

DOI:10.19651/j.cnki.emt.2107071

基于 LSTM 算法的车牌识别系统方法研究

苏云涛 余红英 迟进梓

(中北大学 电气与控制工程学院 太原 030051)

摘要: 针对自然光照以及阴雨天气等复杂环境导致识别精度、响应时间降低问题,提出一种将 LSTM 理论和算法融于复杂环境下的车牌识别算法。通过将采集的车牌图像进行腐蚀算法、灰度化、二值化等预处理,增强车牌区域对比度,减少定位难度;其次,利用图像识别处理算法进行车牌定位、字符分割、字符识别等操作。将车牌分割后所得的字符归一化处理,统一的合适大小,作为长短时记忆网络(LSTM)的输入,对应汉字字符、数字和字母字符作为输出,通过训练所得 LSTM 网络实现车牌识别系统模型。经大量数据采集以及训练验证,与现有车牌识别系统相比,所提出的算法汉字字符识别准确率达 98.90%,字母以及数字字符识别准确率达 99.40%,单图识别速度达 2.65 ms。

关键词: 字符识别;车牌识别;图像处理;LSTM

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.99

Research on license plate recognition system based on LSTM algorithm

Su Yuntao Yu Hongying Chi Jinzi

(School of Electrical and Control Engineering, North University of China, Taiyuan 030051, China)

Abstract: This paper proposes a license plate recognition algorithm based on LSTM theory and algorithm in complex environment, aiming at the problem that the recognition accuracy and response time are reduced due to natural light and rainy weather. In order to enhance the contrast of license plate area and reduce the difficulty of location, the license plate image is preprocessed with corrosion algorithm, gray scale and binarization. Secondly, image recognition processing algorithm is used for license plate location, character segmentation, character recognition and other operations. The characters obtained after the license plate segmentation are normalized and the appropriate size is unified as the input of the long and short duration memory network (LSTM), and the corresponding Chinese characters, numbers and alphabetic characters are as the output. The model of license plate recognition system is realized through the TRAINED LSTM network. After a lot of data collection and training, compared with the existing license plate recognition system, the algorithm proposed has a recognition accuracy of 98.90% Chinese characters, 99.40% alphanumeric characters and a single image recognition speed of 2.65 ms.

Keywords: character recognition; license plate recognition; image processing; LSTM

0 引言

随着社会经济的发展,智能交通系统在现代城市管理中占有举足轻重的地位^[1]。而车牌识别算法作为建设智能交通系统的重要组成部分,在停车管理、ETC、车辆跟踪等城市管理方面有广泛应用。此外,人工智能技术也推动智能交通系统不断发展和延伸^[2]。

车牌识别的基本过程是:对利用摄像机等手段获得车辆图像或视频进行处理分析,然后使用数字图像处理、识别算法等获得车牌识别结果。但现实中,由于环境复杂,使得部分车牌在不同的光照和不同的天气条件下获得,由此,使得对车牌进行定位、分割等操作变得非常困难。目前常用

车牌字符识别技术有模板匹配字符识别法^[3]、神经网络字符识别法^[4]、特征统计匹配法^[5]等。综合国内外研究,陈玮等^[6]提出了改进型的基于欧拉数的模板匹配,原理上是一种改进型的模板匹配法,去除背景噪声并填充有效背景,提高识别质量,但该研究并未对汉字字符进行有效识别。赵艳芹等^[7]对卷积神经网络 LeNet-5 网络结构中的 C5 层和输出层进行了简单的扩充,达到车牌识别的任务要求。鲁扬^[8]采用改进的误差反向传播算法(error back propagation, BP)神经网络算法,虽然对数字和字母识别效果较好,但对汉字识别准确率并不理想。董峻妃等^[9]去除了传统 LeNet-5 网络结构中的 C5,但是该算法也同样没有实现车

