集和合成数据集上分别比使用常规卷积的模型高 0.27% 和 0.52%，两者的性能差距证实了本文提出的识别模型有助于提升车牌识别的准确率。绘制了合成数据集中测试准确率与训练次数关系图，如图 5 所示。从该图中可以看出，在训练过程中使用 Do-Conv 卷积层的网络的准确率上升速度更快，且明显高于使用传统卷积层 Conv 的准确率曲线，表明使用 DO-Conv 层不仅可以提高车牌识别的准确率，还可以加快神经网络的训练。

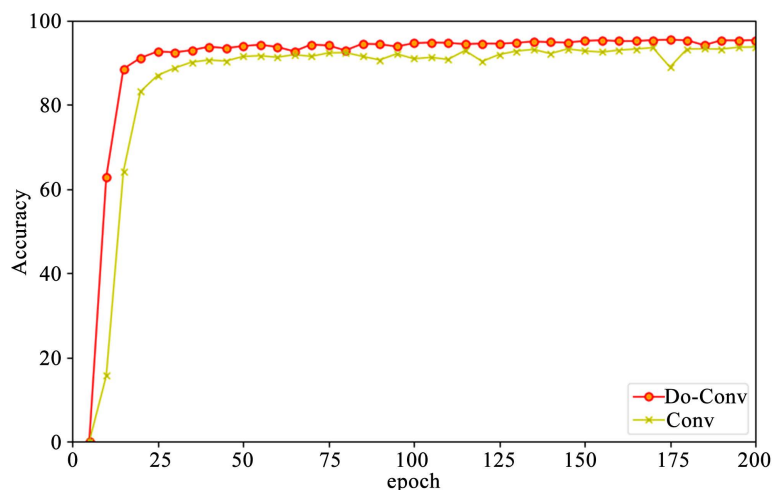


Figure 5. Relationship between test accuracy and training times in synthetic data set

图 5. 合成数据集中测试准确率与训练次数关系

此外，为了进一步验证本文模型的有效性，还与文献[9]的方法进行了对比，实验的结果如表 3 所示。与文献[9]的方法相比，本文提出的基于 DO-Conv 改进的端到端车牌识别方法在两个数据集里都获得了较高的准确率。

Figure 3. Performance comparison of different methods
表 3. 不同方法性能对比

模型结构	SYSU 数据集识别准确率	合成数据集识别准确率
文献[9]	94.06%	94.17%
本文模型	95.08%	97.42%

5. 结论

由于大多数基于 CNN 的车牌识别算法主要研究不同网络结构对车牌识别准确率的影响,很少有方法研究不同卷积层对识别准确率的影响。本文提出了基于 DO-Conv 改进的端到端车牌识别算法。该算法对传统的端到端算法进行改进,在特征提取阶段将原有的卷积方式替换为为深度过参数化卷积 DO-Conv,提升图像特征提取的有效性,然后使用 LSTM 网络对提取的特征进行序列化,最后使用 CTC 算法进行解码输出。通过在合成车牌数据集和 SYSU 数据集上的实验表明,改进后的网络模型对于车牌识别的准确率更高,分别达到了 97.42% 和 95.08%。同时,改进后的模型的训练效果更好。与现有方法的对比也进一步表明了本文所提方法的有效性。

参考文献

- [1] 楚天鸿,唐瑞尹. 基于 MATLAB 平台下的车牌识别系统设计[J]. 科技与创新, 2020(14): 20-22.
- [2] 肖雄,黄樟灿. 基于统计的车牌字符识别[J]. 数字技术与应用, 2016(4): 74.
- [3] 杨洪臣,党京,蔡能斌,王华鹏,翟金良. BP 神经网络车牌识别技术优化研究[J]. 中国刑警学院学报, 2019(3): 120-124.
- [4] Wang, W.W., Yang, J., Chen, M. and Wang, P. (2019) A Light CNN for End-to-End Car License Plates Detection and Recognition. *IEEE Access*, **7**, 173875-173883. <https://doi.org/10.1109/ACCESS.2019.2956357>
- [5] Zou, Y.J., Zhang, Y.J., Yan, J., Jiang, X.X., Huang, T.J., Fan, H.S., et al. (2020) A Robust License Plate Recognition Model Based on Bi-LSTM. *IEEE Access*, **8**, 211630-211641. <https://doi.org/10.1109/ACCESS.2020.3040238>
- [6] Li, H., Wang, P., You, M.Y. and Shen, C.H. (2018) Reading Car License Plates Using Deep Neural Networks. *Image & Vision Computing*, **72**, 14-23. <https://doi.org/10.1016/j.imavis.2018.02.002>
- [7] Wang, X., Han, Y., Leung, V.C.M., Niyato, D., Yan, X.Q. and Chen, X. (2020) Convergence of Edge Computing and Deep Learning: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, **22**, 869-904. <https://doi.org/10.1109/COMST.2020.2970550>
- [8] Huang, X.H., Qiao, L.S., Yu, W.T., Li, J. and Ma, Y.Z. (2020) End-to-End Sequence Labeling via Convolutional Recurrent Neural Network with a Connectionist Temporal Classification Layer. *International Journal of Computational Intelligence Systems*, **13**, 341-351. <https://doi.org/10.2991/ijcis.d.200316.001>
- [9] Zherzdev, S. and Gruzdev, A. (2018) LPRNet: License Plate Recognition via Deep Neural Networks. arXiv preprint arXiv:1806.10447.
- [10] Zhang, L.J., Wang, P., Li, H., Li, Z., Shen, C.H. and Zhang, Y.N. (2020) A Robust Attentional Framework for License Plate Recognition in the Wild. *IEEE Transactions on Intelligent Transportation Systems*, 1-10. <https://doi.org/10.1109/TITS.2020.3000072>
- [11] Chen, Y.P., Fan, H.Q., Xu, B., Yan, Z.C., Kalantidis, Y., Rohrbach, M., et al. (2019) Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 3434-3443. <https://doi.org/10.1109/ICCV.2019.00353>
- [12] Zhu, X.Z., Hu, H., Lin, S. and Dai, J.F. (2019) Deformable ConvNets V2: More Deformable, Better Results. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 9300-9308. <https://doi.org/10.1109/CVPR.2019.00953>
- [13] Cao, J., Li, Y., Sun, M., Chen, Y., Lischinski, D., Cohen-Or, D., et al. (2020) DO-Conv: Depthwise over Parameterized Convolutional Layer. arXiv preprint arXiv:2006.12030.
- [14] Zhao, Y., Yu, Z., Li, X. and Cai, M. (2019) Chinese License Plate Image Database Building Methodology for License

-
- Plate Recognition. *Journal of Electronic Imaging*, **28**, Article ID: 013001. <https://doi.org/10.1117/1.JEI.28.1.013001>
- [15] Wang, J., Huang, H., Qian, X., Cao, J. and Dai, Y. (2018) Sequence Recognition of Chinese License Plates. *Neuro-computing*, **317**, 149-158. <https://doi.org/10.1016/j.neucom.2018.08.023>