

快准狠！Intel论文揭示自家车牌识别算法:LPRNet

52cv.net2018年7月20日

(关注52CV——有价值有深度的公众号~)

来自工业界的最佳实践。

车牌识别是一个老生常谈的话题，在工业界已经得到广泛应用。当深度学习在各种视觉识别任务上刷新更高精度的时候，却常常被认为计算量远大于传统方法。Intel公司俄罗斯IOTG计算机视觉组的工程师最近发布了一篇论文，揭示了自家已经商用的车牌识别算法，使用轻量级深度神经网络进行车牌识别，达到快准狠的新高度，即速度超快、精度超准、硬件利用率超狠。

本文来自于论文《LPRNet: License Plate Recognition via Deep Neural Networks》。

LPRNet: License Plate Recognition via Deep Neural Networks

Sergey Zherzdev
ex-Intel*

IOTG Computer Vision Group
sergeyzherzdev@gmail.com

Alexey Gruzdev
Intel

IOTG Computer Vision Group
alexey.gruzdev@intel.com

文章的第一作者已经离开Intel，这篇论文是他们17年的工作，通过Linkedin得知，两位作者来自Intel收购的Itseez公司，也就是之前维护OpenCV的俄罗斯公司。

该文提出了LPRNet – 自动车牌识别的end-to-end方法，识别之前无需进行初步的字符分割。该方法使用了深度神经网络，能够实时运算，在中国车牌识别准确度上高达95%，速度上在nVIDIA GeForce GTX 1080显卡运算每个车牌3ms，在英特尔酷睿i7-6700K上每个车牌1.3ms。LPRNet由轻量级卷积神经网络组成，因此可以端到端的方式进行训练。论文称，LPRNet是第一个不使用RNN的实时车牌识别系统。因为速度快，LPRNet算法可用于自动车牌识别的嵌入式解决方案，即使在具有挑战性的中国车牌上也具有高精度。

需要说明的是，LPRNet解决的是识别的问题，文中车牌检测使用的是LBP-cascade。

LPRNet特性

- 1.实时、高精度、支持车牌字符变长、无需字符分割、对不同国家支持从零开始end-to-end的训练；
- 2.第一个不需要使用RNN的足够轻量级的网络，可以运行在各种平台，包括嵌入式设备；
- 3.鲁棒，LPRNet已经应用于真实的交通监控场景，事实证明它可以鲁棒地应对各种困难情况，包括透视变换、镜头畸变带来的成像失真、强光、视点变换等。

车牌识别的挑战

图像模糊、很差的光线条件、车牌数字的变化（比如中国和日本的车牌有一些特殊字符）、车牌变形、天气影响（比如雨雪天气）、车牌上的字符个数有变化。



Figure 1. Example of LPRNet recognitions

空间变换预处理LocNet

这是对检测到的车牌形状上的校正，使用 Spatial Transformer Layer[1]，这一步是可选的，但用上可以使得图像更好得被识别。

Layer Type	Parameters	
Input	94x24 pixels RGB image	
AvgPooling	#32 3x3 stride 2	—
Convolution	#32 5x5 stride 3	#32 5x5 stride 5
Concatenation	channel-wise	
Dropout	0.5 ratio	
FC	#32 with TanH activation	
FC	#6 with scaled TanH activation	

Table 1. LocNet architecture



LPRNet的基础构建模块

LPRNet的基础网络构建模块受启发于SqueezeNet Fire Blocks[2]和Inception Blocks[3],如下图所示。

Layer Type	Parameters/Dimensions
Input	$C_{in} \times H \times W$ feature map
Convolution	# $C_{out}/4$ 1x1 stride 1
Convolution	# $C_{out}/4$ 3x1 strideh=1, padh=1
Convolution	# $C_{out}/4$ 1x3 stridew=1, padw=1
Convolution	# C_{out} 1x1 stride 1
Output	$C_{out} \times H \times W$ feature map