牌汉字字符的识别。欧先锋等^[10]提出了一种基于 Caffe 深度学习框架算法,虽然较好地实现数字字符的识别,但限于样本数据集多样性,暂时未提及对某些字母字符及汉字字符的识别。Parekh 等^[11]提出一种新的识别算法,它以动态生成的车牌字符作为数据库模板,对字符进行识别。

针对文献[6-11]研究中的识别率低或识别字符种类有限等缺点,本文对目标车牌原始图像进行腐蚀算法、灰度化、二值化等预处理,然后针对车牌定位、字符分割等操作,利用图像识别处理算法完成,并将车牌分割后所得的字符作为长短时记忆^[12,16]网络(long short-term memory, LSTM)的输入,对应汉字字符、数字和字母字符作为输出。利用通过训练所得 LSTM 网络实现快速且准确的车牌识别,能实现在复杂自然条件下对车牌的准确、快速识别。

1 基于 LSTM 的车牌识别方法流程

基于 LSTM 的车牌识别方法流程结构如图 1 所示。其流程具体步骤如下:

1)图像预处理。针对获取原始车牌图像过程中,车牌质量受自然环境以及光照条件影响的情况,需要对车牌图像进行灰度化、边缘检测、二值化等图像预处理操作,增强车牌区域对比度,减少定位难度。

2)车牌定位。首先在 RGB 颜色空间下提取颜色特征,根据颜色特征信息完成粗定位。然后利用灰度跳变将车牌区域大体确定,最后在垂直方向和水平方向筛选区域,满足车牌长宽比的区域即为目标区域。

3)字符分割。采用垂直投影法对确定的目标区域进行二值化车牌图像像素分布情况分析,然后选择合适阈值,完成字符分割。

4)训练网络。将分割所得的车牌字符归一化处理,并统一的合适大小,作为训练样本,输入 LSTM 网络进行训练,得到训练后的 LSTM 网络。

5)字符识别。将获取的车牌图片通过步骤 1)~3)处理后,输入到步骤 4)中训练好的 LSTM 网络,得到识别结果。

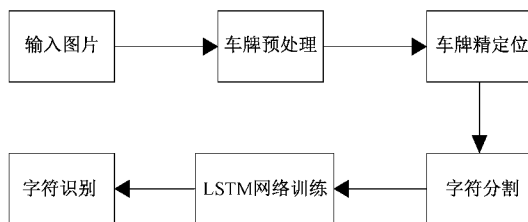


图 1 基于 LSTM 的车牌识别方法流程结构

2 LSTM 算法

2.1 算法描述

为了解决梯度爆炸和消失问题,通常采用 LSTM 块将门嵌入循环神经网络(recurrent neural network, RNN)的

隐藏神经元中。从这个意义上讲, LSTM 框架的一个优点是可以操作引入的门来存储或更新关键信息。此外, LSTM 模型能够在不消失梯度的情况下长时间保持信息。

LSTM 网络结构如图 2 所示。由图 2 可知,它包括具有循环连接的存储单元,并用 3 个乘法门:输入门 i 、遗忘门 f 和输出门 o 来控制信息的流通。 c 为细胞状态或称为记忆单元,控制信息的传递。 x_t 是当前时刻输入,而 h_{t-1} 是前一时刻隐藏状态。而作为 LSTM 关键的细胞状态类似于传送带,在整个链上运行,使信息保持不变的流过整个链。

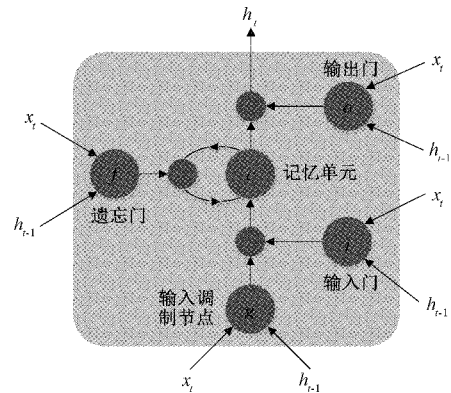


图 2 LSTM 网络结构

LSTM 中 3 个门的详细步骤可概括为:

1)遗忘门 f :决定什么样的信息将被剔除,即保留多少前一时刻的细胞状态 c_{t-1} 至当前的 c_t 。在每一个状态下输出(0,1)区间数值,其中,0 代表“完全遗忘”,1 与之相反。遗忘门计算如下:

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (1)$$

其中, f_t 是遗忘门 f 的激活; W_{fx} 为输入层到 f 的权重向量; W_{fh} 为隐藏层到 f 的权重向量; b_f 是 f 的偏移量; $\sigma(*)$ 为 sigmoid 激活函数。

2)输入门 i :它决定了当前 x_t 保留多少信息给当前状态 c_t 。这里包含两个部分。第一, sigmoid 层(输入门层)决定下一时刻更新什么值。然后, tanh 层创建一个新的加入到状态中的候选值向量。其计算如下:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (2)$$

其中, i_t 是输入门 i 的激活; W_{ix} 为输入层到 i 的权重向量; W_{ih} 为隐藏层到 i 的权重向量; b_i 是 i 的偏移量。

3)输出门 o :决定 c_t 传递多少至当前状态的输出 h_t 。这个输出将会基于细胞状态。首先,运行 sigmoid 层确定细胞状态的输出。接着,通过 tanh 层处理细胞状态,得到遗传信息:位于(0,1)区间数值。并将该数值乘以 sigmoid 门的输出,最终仅会输出目标部分。

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (3)$$

其中, o_t 是输出门 o 的激活; W_{ox} 为输入层到 o 的权重向量; W_{oh} 为隐藏层到 o 的权重向量; b_o 是输出门的偏移量。

输入门、遗忘门和输出门结果的计算大相径庭,均由当前输入序列 x_t 和前一时刻状态输出 h_{t-1} 乘以对应权重加上对应偏移量,最后经过 sigmoid 激活函数求解得到。而即时状态 \overline{C}_t 通过使用 tanh 激活函数(双曲正切激活函数)激活,计算公式如下:

$$\overline{C}_t = \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (4)$$

结合式(1)~(3)更新旧细胞状态,把旧状态与 f_t 相乘,遗忘过期信息。然后加上 $i_t * \overline{C}_t$ 得到新的候选值,即:

$$C_t = f_t * C_{t-1} + i_t * \overline{C}_t \quad (5)$$

则 LSTM 单元的输出 h_t 的计算式为:

$$h_t = o_t * \tanh(C_t) \quad (6)$$

2.2 网络训练

如图 3 所示为 LSTM 分类模型,其输入序列为归一化处理后的 32×16 像素大小的图片,相比于网络结构的输入图像为 32×32 大小,将输入字符图像改进为 32×16 大小更加符合车牌字符的长宽比,能够避免归一化为 32×32 大小时而产生的信息丢失的情况。同时,减少输入图像像素能够优化网络模型的训练参数,提高模型的运算速率。输入经过 LSTM 单元后,对应隐藏层会输出 h_0, \dots, h_t 。同时,运用 Attention 机制计算该时刻的隐藏层输出与字符特征向量的匹配得分占总体得分的比重,得到每个输入注意力概率分布值。设 h_t 为 t 时刻隐藏层输出状态, \overline{h}_t 为高一级特征向量。将 \overline{h}_t 进行随机初始化,作为参数输入 LSTM 训练网络,并逐步更新, $\alpha_t, i_t \in [0, 1]$ 的计算式为:

$$\alpha_i = \frac{\exp(\beta_i)}{\sum_{j=1}^n (\beta_j)} \quad (7)$$

式中: β_i 表示第 i 个隐藏层输出 h_i 在 \overline{h}_t 所占比重。输入在整体中的注意力与 β_i 正相关, β_i 计算公式为:

$$\beta_i = \mathbf{V}^T \tanh(\mathbf{W}h_i + \mathbf{U}\overline{h}_t + b) \quad (8)$$

