

复杂环境下基于角点回归的全卷积神经网络的车牌定位

罗斌 郜伟 汤进 王文中 李成龙

(安徽大学计算机科学与技术学院, 合肥, 230601)

摘要: 车牌定位是车牌识别系统中核心部分, 具有较高的研究和应用价值。尽管近些年来该研究取得了很大的进展, 但仍无法很好地解决低亮度、低分辨率和车辆倾斜等环境下的定位问题。本文提出了一种新的全卷积神经网络, 通过回归车牌角点的方式准确地进行车牌定位。为了保证训练的有效性, 对 45 000 幅含有车牌的图像进行人工标注。同时, 对标注的图像随机进行平移、缩放、旋转和加噪, 提高训练样本的数量和多样性。在本文构建的卡口图像数据集和复杂环境数据集上与两种方法进行了比较, 验证了本文方法的有效性。

关键词: 卷积神经网络; 车牌定位; 深度学习; 角点回归; 复杂环境

中图分类号: TP391.4 **文献标志码:** A

Learning Corner Regression-based Fully Convolutional Neural Network for License Plate Localization in Complex Scene

Luo Bin, Gao Wei, Tang Jin, Wang Wenzhong, Li Chenglong

(School of Computer Science and Technology, Anhui University, Hefei, 230601, China)

Abstract: License plate localization, the core component of license plate recognition system, is valuable in both academic development and potential applications. Though much progress has been made in recent years, challenging problems still exist in the complex scenes, such as low luminance, low resolution and inclination scene of vehicle. This paper proposes a novel fully convolutional neural network to localize license plates accurately by a corner regression algorithm. To guarantee effective training in the proposed model, 45 000 sample images are annotated by one person. Meanwhile, the annotated sample images are processed by four operators, including translating, scaling, rotating and noising, to increase the number and diversity of the training samples. Extensive experiments on the newly collected datasets, traffic monitoring dataset and the complex scene dataset, demonstrate the effectiveness of the proposed method against other two license plate localization methods.

Key words: convolutional neural network; license plate localization; deep learning; corner regression; complex scene

引 言

近年来, 车牌识别技术受到了广泛的关注和研究, 并得到了大量的应用实践。传统的车牌识别技术

主要包含车牌定位、字符分割和识别 3 部分。其中车牌定位是整个技术的核心步骤,常用的方法主要采用边缘检测^[1,2]、水平灰度变化特征以及车牌颜色特征^[3],但都基于标准交通卡口数据进行,且定位限制条件较多。另一方面,随着建设平安城市步伐加快,监控摄像头分布越来越广,这些监控环境比标准卡口复杂,但安放和维修成本较低。而这些复杂环境下的车牌定位问题无法通过传统方法解决,研究不局限于环境的车牌定位算法已成为趋势。

文献[3]综合应用了常用的车牌定位方法解决复杂车辆图像的车牌定位问题。主要过程为基于边缘检测和数学形态学方法,去噪获得粗定位候选区域,并基于颜色信息对非车牌进行过滤,最后利用投影和谷值分析对车牌区域进行精细化定位。这些定位方法对环境限制较高,如车牌位置固定、亮度较高、车牌颜色没有失真、车牌位置较为方正以及车牌图像非常清晰等^[4]。对于其他有挑战并常见的场景,如夜间拍摄补光灯不强、远处车牌以及侧拍车牌,如图 1 所示,以上方法都不能较好地定位车牌。



图 1 复杂环境下车牌
Fig. 1 Plates of complex scene

文献[5]利用卷积神经网络解决分类、定位和检测问题。该文定位使用分类训练完成的前 5 层卷积神经网络参数,并在其后添加两层全连接作为定位回归网络,经过两层全连接后,每 1 类输出 4 个值作为定位边界框的坐标值。该文虽解决了分类、检测和定位问题,但是利用滑动窗对特征图进行操作效率较为低下,需要对多个检测框进行合并操作。文献[6]的定位方式为根据卷积层产生的特征图进行定位,但该文并未对如何选择特征图以及特征图如何处理进行详细研究,故并不能直接完成定位任务。本文设计的卷积神经网络模型不含全连接层,故测试图像可为任意尺度,模型输出特征图为两幅车牌角点置信图,根据置信图生成车牌的左上角点和右下角点,从而确定车牌位置。为了提高该网络的鲁棒性,本文对 45 000 幅图像进行了标注,并进行数据增强处理。实验表明,本文算法能够在复杂环境中较好地定位车牌,且精度大幅度地优于其他方法。

