

---

# 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方法的车牌识别算法

评分：\_\_\_\_\_

学生姓名：黄霄童

学生学号：520030910229

专    业：信息安全

学院(系)：电子信息与电气工程学院

---

# 目录

目录.....	0
第一章    绪论.....	1
1.1    车牌定位.....	1
1.2    车牌字符分割.....	1
1.3    车牌字符识别.....	2
1.4    论文的结构.....	2
1.5    实验要求分析.....	2
第二章    算法步骤.....	4
2.1    图像预处理.....	4
2.2    形态学操作（膨胀和腐蚀）.....	5
2.3    截取车牌并分辨车牌类型.....	6
2.4    去除车牌框和铆钉等干扰.....	6
2.5    基于先验知识和投影方法字符分割.....	7
2.6    CNN 网络识别图片.....	8
第三章    实验测试结果与分析.....	10
3.1    车牌识别测试结果.....	10
3.2    实时性分析.....	10
3.3    评价指标.....	11
第四章    结论.....	12
参考文献.....	13

## 第一章 绪论

随着城市的日益现代化,智能交通系统逐渐成为交通管理系统的重要发展方向,其中,车牌自动识别系统是智能交通系统的重要组成部分。当前主流的车牌自动识别算法通常分为三个部分:车牌定位、字符分割和字符识别。

### 1.1 车牌定位

车牌定位是从所获取图像中运用某些算法,通过图像处理找出车牌在图像中的位置。目前,最常用的办法主要有:基于颜色特征<sup>[1, [2]]</sup>、基于边缘特征和形态学变换<sup>[4, [5]]</sup>、基于颜色和边缘混合特征和基于神经网络的车牌定位<sup>[6, [7]]</sup>。

由于车牌通常具有固定的颜色,因此我们可以通过颜色特征来判断图像中某块区域是否可能是车牌,也可以根据颜色的比例识别出车牌的类型,中国的车牌总共有三种类型:蓝色、绿色和黄色,不同类型的车牌字符特征并不完全相似。由于 RGB 三原色容易受到光照影响,为了更好的提取色彩特征,很多论文会尝试将图像转换到 HSV 空间,在 HSV 空间上基于颜色特征定位车牌。但是,基于颜色特征的识别仍然容易受到拍摄环境条件的影响,因此实际应用中通常结合其他方法一起使用。

基于边缘特征的车牌定位根据字符特点、车牌的边缘等信息来对图像中的车牌进行定位。车牌区域由于字符的存在具有强烈的边缘,表现为从背景到车牌、车牌背景到车牌字符有较大边缘以及密度较大等边缘特征。此外,由于车牌通常在竖直方向的边缘特征更加丰富,为了更加准确的定位车牌,排除干扰,很多论文选用 Sobel 算子仅提取图像在竖直方向的边缘特征。在实际生活中,通常还会结合形态学操作增强车牌的边缘特征,方便更准确的定位车牌位置。在本篇论文提出的算法中,我们借鉴了这一思想,尝试使用 Sobel 算子结合形态学操作的方式定位车牌位置。

近些年,随着对机器学习和模式识别研究的深入,使得利用神经网络模型<sup>[8, [9]]</sup>和遗传算法等理论进行车牌定位成为当前研究的热点之一,神经网络有许多优势:具有良好的抗噪性,容错、自适应和自学习能力强等。常见的分割网络有 Unet, YOLO 等。但是,深度学习方法通常由于需要大量的数据集进行训练,或网络参数过于复杂耗时较长等原因,不能满足实时性要求或者不能适应某些固定的场景。

### 1.2 车牌字符分割

字符分割是从含有字符的车牌中分割出单独字符。实际拍摄的车牌存在图像光照变化、图像被阴影遮挡、车牌不清洁、字符断裂、单双层车牌以及边框等现象,使分割的难度大幅度增加。如本次大作业的图片存在车牌倾斜,车牌表面光照不均以及车牌铆钉干扰分割等问题。常见的字符分割方法有基于投影、边缘和轮廓提取等方法。

基于投影的方法通常是基于车牌在不同字符之间存在跳变属性,将车牌水平和垂直投影到两个坐标轴上,结合车牌每个字符宽度和长度相同,车牌字符固定比例间距等先验知识,根据直方图的峰谷特征分割字符。基于投影的方法操作简单,但是易受车牌定位准确性、车牌表面污损等因素的影响。可以通过如 hough 变换等方式,将定位的车牌矫正后再使用投影方法分割字符<sup>[10]</sup>。

