

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

计算机视觉课程报告



题目：基于传统方法和深度学习的车牌检测识别算法

评分：_____

学生姓名： 黄榕基

学生学号： 519030910026

专 业： IEEE 试点班(信息安全)

学院(系)： 电子信息与电气工程学院

目录

1 绪论	1
1.1 实验要求分析	1
1.2 实验内容	1
2 基于传统方法的车牌识别	2
2.1 整体流程	2
2.2 算法原理	2
2.2.1 基于车牌颜色特征的图片二值化	2
2.2.2 基于轮廓识别和车牌形态学特征的车牌定位	3
2.2.3 基于基本图像变换的车牌透视变换	4
2.2.4 基于字符分割和模板匹配的车牌文字识别	4
3 基于深度学习的车牌识别	6
3.1 整体流程	6
3.2 算法原理	6
3.2.1 基于 U-net 图像分割网络的车牌定位	6
3.2.2 基于基本图像变换的车牌透视变换	7
3.2.3 基于 CNN 卷积神经网络的车牌文字识别	8
4 实验测试结果与分析	8
4.1 基于传统方法的车牌识别测试结果与分析	8
4.2 基于深度学习的车牌识别测试结果与分析	9
4.3 算法评价指标	10
5 总结	10
参考文献	11
谢辞	11

1 绪论

车牌检测识别是计算机视觉的一项经典应用，其算法可分为传统方法和深度学习方法。其中传统方法是基于形态学和色彩等特征对车牌进行检测与识别，其鲁棒性较差；而深度学习方法需要大量的数据集和计算资源进行训练才能达到较好的效果，随着大数据和深度学习技术的发展，该技术已日趋成熟，常见的有百度飞桨提供的车牌识别模型。

1.1 实验要求分析

要识别出一张图片中所包含的车牌上的文字，其难点主要为以下三点：

1. 如何准确定位图片中车牌的位置；
2. 如何对定位出的车牌进行透视矫正；
3. 如何准确识别透视矫正后的车牌上的文字。

第 1 个难点可根据车牌的形态学特征和颜色特征来对图片进行分割判断得到车牌图片，也可采用 YOLO、SSD、MaskRCNN 等目标检测算法进行检测，或是 Unet 图像分割算法来分割出车牌图片。第 2 个难点可根据分割出的车牌照片得知车牌的四个顶点和倾斜角度，因此可通过 OpenCV 提供的透视变换函数进行矫正。第 3 个难点可将车牌中每个字符分割开分别通过模板匹配来识别，也可采用 CNN 卷积神经网络训练识别。

对于给定的测试图片，难度被划分为 easy、medium 和 difficult 三个等级。其中 easy 等级是已经分割出来且矫正后的车牌，故只需要考虑第 3 个难点；medium 等级是从正面拍摄的车辆照片，故只需要考虑第 1 个和第 3 个难点；而 difficult 等级则是从斜方向拍摄的车辆照片，故需要考虑全部三个难点。

1.2 实验内容

我采用了传统方法和深度学习两个方法分别完成了车牌识别的任务。

对于传统方法，我基于形态学和色彩特征挑选出长宽比合适且颜色在范围内的矩形将车牌从图片中分割出来，再用准备好的模板通过模板匹配识别文字。

对于深度学习方法，我从 CCPD 数据集^[1] 中选取了一部分照片作为训练集，选择了 U-net 图像分割算法^[2] 来训练模型分割出图片中的车牌区域，再通过 CNN 卷积神经网络训练模型^[3] 来识别车牌中的文字。

2 基于传统方法的车牌识别

由于识别 difficult 等级的图片包含了该实验完整的流程和所有的难点，故将以 difficult 等级的图片 3-1.jpg 为例进行整体流程和实验原理的具体阐述。