1 基于角点回归的全卷积神经网络

1.1 基本概念

文献[7]详细阐述了深度卷积神经网络的概念。卷积神经网络工作过程分前向传播和反向传播两个步骤。前向传播对输入的图像数据进行多层卷积运算,利用设计的损失函数对卷积结果进行计算获得损失值;基于损失值沿卷积方向进行反向传播,主要为更新卷积核权值的过程^[8]。本文的全卷积神经网络的任务就是回归出测试图像两幅车牌角点置信图。车牌角点指的是车牌矩形区域的两个正对角线顶点,即车牌左上角点和右下角点,此两点描述了车牌位置信息。通过对车牌两角点进行匹配即可获得车牌的矩形框位置。置信图和测试图像具有一定的线性关系,置信图中每一个值表示对应原图中某个像素点或者某块区域为车牌某角点的置信值。利用相关算法处理置信图能够找到焦点在测试图像中的位置。

1.2 模型设计

由于含有全连接网络的卷积神经网络限制了输入图像的尺度且全连接需要的权值数据繁多,导致训练和测试卷积神经网络模型时效率低下。为了解决输入图像尺度限制和模型收敛慢等问题,本文使用了深度全卷积神经网络模型。深度全卷积神经网络指的是深度卷积神经网络层中除了卷积层、下采样层和 ReLU^[8]层外,不含有其他层网络结构。经过多次实验,本文采用的全卷积神经网络如图 2 所示。

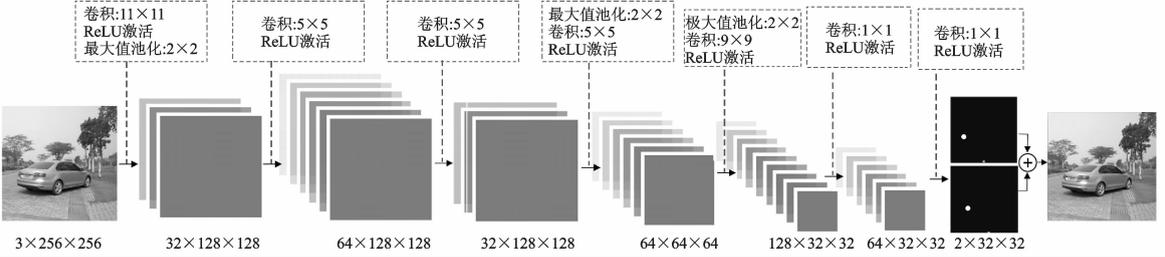


图 2 全卷积神经网络结构

Fig. 2 Fully convolutional neural network structure

本神经网络模型共由 7 层卷积层和 3 层池化层构成。第 1 层卷积具有 32 个 $11 \times 11 \times 3$ 的卷积核,第 2 层卷积具有 64 个 $5 \times 5 \times 32$ 的卷积核,第 3 层卷积具有 32 个 $5 \times 5 \times 64$ 的卷积核,第 4 层卷积具有 64 个 $5 \times 5 \times 32$ 的卷积核,第 5 层具有 128 个 $9 \times 9 \times 64$ 的卷积核,第 6 层具有 64 个 $1 \times 1 \times 128$ 的卷积核,第 7 层卷积具有 64 个 $1 \times 1 \times 64$ 的卷积核。其中在第 1 层卷积、第 3 层卷积和第 4 层卷积后以步长为 2,大小为 2×2 的窗口求最大值的方式进行下采样操作,并且在每层卷积后利用 ReLU 函数进行激活,以加快模型训练收敛。

如图 2 所示,输入为含有车辆的原始图像,图像进入 7 层卷积后,会产生两幅图像中车牌的左上角点和右下角点对应的置信图,值接近 1,图中圆圈表示此处出现车牌特征点的概率较大。利用输出的置信图进行线性变换可获得在测试图像上车牌对应的左上角点和右下角坐标值,将两点信息匹配合并即可获得车牌位置信息。