基于边缘和轮廓提取的方法通常是利用字符连通域的特性,使用形态学膨胀和腐蚀操作提取轮廓特征,从而分割字符。该方法的缺点在于对于连通性较差的字符可能会不能正确分割,出现将一个字符分割成两个字符的情况。

### 1.3 车牌字符识别

车牌字符识别是将分割好的字符图片准确的识别出对应的字符,通常会结合首字符为一定范围内的汉字,第二个字符为英文字母等先验知识,辅助字符识别,提高识别的鲁棒性和准确性。常用的方法通常有基于模板匹配的方法、基于特征的识别和基于神经网络的识别方法。

基于模板匹配的方法即将分割的字符图片和样本库中的样本一一对比,选择相似度最高的字符作为匹配结果。该方法操作简单,但是性能会极大取决于样本库中字符的质量,同时,当样本库的字符较大时,字符的实时性会随之降低。

基于特征的识别通常是先求取字符特征,再设定分类器如 SVM 等识别字符。字符特征通常为字符像素的统计特征和结构特征<sup>[11]</sup>,字符的统计特征包括像素块数、字符的轮廓数、轮廓的形状等,结构特征则是根据如汉字的笔画不同,提取横、撇、竖、捺等特征的方式分辨字符。在实际应用中,倘若字符出现了模糊、旋转等问题,该方法可能会导致部分字符无法正确识别。

基于神经网络的识别<sup>[12]</sup>通常采用 BP 神经网络,在使用分割好并做好标注的人工数据集训练好后,直接将归一化后的字符图像输入到网络中得到分割结果。该方法严重依赖于参与训练的数据集的质量和泛化性,比如普通车牌的 6 和新能源车牌的 6 差别较大,倘若参与训练的数据集中不存在新能源车牌的 6,神经网络很大概率会将 6 误识别成更像的字符 G。此外,还存在一种更加直接的方式,即直接将分割好的车牌放入如 RNN 等神经网络中,直接识别车牌中的字符的方法,该方法可以跳过字符分割,但是需要良好的网络设计和海量的数据以期待该网络收敛,难度较大。

### 1.4 论文的结构

该论文第一章绪论部分主要总结了当前车牌识别领域在车牌定位、车牌字符分割和字符识别方面常见的方法,并分析了各个方法的优缺点和在实际应用中的应用思路;第二章主要叙述了我们提出的车牌识别算法的基本步骤;第三章实验部分展示了本算法的泛化性和准确率,证明了我们的算法具有一定的鲁棒性;第四章总结部分则总结了本算法的优缺点,并简述了可能的改进思路。

### 1.5 实验要求分析

按照给定测试集的难度等级,我们不难分析出车牌识别系统至少应该克服的难点有:

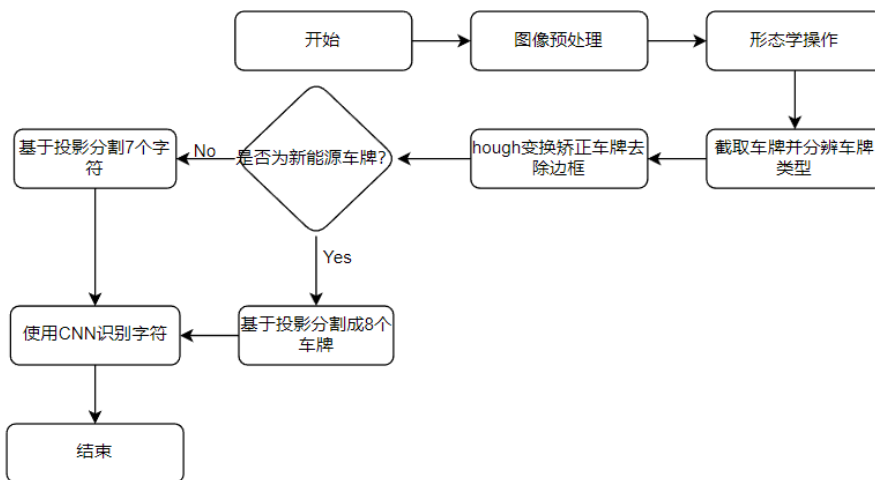
- (1) 如何精准的识别车牌上的字符,排除车牌如铆钉等其他区域的干扰;
- (2) 如何从一张照片中精准定位出车牌的位置;
- (3) 如何检测出车牌是否存在倾斜,并精准矫正;
- (4) 如何分辨蓝色车牌和绿色车牌的不同,并给出相对应的处理方式。