Table 2. Small Basic Block

特征提取骨干网架构

骨干网将原始的RGB图像作为输入，计算得到空间分布的丰富特征。为了

利用局部字符的上下文信息，该文使用了宽卷积（ 1×13 kernel）而没有使用LSTM-based RNN。骨干网络最终的输出，可以被认为是一系列字符的概率，其长度对应于输入图像像素宽度。

由于解码器的输出与目标字符序列长度不同，训练的时候使用了CTC Loss[4]，它可以很好的应对不需要字符分割和对齐的end-to-end训练。

Layer Type	Parameters
Input	94x24 pixels RGB image
Convolution	#64 3x3 stride 1
MaxPooling	#64 3x3 stride 1
Small basic block	#128 3x3 stride 1
MaxPooling	#64 3x3 stride (2, 1)
Small basic block	#256 3x3 stride 1
Small basic block	#256 3x3 stride 1
MaxPooling	#64 3x3 stride (2, 1)
Dropout	0.5 ratio
Convolution	#256 4x1 stride 1
Dropout	0.5 ratio
Convolution	# class_number 1x13 stride 1

Table 3. Back-bone Network Architecture

为了进一步提高性能，使用了论文[5]中global context嵌入。

推理阶段对上述一系列字符的概率进行解码，使用beam search[6],它可以最大化输出序列的总概率。

后过滤（post-filtering）阶段，使用面向任务的语言模型实现作为目标国家车牌模板的一组集合，后过滤阶段是和beam search 结合在一起用的，获得通过beam search找到的前N个最可能序列，返回与预定义模板集合最匹配的序列，该模板取决于特定国家的车牌规则。

识别实验结果

训练时，使用一个来自监控场景的中国车牌的私有库，总共有11696幅经过LBP级联检测器检测出来的车牌，并进行了数据增广（data augmentation）即随机旋转、平移、缩放，下图报告了上述各种tricks对识别精度的影响。

Approach	LPRNet							
Global context				✓	✓	✓	✓	✓
Data augm.	✓	✓	✓		✓	✓	✓	✓
STN-alignment		✓	✓			✓	✓	✓
Beam Search			✓				✓	✓
Post-filtering			✓					✓
Accuracy, %	53.4	58.6	59.0	62.95	91.6	94.4	94.4	95.0

Table 5. Effects of various tricks on LPRNet quality.

最大的识别精度增益来自于global context (36%)，其次是data augmentation (28.6%)，STN-based alignment即预处理也带来了显著提高 (2.8-5.2%)，Beam Search联合post-filtering进一步提高了0.4-0.6%。

识别速度

Intel将LPRNet在CPU\GPU\FPGA上都进行了实现，每个车牌的识别时间如下：

Target platform	1 LP processing time
GPU + cuDNN	3 ms
CPU (using Caffe [22])	11-15 ms
CPU + FPGA (using DLA [23])	4 ms ¹
CPU (using IE from Intel OpenVINO [24])	1.3 ms

Table 6. Performance analysis.

这里GPU用的是nVIDIA GeForce1080, CPU是Core i7-6700K SkyLake, FPGA是Intel Arria10, 推断引擎IE来自Intel OpenVINO.

虽然这篇文章本身没有什么新的发明，但52CV还是认为非常值得推荐给大家的，它绝不属于水文，因为Intel已经将其商用了，足以证明它的优势和价值。

该文没有开源代码，论文地址：

<https://arxiv.org/abs/1806.10447v1>

在“我爱计算机视觉”公众号后台回复“lprnet”可以直接获取论文的百度网盘

盘下载地址。

参考文献

[1]“Spatial Transformer Networks,” arXiv:1506.02025

[2]“SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5mb model size,” arXiv:1602.07360

[3]“Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning,” arXiv:1602.07261

[4]Connectionist temporal classification:labelling unsegmented sequence data with recurrent neural networks. ICML 2006:369-376

[5]“ParseNet: Looking Wider to See Better,” arXiv:1506.04579

[6]Supervised Sequence Labelling with Recurrent Neural Networks, 2012th ed. Heidelberg ; New York:Springer, Feb. 2012.

转载请注明：《[快准狠！Intel论文揭示自家车牌识别算法:LPRNet](#)》