式中: b 为偏移量, \tanh 为双曲正切激活函数, $\mathbf{V}, \mathbf{W}, \mathbf{U}$ 为权值矩阵。

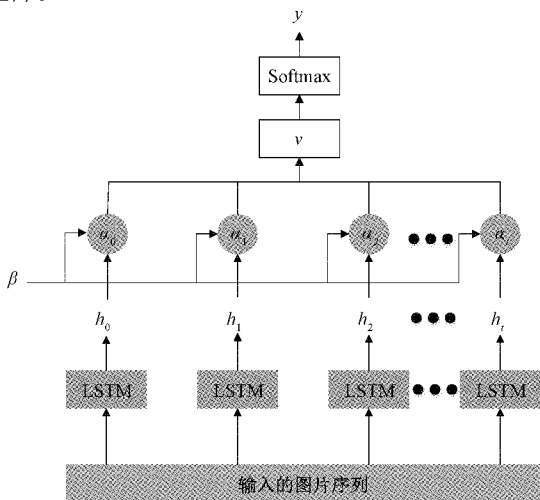


图 3 LSTM 分类模型

得到各字符图片注意力概率分布值后,再计算特征向量 ϵ , 公式如下:

$$\epsilon = \sum_{j=1}^l \alpha_j h_j \quad (9)$$

ϵ 包含字符特征信息,最后运用 Softmax 分类函数求得预测类别 y , 计算式为:

$$y = \text{softmax}(\mathbf{W}_v v + b_v) \quad (10)$$

式中: b_v 是偏置量; \mathbf{W} 为权值矩阵。

在本文 LSTM RNN 训练中引入了小批量梯度下降的优化算法,该算法对 n 个小批量样本进行梯度更新,从而充分利用传统的批梯度下降和随机梯度下降(SGD)。利用高度优化的矩阵优化算法同时计算 n 个样本的梯度,具有很高的效率,特别是对于深度学习更是如此。计算减少梯度更新过程中的变化,使得网络更稳定的收敛。模型使用的损失函数为交叉熵,计算式为:

$$H_{y'}(y) = - \sum y'_i \log y_i \quad (11)$$

其中, y'_i 是实际类别标签值, y_i 是预测类别标签值。

3 实验

3.1 实验数据

实验环境为 CPU 为 Intel i5-4570,主频 3.2 GHz、四核,4 GB 内存;64 位 Microsoft Windows 10 操作系统, MATLAB R2014a 编程语言。

我国的标准汽车牌照由汉字、数字、阿拉伯字母共同组成,共计 65 个不同的字符;同时由于我国车牌的特殊性,复杂环境下给车牌识别增加了难度。

MATLAB 是国际上公认的优秀数值计算以及仿真分析软件,研究使用该软件使得一些原本的繁杂计算变得简单,研究中运用 MATLAB 实现了车牌识别模块。

由于 LSTM 网络的训练要以大量带标签的训练集为基础,为了识别结果的公正性,使用的数据集不能过于单一,需要全国各地车辆图片。由于使用的数据集需要全国各地车辆图片,难以全部进行拍照采集,因此缺乏标准的数据集训练本研究提出的模型;本文将训练数据分为拍摄和收集两部分,且数据集包含环境复杂状态下获得的车牌图片,例如:不同自然光照、阴雨天气等各种复杂环境下的车牌图片。第 1 部分是通过手机拍摄、无人机拍摄等获得 200 张高清车牌图像;第 2 部分是通过百度搜索、谷歌搜索各省车牌图片下载 700 张全国各地车牌图像。为了增加数据集样本的多样性,将采集到的车辆牌照图像进行了拉伸、缩放、切斜等处理,然后所有车牌图像经过第 1 小节步骤 1)~3)后,由此得到 LSTM 网络训练集 10 000 多张图片,其中 6 000 多张图片为数字和字母字符,4 000 多张图片为汉字字符。作为 LSTM 网络的训练样本,每个字符图像均为相同大小。如图 4 所示,实验中用到的部分车牌样本。