对于(1)我们可以考虑使用 CNN 神经网络来对分割好的字符分别识别,并利用好车牌上不同位置固定字符类型的先验知识;对于(2)我们可以考虑结合边缘特征和颜色特征定位出车牌的位置;对于(3)我们可以考虑使用 hough 变换拟合车牌的角度,并进行矫正处理;对于(4)我们可以考虑利用颜色特征分辨车牌类型,并对蓝色车牌主动分割 7 个字符,对于绿色车牌主动分割 8 个字符。

显然 difficult 数据集里的图片涵盖了 easy 和 medium 里所有的难点,因此下面演示算法执行过程我们均选用 difficult 数据集中的图片进行效果演示。

## 第二章 算法步骤

为了方便展示，我们默认展示图像 3-2.jpg 的处理结果。我们提出的算法流程图 2-1 如下：



图表 2 - 1 算法流程图

### 2.1 图像预处理

图像预处理主要分为四个步骤：

- (1) **Resize 操作**：由于形态学操作核大小的设置对于图像的尺寸较为敏感，因此我们首先对输入的图片进行 **resize** 操作，将图像的宽设置成 700，图像的长根据宽的缩放比例进行缩放。

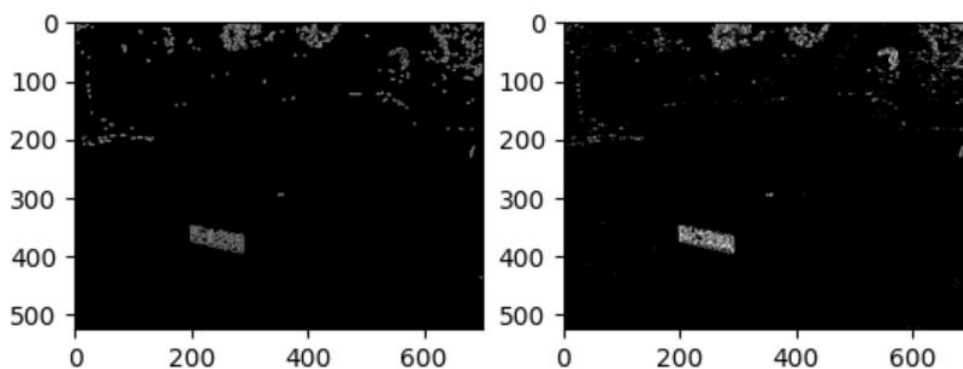


图表 2 - 2 绿色车牌掩膜结果

- (2) **获取 HSV 空间的掩膜**：HSV 模型中颜色的参数分别是：色调(H), 饱和度(S), 亮度(V)。由于 HSV 颜色模型相比于 RGB 颜色空间与人的视觉特性更为接近, 本文采取使用广泛的 HSV 颜色模型对车牌进行初步检测。对于图像中的每一个像素点  $x$ ，若  $x$  的值在  $[100, 115, 115] \sim [124, 255, 255]$ ，视为蓝色车牌可能存在的位置；若  $x$  的值在  $[35, 10, 160] \sim [70, 100, 200]$ ，则视为绿色

色车牌（新能源车牌）可能存在的位置。对于除了上述两个区域以外的图像区域，我们直接置为 0，以排除干扰。使用蓝绿色车牌的掩膜处理结果见图表 2-2。

- (3) 高斯模糊：在使用掩膜后，为了方便后续形态学操作，我们将图片从  $hsv$  空间转到灰度空间，并使用  $3 \times 3$  的核进行高斯模糊，去粗可能存在的噪点。
- (4) 边缘检测：正如绪论中所述，车牌垂直边缘较多，因此我们选择 Sobel 算子的水平方向提取出图像中的垂直边缘，可以在一定程度上过滤掉水平边缘的噪点，方便后续更好的识别出车牌位置，减少干扰。Sobel 水平方向算子处理结果和 Canny 边缘检测的结果对比见图表 2-3，可以发现，相比 Canny 算子，Sobel 算子很好的提取了车牌的边缘特征。



图表 2 - 3 Canny 算子(左)与 Sobel 水平方向算子(右)