2.1 整体流程

使用传统方法进行车牌识别的整体流程如下：

1. 将原始图片 BGR 转换为 HSV，根据是否符合车牌 HSV 色彩特征将图片二值化。
2. 对二值化后的图片进行形态学上的开操作和闭操作，去除小的噪点并将车牌化为一个整体。
3. 再将开闭操作后的二值图中的所有轮廓找出来，用最小的矩形将这些轮廓分别框住，选择出符合车牌长宽比的矩形，则该矩形框选的部分则包含车牌，将其从原图分割出来。
4. 对分割出来的车牌进行透视变换得到正视图的车牌图片。
5. 将车牌图片转化为灰度图，进行（自适应）阈值二值化。
6. 同样对二值化后的车牌图片进行形态学上的开操作和膨胀操作，去除小的噪点并将车牌上的每个字符分别化为一个整体。
7. 检测出膨胀后的每个字符的轮廓并用最小的矩形分别框住，选择出符合字符长宽比的矩形，根据这些矩形对二值化的车牌图片进行分割得到每个字符。
8. 将分割好的字符与准备好的模板进行模板匹配，选择匹配得分最高的模板对应的文字将车牌上的字符逐一文字化从而得到车牌识别结果。

2.2 算法原理

2.2.1 基于车牌颜色特征的图片二值化

将图片从 BGR 转化为 HSV 色彩模式，其中 H 为色调、S 为饱和度、V 为明度，是一种更加直观的颜色模型，在进行车牌颜色特征范围的设计时更加便于调整。

对于蓝色车牌的蓝色主颜色设计的 HSV 色彩范围为 $[100, 115, 115] \sim [124, 255, 255]$ ；绿色车牌的绿色主颜色所在的 HSV 色彩范围则为 $[35, 10, 160] \sim [70, 100, 200]$ 。再通过 OpenCV 提供的 inRange 函数来通过 HSV 色彩范围将图片二值化，在色彩范围内的像素点值化为 255，其余的值为 0，二值化前后图片如图 1 所示。