如图 5 所示,显示了 LSTM 算法的车牌识别同 LeNet-5 车牌识别相比,训练次数与识别准确率的关系,从图中可以



图 4 部分车牌样本

看出,识别准确率与训练次数成正相关,当两者识别准确率不再升高后,LSTM 算法的识别准确率更高。

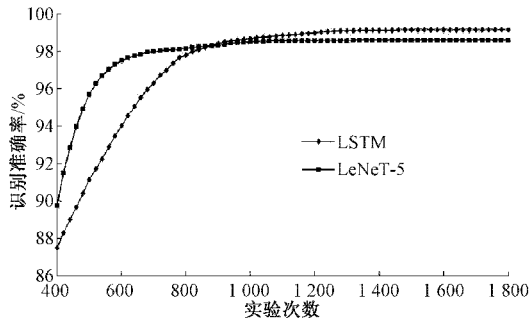


图 5 训练次数与识别率关系

3.2 实验结果

为证明基于 LSTM 算法车牌识别系统的优越性,本文从汉字识别准确率,字母、数字识别准确率以及车牌识别时间 3 个方面去评估各算法的识别效果。本文的测试集包括 90 张来自全国各地的车辆牌照,经过分割后得到 630 张字符图片,包括 90 张汉字字符以及 540 张字母、数字字符。实验结果对比^[17-18]如表 1 所示。

表 1 实验结果比较

算法	汉字识别 准确率/%	字母、数字 识别准确率/%	用时/ ms
文献[6]	—	96.67	5.59
文献[7]	98.68	98.68	3.79
文献[17]	98.53	98.85	3.75
文献[18]	—	94.19	1.25
本文	98.90	99.40	2.65

文献[6]使用改进型基于欧拉数的模板匹配法进行车牌识别。文献[7]在前文已提及,文献[17]使用卷积神经网络算法进行车牌识别。文献[18]采用 LSTM 算法进行车牌识别。由表 1 可知,本文中的车牌识别在汉字识别准确率、字母、数字识别准确率相比实验中其他结果均有提高;同时用时相较而言较短。文献[6]只完成了对字母、数字的

识别。文献[7]和[17]同时对汉字字符以及字母数字字符进行识别。文献[18]对外国车牌进行识别,与我国车牌相比,外国车牌并未涉及汉字字符的识别。

根据评估标准,由表 1 可知,本文所采用的基于 LSTM 网络的车牌识别算法汉字字符识别准确率为 98.90%,字母以及数字字符识别准确率为 99.40%,对单张图片的识别速度为 2.65 ms。相较于文献[6]的实验结果而言,本文算法不光实现了字母以及数字字符识别,还实现了汉字字符的识别,同时识别速度有所提高;相较于文献[7]和[17],具体表现在汉字识别最低升 0.22%,字母、数字识别准确率最低提升 0.55%,算法运行效率最低提升 29.33%。相较于文献[18],本文算法运行效率虽然略微下降,但实现了汉字字符识别,且字母、数字的识别准确率提升较大,达到了 5.21%。通过上述对比实验,体现了本文提出的方法具有更高的识别精度,以及更优秀的算法运行效率。

4 结 论

在现代智能交通管理系统中,车牌识别技术是其中的核心部分。本文论述了一种基于长时记忆网络(LSTM)算法的车牌识别系统的方法研究,对 LSTM 算法着重进行了分析和论述,并通过实验论证了研究的正确性和可靠性。

本文为提高车牌动态识别的准确率以及加快车牌识别的速度,对采集的车牌图像进行腐蚀算法、灰度化、二值化等预处理,然后利用图像识别处理算法进行车牌定位、字符分割、字符识别等操作。将车牌分割后所得的字符作为长时记忆网络(LSTM)的输入,对应汉字字符、数字和字母字符作为输出。通过训练所得 LSTM 网络实现快速且准确的车牌识别系统,本系统在分析验证后识别准确率高,达到 99.15%,算法运行效率高,达到了 2.65 ms。