将已标记的车牌位置信息 A 和 B 进行转换得出 A' 和 B' ,转换方程为

$$A' = \alpha A + \beta \tag{1}$$

式(1)中的 α 和 β 为网络转换参数,和网络结构相关。图 3 表示神经网络模型中的第 i 层计算过程:输入宽为 W_i ,高为 H_i 的数据,利用 $sk \times sk$ 、步长为 stride 的卷积核对输入数据进行卷积,输出宽 W_{i+1} ,高 H_{i+1} 特征图,其中 A_{i+1} 点对应于输入数据中的 A_i 点,推算可知转换方式为

$$A_{i+1} = \frac{A_i + 0.5(\text{Padding} - sk + 1)}{\text{stride}} \tag{2}$$

为了方便卷积计算,对输入图像宽和高分别扩充 Padding 个数值,图 3 中 Padding 部分中 Padding 取值为 1。计算出的特征图宽 W_{i+1} 和高 H_{i+1} 为

$$W_{i+1} = \frac{W_i + \text{Padding} - sk}{\text{stride}} + 1 \tag{3}$$

$$H_{i+1} = \frac{H_i + \text{Padding} - sk}{\text{stride}} + 1 \tag{4}$$

对整个网络模型进行计算后,可得 α 和 β 的值,本文中的网络模型 α 取 0.125, β 取 0。计算出 A' (x_A, y_A) 后与高斯分布进行叠加得出 GT_1 置信图, x_A 和 y_A 表示点 A' 的 x 轴和 y 轴坐标。设定 A' 点的置信值为 1,其他点符合高斯分布, $f(x, y)$ 为置信图中坐标为 (x, y) 点的置信值,计算式为

$$f(x, y) = e^{-\frac{1}{2}(x-x_A)^2 + (y-y_A)^2} \quad (5)$$

同理, B 点利用式(1)和式(5)计算得出观测置信图值 GT_2 。训练模型时,以 GT_1 和 GT_2 作为观测值进行损失值计算。利用全卷积神经网络输出的两幅角点置信图可以确定车牌位置信息。将置信图以阈值0.6进行二值化,利用8连通区域检测算法进行连通区域计算,获取连通区域中置信值最大的点为该区域中的角点映射,获取角点映射坐标,并利用式(1)的反函数求得测试图像中对应位置。最终根据两幅特征图获得的左上角点和右下角点空间关系进行配对,得出车牌区域即为车牌位置信息。

2 模型训练

全卷积神经网络结构设计完成后,需要适当的训练方式才能获得有效模型。本文的网络模型训练过程如图4所示。主要步骤为将标记的车牌特征点(图4中 A, B 点)生成两幅置信图 GT_1 和 GT_2 ,分别表示 A 点和 B 点对应的置信图,作为模型的观察值。 FM_1 和 FM_2 为经过卷积神经网络生成的特征图作为置信图的预测值。根据观察值和预测值进行损失函数计算获得损失值,将损失值进行反向传播进行权值更新从而获得训练完成的模型。