通过这种方式对图片二值化，可得到符合车牌颜色特征的区域，可以排除很大一部分的干扰区域。

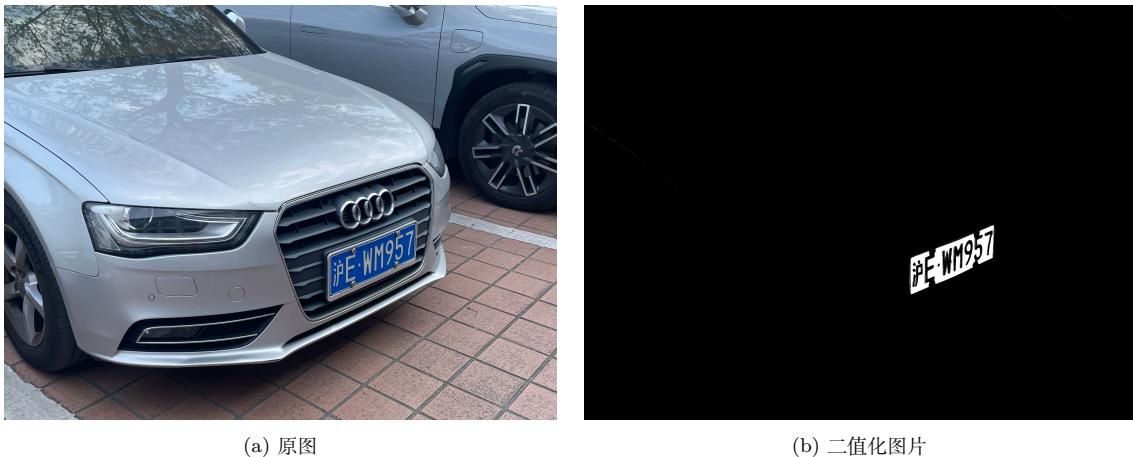


图 1: 基于 HSV 色彩范围的图片二值化

2.2.2 基于轮廓识别和车牌形态学特征的车牌定位

对已经二值化的图片先进行形态学开操作和闭操作，去除掉一些小的噪点，并将车牌部分化为一个整体，如图 2 所示，便于边缘检测得到整个车牌的轮廓。

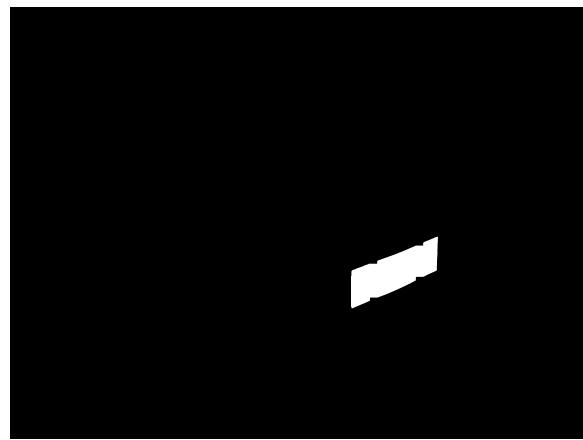


图 2: 开闭操作后的二值图

再对开闭操作后的二值图进行轮廓检测，用 OpenCV 提供的 `findContours` 函数找到二值图中所有的轮廓，再通过 OpenCV 的 `boundingRect` 函数以及 `minAreaRect` 函数将这些轮廓逐一用最小的正方形和斜矩形框住，同时得到车牌在图片中倾斜的角度。根据车牌的长宽比判断矩形是否符合要求，符合要求可将包含车牌的区域从原图和二值图中裁切出来，如图 3 所示。



图 3: 裁切出的车牌区域

这样就成功将包含车牌的区域裁切出来了，接下来需要通过透视变换得到正视图的车牌图片，方便最后的文字识别。

2.2.3 基于基本图像变换的车牌透视变换

根据裁切出来的二值图可得知车牌在 x 轴上的最小值 x_1 和最大值 x_2 ，差值 $dx = x_2 - x_1$ ，同时可得到其在 y 轴上的最小值 y_1 和最大值 y_2 ，但差值 dy 并不等于 $y_2 - y_1$ 。因此可采用几何方法，通过二值图中近似的平行四边形面积除以高 dx 得到底边长 dy 。又根据得到的倾斜角度可判断车牌在图片中是向左侧倾斜还是右侧倾斜，不同的倾斜方向在车牌顶点的表示方法上有所不同。根据 dx 、 dy 和倾斜角度可表示出车牌的四个顶点坐标。

若车牌是向右倾斜，那么其四个顶点可近似表示为 $(x_1, y_2-dy), (x_1, y_2), (x_2, y_1+dy), (x_2, y_1)$ ；若车牌是向左倾斜，那么其四个顶点可近似表示为 $(x_1, y_1), (x_1, y_1 + dy), (x_2, y_2 - dy), (x_2, y_2)$ 。

根据车牌的四个顶点，通过 OpenCV 的透视变换函数 `getPerspectiveTransform` 函数和 `warpPerspective` 函数得到透视变换矩阵并对裁切出的车牌区域图进行透视变换，从而得到正视图的车牌图片，如图 4 所示。



图 4: 正视图的车牌图片

2.2.4 基于字符分割和模板匹配的车牌文字识别

将得到车牌图片灰度化，并进行（自适应）阈值二值化，如图 5 所示。此处需要注意由于蓝色车牌和绿色车牌由于文字颜色和车牌背景色均不相同，绿色车牌二值化后得到的结果文字与背景颜色相较蓝色车牌二值化后的结果正好相反。因此需要将绿色车牌二值化后的图片颜色翻转，保证文字部分是白色（255），而背景部分是黑色（0）。

为了增加传统方法的鲁棒性，我设计了一个函数 `isBlue` 来判断车牌是蓝色车牌还是绿色车牌。该函数将整个车牌图片的通道 0 (blue) 和通道 1 (green) 的值分别求和，若 blue 值大于 green 值则判断该车牌为蓝色车牌，反之则判断为绿色车牌。从而对于不同颜色的车牌进行不同的操作。