由于本人学识、时间以及自身技能的受限,就目前来说,对车牌识别算法的研究还并不完善,仍在以后需要对很多方面进行改进;同时本文算法还存在有一些不足,需要进行图像处理分割出字符作为输入,而不是以整张车牌图像作为网络输入,还需要在之后研究中进一步优化算法,提高车牌识别效率。

参考文献

- [1] 赵伟,张南楠. 深度学习在复杂环境下车牌定位算法中的应用[J]. 现代电子技术,2019,42(17):38-42.
- [2] 张荣梅,张琦,陈彬. 基于改进 LeNet-5 的车牌识别算法[J]. 科学技术与工程,2020,20(12):4775-4779.
- [3] PURANIC A, DEEPAK K, UMADEVI V. Vehicle number plate recognition system: A literature review and implementation using template matching [J]. International Journal of Computer Applications, 2016, 134(1):12-16.
- [4] GUAN X Z, ZHANG L. License plate recognition based on improved BP neural network[J]. Techniques

- of Automation & Applications, 2015, 34(7):66-68.
- [5] SANJANA S, SHRIYA V R, VAISHNAVI G, et al. A review on various methodologies used for vehicle classification, helmet detection and number plate recognition[J]. Evolutionary Intelligence, 2020, DOI: 10.1007/s12065-020-00493-7.
- [6] 陈玮,曹志广,李剑平.改进的模板匹配方法在车牌识别中的应用[J].计算机工程与设计,2013,13(5):1808-1811.
- [7] 赵艳芹,童朝娣,张恒.基于 LeNet-5 卷积神经网络的车牌字符识别[J].黑龙江科技大学学报,2019,29(3):382-386.
- [8] 鲁扬.基于 BP 神经网络的车牌识别算法研究[D].大庆:东北石油大学,2018.
- [9] 董峻妃,郑伯川,杨泽静.基于卷积神经网络的车牌字符识别[J].计算机应用,2017,37(7):2014-2018.
- [10] 欧先锋,向灿群,郭龙源,等.基于 Caffe 深度学习框架的车牌数字字符识别算法研究[J].四川大学学报(自然科学版),2017,54(5):971-977.
- [11] PAREKH R, CHAKRABORTY S. An improved template matching algorithm for car license plate recognition [J]. International Journal of Computer Applications, 2015, 118(25):16-22.
- [12] 王丽君,刘彦戎,王丽静.基于卷积长短时深度神经网络行为识别方法[J].电子测量与仪器学报,2020,34(9):160-166.
- [13] SHU T X. Sensory data prediction using spatio temporal correlation and LSTM recurrent neural network [J]. Instrumentation, 2019, 6(3):10-17.
- [14] WIJESINGHE S. Time series forecasting: Analysis of LSTM neural networks to predict exchange rates of currencies[J]. Instrumentation, 2020, 7(4):25-39.
- [15] 陈思佳,罗志增.基于长短时记忆和卷积神经网络的手势肌电识别研究[J].仪器仪表学报,2021,42(2):162-170.
- [16] 秦敏莹,肖秦琨.基于长短时记忆网络的多媒体教学手势识别研究[J].国外电子测量技术,2019,38(6):80-85.
- [17] 刘华春.卷积神经网络在车牌识别中的应用研究[J].计算机技术与发展,2019,29(4):128-132.
- [18] HUI L, PENG W, MINGYU Y, et al. Reading car license plates using deep neural networks[J]. Image and Vision Computing, 2018, 72:14-23.

作者简介

苏云涛(通信作者),硕士研究生,主要研究方向为计算机控制。

E-mail:870750844@qq.com

余红英,教授,主要研究方向为电气传动与控制、系统监测与控制。

E-mail:yuhongying@nuc.edu.cn

迟进梓,硕士研究生,主要研究方向为计算机控制。

E-mail:1073566427@qq.com