## 2.2 形态学操作（膨胀和腐蚀）

膨胀是指将图像中的一部分区域  $A$  与核  $B$  进行卷积，即计算核  $B$  覆盖的区域的像素点的最大值，并把得到的最大值赋给参考点指定的像素。这样就会使得图像中的亮度区域变得更强。在试图寻找连通分支（即具有相似强度或颜色的像素点的块状互相分离的区域）时通常用膨胀操作。腐蚀是膨胀的反操作，即腐蚀操作是计算核  $B$  覆盖的区域的最小值，并且把该最小值赋值给参考点。腐蚀可以将图像上的斑点腐蚀掉，且能确保图像内的大块区域仍然存在。

在二值化边缘检测的结果后，我们先使用膨胀操作将车牌化为一个整体，再使用腐蚀操作去除图像中一些小的噪点。经过实验发现，我们首先膨胀 2 次，再腐蚀 3 次，最后膨胀 1 次可以较好的将所有测试图片中的车牌区域变成整体，并消除噪点，且该步骤具备鲁棒性。形态学操作结果见图表 2-4。



图表 2 - 4 形态学处理



## 2.3 截取车牌并分辨车牌类型

在形态学操作后，我们需要设计特征来过滤掉不是车牌的结果。在这里，我们设计了三个特征：

- (1) 截取的矩形区域长宽之比应该在某一个范围之间，根据先验知识，车牌外廓尺寸为  $440 \times 140(\text{mm}^2)$ ，因此宽长之比应该在 3.5 左右，由于车牌可能存在倾斜的情况，因此括住整个车牌的宽长比我们设置在  $[2.5, 5.5]$  之间；
- (2) 车牌区域的角点应该较多，因此我们要求矩形内部的角点面积应该超过 2500；
- (3) 我们定义矩形度 (rectangularity) 为角点面积除以该矩形框的面积，该特征说明了矩形框中角点面积的占比情况，显然，由于车牌位置角点数量应该较稠密，因此我们认为车牌位置角点面积在矩形框的占比应该超过某一个阈值，这里我们设置为 0.4；
- (4) 蓝色车牌的 Blue 通道之和除以 255，和矩形框的面积之比应该超过某一个阈值，同理，绿色车牌的 Green 通道之和除以 255，然后和矩形框的面积之比也应该超过某一个阈值，这里我们设为 0.45，这个特征描述了颜色在图像中的占比，超过该阈值说明这个矩形框中存在大部分颜色为蓝色或绿色的物体，我们认为其大概率是车牌。此外，若蓝色占比比例大于绿色，则应该是蓝色车牌；反之，我们可以判断为绿色车牌。

在经过上述四个特征筛选后，我们可以将包含车牌的矩形区域从图片中筛选出来。为了证明上述特征是合理且有效的，如图 2-5，在形态学结果处理后，我们可以一一验证每个矩形框内不同的特征。



图表 2 - 5 形态学处理后可能的车牌位置

## 2.4 去除车牌框和铆钉等干扰

在截取出车牌框后，我们首先将该框内的图像转为灰度图，并二值化。我们发现，绿色车牌的字符是黑色的，因此，对于绿色车牌，我们在二值化后还需要反转像素，使得字符区域总是白色。

但是，二值化后的图片仍然不能直接使用垂直投影将字符分割，因为诸如车牌上的铆钉，以及矩形框中其他具有和字符一样的颜色的区域，我们需要通过一些方式去除掉。



当车牌水平时，如图表 2-6 右上图片所示，我们不难观察到，干扰的白色区域一般存在两种类型：

- (1) 这一行连续的黑色段或者白色段总是较长，而字符区域，由于字符的形体特征，不可能存在白色段或是黑色段很长的情况。这里，我们认为连续的白色段超过有效长度（即一行中第一个白色点到最后一个白色点的距离）的  $1/3$ ，则显然不可能是字符段，将这一段全部置 0；
- (2) 这一行虽然连续的黑色段或是白色段不长，但是离散的白色点较多几乎占满整行，显然字符区域不可能出现白色区域远远超过黑色区域的情况。这里，我们认为白色点占据整个有效段的比例超过 0.8，则显然也不可能是字符段，同样置 0。