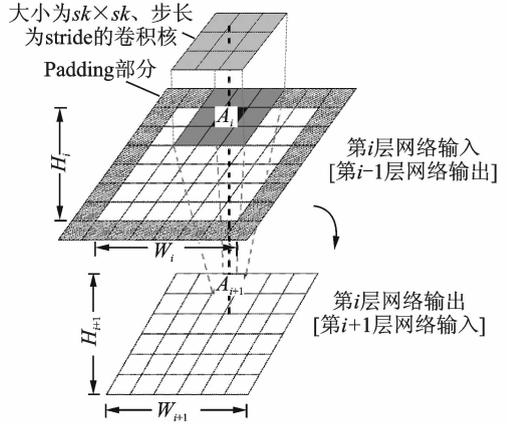


图3 坐标点转换原理

Fig. 3 Principle of coordinate transformation

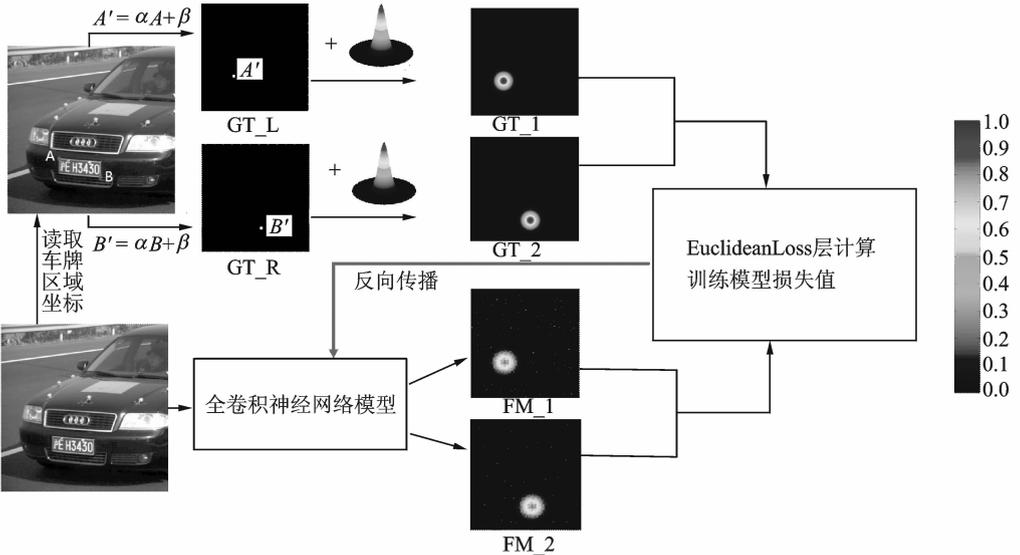


图4 模型训练流程图

Fig. 4 Training flow diagram of model

学习检测车牌特征点的卷积神经网络模型必须要具备大量的数据,本文原始数据分成交通卡口数据和监控摄像数据,为了在大量数据中检测出车辆图像,本文使用了自动标注方法^[9]来对车辆图像进行过滤和标注。最终获取45 000幅车辆图像数据后,设计实现了车牌标注软件并进行标注。为了使训练样本具备多样性,对标注的图像随机进行平移、缩放、旋转、添加高斯噪声以及椒盐噪声去除等操作,使原始数据增强到100 000幅,其中椒盐噪声去除采用了文献^[10]中的算法。

为了准确回归出两幅置信图,本文采用了计算 Euclidean 的距离作为损失函数,表示为

$$\text{EuclideanLoss} = \frac{1}{n} \left(\sum_{i=0}^{W_{\text{gt}} \times H_{\text{gt}}} (vp_j - vs_j)^2 \right)^{0.5} \quad (6)$$

式中:EuclideanLoss 为通过损失函数计算出的损失值; H_{gt} 和 W_{gt} 为置信图高度和宽度; vp_j, vs_j 分别为特征点置信图中第 j 个预测值和观察值; n 为特征点置信图像素点个数,取值为 $H_{\text{gt}} \times W_{\text{gt}}$ 。

实验证明,当置信值大于 0.6 时,该点对应的像素点为车牌特征点概率较高。根据图 3 可知,在一幅置信图中,置信值接近 0 的个数为 90% 以上,只有 2% 左右的置信值大于 0.6。该情况导致样本学习分布不均匀,无法使模型最优化。为了解决此问题,在实际训练过程中,首先计算置信值大于 0.6 的数据个数 N ,再从置信值小于等于 0.6 的数据中随机不重复选取 N 个数据用来进行反向传播更新权值,未选择的数据将不参与反向传播过程。实验证明此方法有效地加速了模型收敛。

3 实验结果与分析

3.1 数据集

车牌定位方法不仅需要在复杂环境正常运作,还需要在标准环境下具有很高的鲁棒性。为了验证以上两点,本文实验测试数据集共有 3 047 幅,主要包含了复杂环境数据和标准交通卡口图像数据。标准交通卡口用于测试在非复杂环境下本算法的鲁棒性,复杂环境数据用于验证本文方法和传统车牌定位方法相比的优越性。复杂环境数据集部分数据采集算法使用了文献[11]的抓拍算法。标准交通卡口图像指的是交通路口中卡口相机获取的图像,主要场景为高速路口以及车库入口处图像数据,利用传统车牌定位方法即可有效完成对标准卡口图像定位操作,虽本文方法用于解决复杂环境车牌识别问题,但同时需要在标准卡口图像数据上具有有效性,才能保证本文方法在绝大部分场景的鲁棒性。复杂环境数据主要分为低光照、低分辨率以及车牌倾斜度较大的图像数据,实验证明这些数据利用其他的车牌定位方法无法完成正常定位,故下文试验中并未利用传统方法与本文方法作对比试验,仅对比了标准交通卡口图像的数据与传统方法的优劣。