(a) 车牌灰度图

(b) 车牌二值图

图 5: 处理过程图 1

然后，对车牌二值图进行形态学开操作，去除小的噪点，而后进行膨胀操作，让每个字符膨胀为一个整体，如图 6 所示。此后再对每个字符的整体进行边缘检测并用矩形框住检测出的边缘，根据字符长宽比判断得到符合要求的矩形框，此处会忽略掉车牌中的“·”符号，会在模板匹配过程中自动添补。通过矩形框在开操作后的车牌二值图上裁切，将车牌上的字符逐一分割出来，如图 7 所示。



(a) 开操作后的车牌二值图

(b) 膨胀为整体的每个字符

图 6: 处理过程图 2

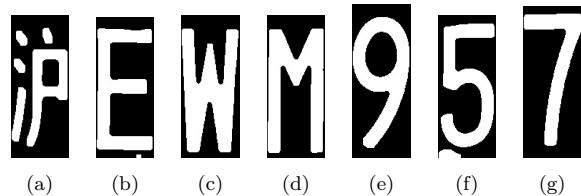


图 7: 分割开的字符

最后通过准备好的字符模板对分割开的字符图片逐一进行匹配，先将模板 `resize` 为和分割开的字符同样的宽度，再采用 OpenCV 的 `matchTemplate` 函数计算匹配相关系数，选择相关系数最大的模板对应的文字作为匹配的结果。每个字符图片匹配完成后输出最后的识别结果。

3 基于深度学习的车牌识别

采用深度学习的方法需要大量数据来进行训练，这里选取了 CCPD 数据集^[1] 中的部分图片作为数据集。我选择了 U-net 图像分割网络来训练模型分割出图片中的车牌区域，再通过 CNN 卷积神经网络训练模型来识别车牌中的文字，均通过 TensorFlow 的 keras 实现^[3]。

3.1 整体流程

基于深度学习的车牌识别整体流程^[3] 如下：

1. 使用 labelme 处理训练数据集，将照片中的车牌区域框选出来得到的二值图作为训练集标签，使用 U-net 图像分割网络进行训练，训练得到的模型可检测出图片中的车牌位置并生成对应的二值图。
2. 使用 OpenCV 对二值图进行边缘检测获得车牌区域，将车牌裁切出来并进行视角矫正。
3. 将得到的车牌正视图作为 CNN 卷积神经网络的训练数据集，将车牌上的文字经过一一对应为数字作为训练集标签，使用 CNN 进行训练，训练得到的模型可识别车牌上的文字。

3.2 算法原理

3.2.1 基于 U-net 图像分割网络的车牌定位

U-net 的网络架构^[2] 如图 8 所示，前半部分就是特征提取，后半部分是上采样。这种结构也被称为编码器-解码器结构，由于网络的整体结构是一个大些的英文字母 U，所以叫做 U-net。

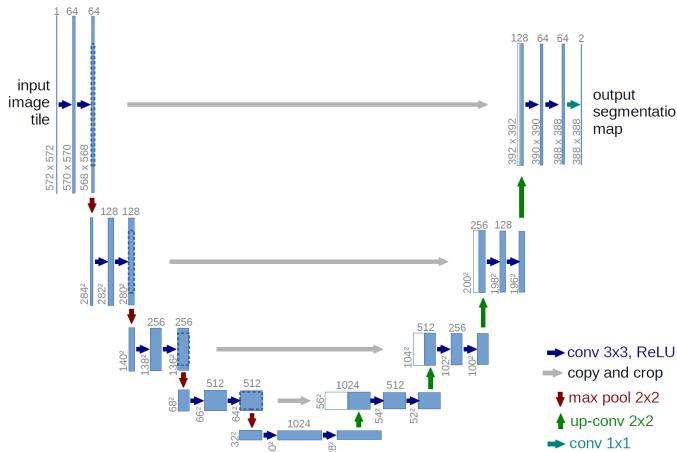


图 8: U-net 网络架构

其前半部分是由两个 3×3 的卷积层 (ReLU) 再加上一个 2×2 的 maxPooling 层组成一个下采样的模块；其后半部分是由一个上采样的卷积层 (去卷积层) 和特征拼接 concat 层以及两个 3×3 的卷积层 (ReLU) 反复构成。