当车牌倾斜时，我们考虑将倾斜的车牌矫正为水平，再依据上述分析去除干扰区域。对于倾斜的车牌，我们考虑使用 hough 变换进行车牌矫正。我们考虑拟合出 hough 变换中所有和水平线夹角在  $\pm 30^\circ$  之内的直线，根据这些直线与水平线的夹角的均值判断车牌应该旋转的角度，并使用透视变换将图片转为水平。需要注意的是，在删除冗余水平线后，我们需要将图片旋转回来，因为车牌在水平旋转后，字符将变成倾斜的方向，这不方便我们后续利用垂直投影分割字符。而我们假设原图中，车牌上的字符无论车牌怎么倾斜，总是可以找到垂直分割线去分割字符，因为在图片中，如果摄像机是正的，垂直方向总是不会产生倾斜。总之，整体步骤为：旋转-去除冗余线-逆旋转回原图，如图 2-6，我们可以较好的删除冗余的白线。



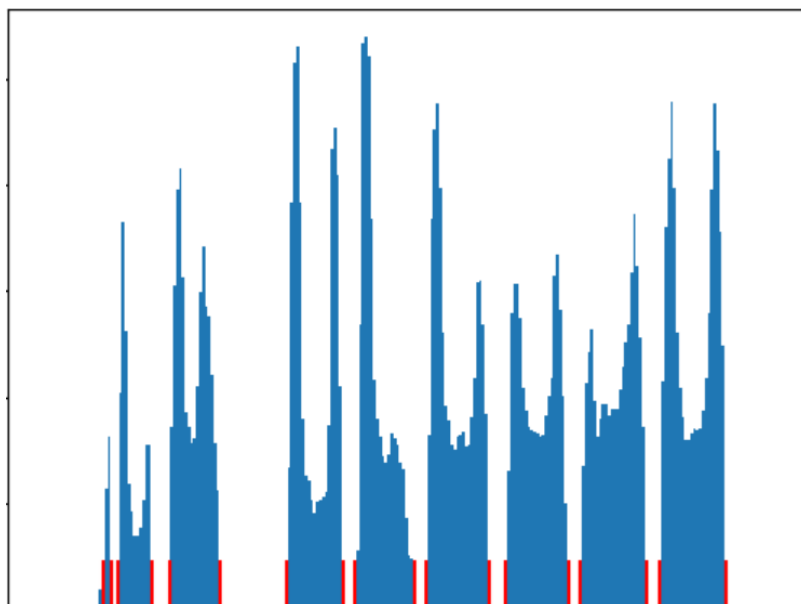
图表 2 - 6 左上为二值化的图像，右上为利用 hough 变换矫正车牌的图像，左下为删除冗余白线后的效果，右下为逆旋转的结果。

## 2.5 基于先验知识和投影方法字符分割

首先，我们利用垂直投影来分割字符的竖直方向，为了避免被车牌中间的原点干扰分割结果，我们根据车牌的先验知识，从左往右分割两个字符，再从右往左根据车牌的类型分割对应个数的字符：蓝色车牌分割 5 个，绿色车牌分割 6 个。

如图 2-7 垂直投影后的直方图所示，由于我们之前去除了冗余的白点，峰与峰之间的间隔较为明显，因此我们简单的设定直方图的均值/3 作为阈值，小于该阈值的位置视为波谷。对于距离很近的波谷，我们直接删除，以为一个字符的左右两个波谷不可能很

小；对于两个距离较近的波峰，我们认为是属于同一个字符，只是因为如“沪”字的偏旁分隔较开因此中间存在波谷。



图表 2 - 7 垂直投影直方图，红线为选定的可能的分割线

最后，我们假设前面冗余的白线基本被清除，我们对每一个字符进行水平投影，切除掉冗余的黑色区域，即可保证字符在分割图像的正中央，方便后续神经网络的识别。分割字符的结果如图 2-8 所示，我们的算法符合我们想法。



图表 2 - 8 字符分割结果

## 2.6 CNN 网络识别图片

我设计了一个简单的神经网络，可以根据输入的图片输出所有可能的字符的概率分

布，取最大的概率分布作为分类结果，并利用人工标注好的字符集（图片大小为 20x20）进行训练，在代码中称为 PNet，网络结构如图 2-9 所示。

需要注意的是，为了得到更加准确的分类结果，我们对于汉字和其他字符分开训练，得到一个可以识别汉字的神经网络和一个可以识别字母和数字的神经网络；对于车牌的第一个字符，放入可以识别汉字的神经网络中获得可能的分类结果，第二个字符由于一定是字母，因此对于第二个字符我们仅仅使用第二个神经网络中对于字母的分类结果，后面的字母则直接使用第二个神经网络训练，并得到相应的分类结果。