本文实验硬件环境为型号 i7-4770,频率 3.40 GHz 处理器,16 GB 的内存。由于 GPU 加速图像处理^[12],实验采用了型号 GTX980 的 GPU,卷积神经网络在 Caffe^[13] 开源框架下实现,对比实验在 Visual Studio 2010 上完成。实验内容主要分成两部分:标准交通卡口对比实验和复杂环境验证实验。标准卡口数据集下的测试实验主要用来验证本文方法在标准环境中的鲁棒性,并和其他传统算法做对比实验。复杂环境数据测试用来验证本文方法在光照低、车辆倾斜和低分辨率的环境中的有效性。在标准卡口数据集测试中,本文实现了两种常用定位算法:基于灰度变化特征算法和基于边缘检测定位算法^[14]。基于灰度变化特征算法主要过程为将车牌图像转换成灰度图,并将灰度图转成 0,1 二值图像,计算每一行 0 值和 1 值的跳变次数,当次数大于一定阈值时,则估计为车牌候选行,在候选行进行垂直方向上投影能够初步定位车牌位置。根据文献[14]实现了基于边缘检测定位算法。利用这两种算法与本文方法在标准交通数据卡口上进行对比和分析,得出本文算法的优缺点。

3.2 实验

3.2.1 模型特征图提取与分析

利用本文提出的卷积神经网络结构经 10 000 次迭代后趋于收敛。模型收敛并不能表明模型有效性,还需要验证模型对车牌区域的敏感度,一般而言,模型对该对象的敏感度越高,则该模型对该对象的识别或定位效果越好。通过实验提取第 1 层卷积网络权值如图 5 所示。从图 5 可以分析出卷积网络对边缘和车牌颜色等敏感。进一步实验得到经过每层卷积网络产生的特征图,如图 6 所示。图 6 显示了两幅原

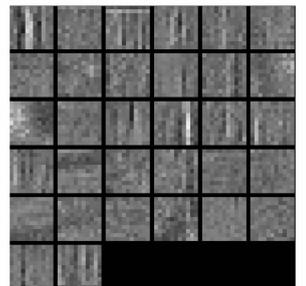


图 5 卷积网络权值图像
Fig. 5 Images of convolutional network weights

始图像经过卷积后生成的特征图,其中右上为第5层卷积后的特征图,第2行两幅图像为模型输出的特征图,分别为车牌左上角点和右下角点的置信图。每层卷积的特征图高亮部分表明了卷积网络敏感的区域,从图中每层产生的特征图可知,本网络模型对于车牌区域响应强烈,适用于解决车牌定位的问题,也验证了训练模型的有效性。