由于能够获得的数据集数量有限，采用 U-net 这种小型简单的网络结构能更好地防止过拟合，能更多地关注图像表层的纹理信息，从而更好地检测出车牌的位置。

考虑到蓝色车牌与绿色车牌的特征并不相同，因此要进行两次训练，一次针对蓝色车牌进行训练，一次针对绿色车牌进行训练。通过 U-net 网络训练得到的模型进行预测，可以得到车牌位置值为 255、其他位置值为 0 的二值图，如图 9 所示。



图 9: U-net 预测车牌定位

3.2.2 基于基本图像变换的车牌透视变换

该部分算法原理与传统方法基本相同，但由于深度学习定位出来的车牌位置不一定精准，故进行透视变换后的结果也有可能有一定倾斜，如图 10 所示。



图 10: 车牌视角矫正

3.2.3 基于 CNN 卷积神经网络的车牌文字识别

将车牌上所有会出现的字符与数字一一对应，则一共有 65 种标签，而每个蓝色车牌有 7 个标签，每个绿色车牌有 8 个标签，且为了方便训练，我将绿色车牌也一并转化为蓝底白字的蓝色车牌，在此基础上进行训练。根据参考^[3] 将 CNN 卷积神经网络设计为

```
Input → Conv → ReLU → Pool  
→ [Conv → ReLU → Conv → ReLU → Pool] × 3  
→ FC → softmax
```

而由于训练数据和计算资源不足，我通过 CNN 训练出的模型效果并不好，预测结果准确率不高，这一点将在第 4 部分说明。

4 实验测试结果与分析

4.1 基于传统方法的车牌识别测试结果与分析

基于传统方法的车牌识别的测试结果如图 11 所示：

```
PS D:\大三下\Computer Vision\hw\project\code> d:; cd 'd:\大三下\Computer Vision\hw\project\code'; & 'C:\Users\30403\conda\envs\cv\python.exe' 'c:\Users\30403\.vscode\extensions\ms-python.python-2022.4.1\pythonFiles\lib\python\debugpy\launcher' '52159' '--' 'd:\大三下\Computer Vision\hw\project\code\traditional_method.py'

testing images/easy/1-1.jpg
result: 沪E·WM957
-----
testing images/easy/1-2.jpg
result: 沪A·F02976
-----
testing images/easy/1-3.jpg
result: 鲁N·BK268
-----
testing images/medium/2-1.jpg
result: 沪E·WM957
-----
testing images/medium/2-2.jpg
result: 豫B·20E68
-----
testing images/medium/2-3.jpg
result: 沪A·93S20
-----
testing images/difficult/3-1.jpg
result: 沪E·WM957
-----
testing images/difficult/3-2.jpg
result: 沪A·DE6598
-----
testing images/difficult/3-3.jpg
result: 皖S·J6M07
-----
```

图 11：基于传统方法的车牌识别测试结果

可以看见，基于传统方法的车牌识别测试正确率为 100%。实际上，模板的选取对于识别正确与否影响很大，且车牌透视变化不准确也会影响识别结果。

4.2 基于深度学习的车牌识别测试结果与分析

根据参考^[3]我设计了一个可视化框，展示车牌定位和识别的结果。基于深度学习的车牌识别的测试结果如图 12 所示。



图 12: 基于深度学习的车牌识别测试结果

可以看到，基于深度学习的车牌识别测试结果只有 medium 难度全对，其他难度均有错误。对于 easy 难度，应该是训练 U-net 模型时训练集中单个车牌的照片太少，导致对与单个车牌的判断有误，框选出来的部分就不再是完整的车牌，从而导致文字识别出现误差。而对于 difficult 难度，则是框选出来的部分有误差导致透视变换的结果并不理想，从而导致文字识别出现误差。