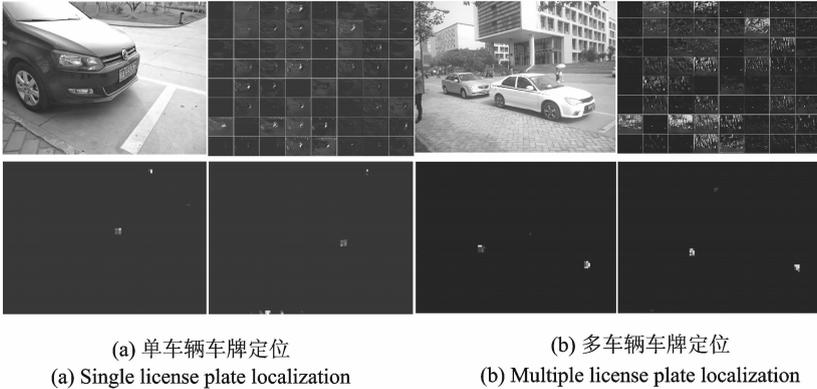


图 6 卷积神经网络特征图
Fig. 6 Convolutional neural network feature map

3.2.2 数据集测试与分析

测试数据由标准交通卡口数据和复杂环境数据构成。标准交通卡口数据用于检测在非复杂环境下本文方法的有效性,并和其他车牌定位算法进行比较和分析。复杂环境为低亮度、低分辨率和车辆倾斜严重的环境数据,由于传统方法无法在此环境中进行有效定位,故在复杂环境并没有和传统方法进行对比实验。经测试,本文算法不仅能够定位出蓝底白字和黄底黑字车牌,还能够同时对颜色失真车牌、多辆车车牌和警车车牌等进行准确定位,部分实验结果如图 7 所示。

本文同时进行了基于灰度变化特征算法和基于边缘检测算法对标准交通卡口图像数据集与本文算法的对比实验,结果如表 1 所示,其中测试图像分辨率为 900×675 。从实验结果分析,在标准交通卡口数据集上本文算法能达到和传统算法的定位率,并且在误定位车牌数量上优于传统方法。主要原因为本文车牌定位需要两幅角点特征图同时存在对应角点才能进行车牌定位,定位条件相对限制较多,且基于大量数据进行训练的结果,故误定位数量较低。而误定位部分主要为与车牌相似的标志牌,如车窗上的蓝色信息牌、停车库告示以及车身黄色广告等。

对比实验的传统方法误定位主要为车灯部分,未定位车牌主要为标准卡口中车牌过曝、模糊和车牌颜色失真的环境,而本文提出的算法并不依赖这 3 种环境,只和车牌区域信息相关,故在定位方面具备更高的鲁棒性。本文算法主要是解决复杂环境下车牌定位问题,故专门对复杂环境数据进行了测试,部分测试结果如图 8 所示。经测试,复杂环境下数据共有 1 540 幅,本文算法未定位车牌 68 幅,误定位车牌图像 8 幅,车牌定位率为 0.956。图 8(a)中左上为黑夜低亮度场景,右上为车牌污损场景,左下为黄牌过曝场景以及右下为图像低分辨率情况。其中第 3 幅卡车右上角标识部分被误定位车牌,主要为此区域和蓝底车牌类似。图 8(b)中为车辆倾斜严重场景数据测试结果。第 4 幅图中右上角车标识部分



图 7 利用卷积神经网络进行车牌定位效果
Fig. 7 License plate localization results via convolutional neural network

被误定位为蓝底车牌,因此区域和车牌区域相似。实验结果表明本文提出的算法不仅能够在标准交通卡口数据上进行有效定位,在光照条件低下、车辆倾斜严重等一些复杂环境下,同样具有较高的鲁棒性。

表 1 3种定位算法实验结果

Tab. 1 Results of three localization algorithm

算法	未定位车牌数	误定位车牌数	训练时间/min	测试时间/(ms/幅)	车牌定位率
本文算法	16	4	59	53	0.989 4
基于灰度变化特征算法	101	222	0	28	0.933 0
基于边缘检测算法	45	109	0	44	0.970 1



(a) 黑夜、过曝和低分辨率环境

(a) Scene of night, over-exposure and low resolution

(b) 车辆倾斜环境

(b) Inclination scene of vehicle

图 8 复杂环境实验效果

Fig. 8 Experimental results of complex scene

4 结束语

在复杂环境下,特别是低分辨率、车辆倾斜严重和低亮度环境,传统车牌识别方法无法进行准确定位。本文基于深度卷积神经网络方法对车牌左上角点和右下角点进行置信图回归,并根据置信图对车牌进行定位。在训练模型过程中为了加快模型收敛速度提出了相应的优化和调整策略。实验证明,在复杂环境下本文方法优于其他车牌定位方法,具有更高的定位准确率和鲁棒性。但场景中和车牌相似度较高的物体,如路牌、标志等,本文算法易误定位,故下一步工作任务将本方法应用于车辆以及车其他部位定位,结合车辆相关信息进一步降低车牌误定位率,并改善本模型解决定位问题的鲁棒性。

参考文献:

- [1] 邓彩霞,王贵彬,杨鑫蕊.改进的抗噪形态学边缘检测算法[J].数据采集与处理,2013,28(6):739-745.
Deng Caixia, Wang Guibin, Yang Xinrui. Improved algorithm of morphology in edge detection to resist the noise[J]. Journal of Data Acquisition and Processing, 2013,28(6):739-745.
- [2] 李国东,王雪,赵国敏.基于五阶 CNN 的图像边缘检测算法研究[J].安徽大学学报:自然科学版,2015,39(3):15-21.
Li Guodong, Wang Xue, Zhao Guomin. The algorithm for detection edge of image based on the 5 dimension CNN[J]. Journal of Anhui University: Natural Science Edition, 2015,39(3):15-21.
- [3] Yao Z, Yi W. License plate detection based on multistage information fusion[J]. Information Fusion, 2014,18:78-85.
- [4] 范春梅.车牌定位技术介绍与分析[J].信息技术,2013,37(11):167-168.
Fan Chunmei. Introduction and analyzing of the license plate pocation technology[J]. Information Technology, 2013,37(11):167-168.

- [5] Sermanet P, Eigen D, Zhang X, et al. Overfeat: Integrated recognition, localization and detection using convolutional networks[J]. arXiv Preprint arXiv:1312.6229,2013.
- [6] Oquab M, Bottou L, Laptev I, et al. Learning and transferring mid-level image representations using convolutional neural networks[C]//Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. Columbus: IEEE, 2014: 1717-1724.
- [7] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[J]. ePrint arXiv:1409.4842, 2014:1-9.
- [8] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems. Lake Tahoe: NIPS,2012:1097-1105.
- [9] 杨阳,张文生. 基于深度学习的图像自动标注算法[J]. 数据采集与处理,2015,30(1):88-98.
Yang Yang, Zhang Wensheng. Image auto-annotation based on deep learning[J]. Journal of Data Acquisition and Processing, 2015,30(1):88-98.
- [10] 王文豪,高尚兵,周静波,等. 图像中椒盐噪声去除算法研究[J]. 数据采集与处理,2015,30(5):1091-1098.
Wang Wenhao, Gao Shangbing, Zhou Jingbo, et al. Research on denoising algorithm for salt and pepper noise[J]. Journal of Data Acquisition and Processing, 2015,30(5):1091-1098.
- [11] 顾占冰,陶海涛,潘燕,等. 快速公交专用车道上非公交车辆抓拍系统[J]. 数据采集与处理,2012,27(6):696-702.
Gu Zhanbing, Tao Haitao, Pan Yan, et al. Image snapshot system of non-bus on rapid bus lane[J]. Journal of Data Acquisition and Processing, 2012,27(6):696-702.
- [12] 孙冬,高清维,卢一相,等. GPU加速的图像实时分形编码[J]. 安徽大学学报:自然科学版,2014,38(3):50-55.
Sun Dong, Gao Qingwei, Lu Yixiang. Real-time image fractal encoding based on GPU[J]. Journal of Anhui University, 2014,38(3):50-55.
- [13] University of California, Berkeley Caffe. Deep learning framework[EB/OL]. <http://caffe.berkeleyvision.org/.fr>, 2015-10-18.
- [14] 李琼,饶俊慧,陈多瑜. 基于边缘检测及灰度跳变的车牌定位算法研究[J]. 玉林师范学院学报,2013,34(5):127-132.
Li Qiong, Rao Junhui, Chen Duoyu. Research of vehicle license plate location algorithm based on edge detection and gray level jump[J]. Journal of Yulin Normal University, 2013,34(5):127-132.

作者简介:



罗斌(1963-),男,教授,研究方向:模式识别与应用、数字图像处理与应用, E-mail: luobin@ahu.edu.cn。



郜伟(1990-),男,硕士研究生,研究方向:图像处理与模式识别, E-mail: great-gaogao@qq.com。



汤进(1976-),男,教授,研究方向:计算机视觉、图像处理和模式识别, E-mail: 13244287@qq.com。



王文中(1978-),男,博士研究生,研究方向:计算机视觉、计算机图形学与虚拟现实, E-mail: 2238124604@qq.com。



李成龙(1988-),男,博士研究生,研究方向:图像处理与模式识别, E-mail: lel1314@foxmail.com。