而且我在训练 CNN 模型时将绿色车牌一并转化为蓝底白字的蓝色车牌进行训练，可能是这一部分样本数量不足，导致训练出来的模型无法区分哪些是 8 个字的车牌那些是 7 个字的车牌，从而全部识别出来只有 7 个字。这些都是还有改进空间的点，但碍于训练数据和计算资源不足难以优化。

4.3 算法评价指标

如果只看完整的车牌识别正确率，传统方法的正确率到了 100%，而深度学习方法的正确率只有 44.44%，这样的评价指标很明显是不太合理的。

因此，可以参考目标检测中常用的 IoU 指标（Intersection over Union，实际结果与测试结果的交集比上实际结果与测试结果的并集），设计一种对于字符识别的 IoU 指标，IoU-CR (Intersection over Union for Character Recognition)，即正确识别的字符个数比上真实结果字符个数和错误识别或漏掉的字符个数的和。

那么可以计算出两个算法的 IoU-CR 准确率，如表 1 所示。

testing image	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3	average
traditional method	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
deep learning	0.75	0.60	0.56	1.00	1.00	1.00	1.00	0.45	0.75	0.79

表 1：基于传统方法和深度学习的车牌识别在测试集上的 IoU-CR 准确率

由于这两个算法都能够保证预测顺序的正确，也就是说只有错误识别和漏掉的识别需要考虑，那么通过 IoU-CR 指标可以有效地衡量算法预测的准确性。在训练集不足的情况下能够有接近 80% 的准确率是可以接受的，只有少量的字符被错误识别。

5 总结

对于车牌识别，本文从传统方法和深度学习两个方面进行了阐述。通过对车牌定位裁切、视角矫正、文字识别三个阶段逐步完成对图片中车牌的识别，并设计了 IoU-CR 指标来评价算法的性能，能够达到较高的 IoU-CR 准确率。

具体来看，本文完整分析了传统方法的细节，基于车牌的色彩和形态学特征通过二值化、轮廓检测、模板匹配等具体方法进行车牌识别，有着非常高的准确率，但鲁棒性较低；同时也分析了深度学习的方法，对于同样的问题可通过 U-net 图像分割网络和 CNN 卷积神经网络来解决，虽然准确率不及传统方法，但有着更高的鲁棒性。

总的来看，传统方法的局限性是无法避免的；若能有更多的训练集和充足的计算资源，通过深度学习得到的模型无论是准确率还是鲁棒性都将得到很大的提升，且能应用于日常大多数需要车牌识别的场景。

参考文献

- [1] Xu Z, Wei Y, Meng A, et al. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline[C/OL].European Conference on Computer Vision. Springer, Cham, 2018. <https://github.com/detectRecog/CCPD>.
- [2] O. Ronneberger and P. Fischer and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation[J/OL].Medical Image Computing and Computer-Assisted Intervention (MICCAI),LNCS. Springer,2015,9351:234–241.(available on arXiv:1505.04597 [cs.CV]).<http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a>.
- [3] duanshengliu. End-to-end-for-chinese-plate-recognition: 基于 u-net, cv2 以及 cnn 的中文车牌定位, 矫正和端到端识别软件 [OL]. 2020-06-14. <https://github.com/duanshengliu/End-to-end-for-chinese-plate-recognition>.

谢辞

感谢计算机视觉的授课老师赵旭老师讲授的计算机视觉知识与应用，给予了我很多深刻的启发，激起了我对于该领域的探索兴趣。感谢这门课的助教们的辛苦付出与认真指导。