我们训练了 20 个 epoch，在测试集上，识别汉字的神经网络准确率为 95.3%[142/149]；识别字母和数字的神经网络准确率为 99.9%[660/664]。对于本次作业中分割好的字符，我们将字符 resize 成 20x20，归一化后作为输入，准确率达到 100%，说明该神经网络具有一定的泛化性。

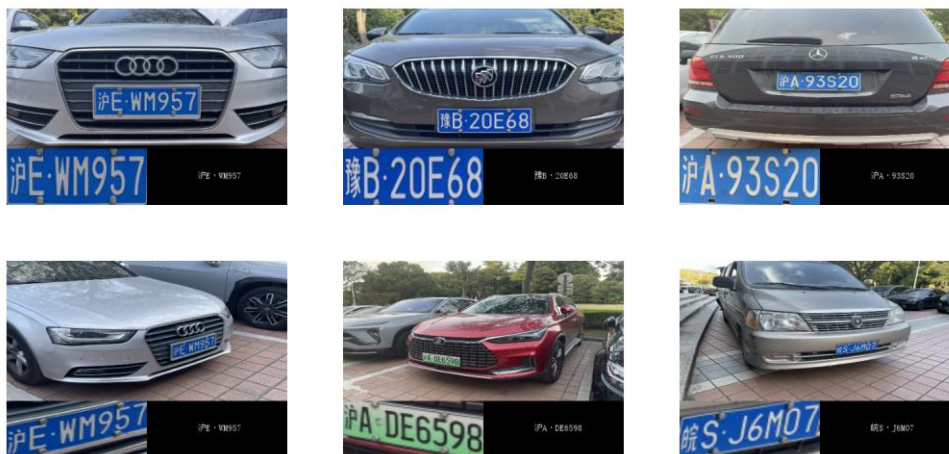
model	hyper-parameters
输入图像	20 x 20
conv	in channels=1, out channels=8, kernel size = 5, padding = 2
BatchNorm	num features = 8
maxPooling	kernel size=2, stride=2
conv	in channels=8, out channels=16, kernel size = 5, padding = 2
BatchNorm	num features = 8
maxPooling	kernel size=2, stride=2
Linear	in features=5516, out features=128
Dropout	p=0.5
Linear	in features=128, out features=64
Dropout	p=0.5
Linear	in features=64, out features=N_classes
Softmax	

图表 2 - 9 PNet 网络结构，在每层卷积和 fc 层后我们均选择 ReLU 作为激活函数

## 第三章 实验测试结果与分析

### 3.1 车牌识别测试结果

对于给定的测试图片集，easy 是等级是已经分离出来并且矫正好的图片，因此在 easy 数据集上我们跳过了前面车牌定位的步骤，而 medium 和 difficult 数据集则是完全遵循于我们提出的算法步骤执行，识别结果如图表 3-1 和 3-2 所示，准确率达到 100%。



图表 3 - 1 车牌识别结果可视化 (medium 和 difficult)

```
#### easy ####
img name: 1-1.jpg Licence plate: 沪E·WM957
img name: 1-2.jpg Licence plate: 沪A·F02976
img name: 1-3.jpg Licence plate: 鲁N·BK268
#### medium ####
img name: 2-1.jpg Licence plate: 沪E·WM957
img name: 2-2.jpg Licence plate: 豫B·20E68
img name: 2-3.jpg Licence plate: 沪A·93S20
#### difficult ####
img name: 3-1.jpg Licence plate: 沪E·WM957
img name: 3-2.jpg Licence plate: 沪A·DE6598
img name: 3-3.jpg Licence plate: 皖S·J6M07
```

图表 3 - 2 识别结果输出

### 3.2 实时性分析

作为一个车牌识别系统，显然算法识别的速度存在一定要求。比如在停车场出入口处，倘若识别速度较慢，很可能造成出入口堵塞。如图表 3-3 所示，我们统计了算法所有步骤的耗时，可以发现，虽然在预训练时需要较长时间，但是在实际应用场景下，

模型参数应该是预先训练好后，加载到该算法中执行，因此我们只需要考虑加载模型-车牌定位-字符分割-字符识别共四个步骤的运行时间即可，四个步骤总计耗时 $\approx 0.036s$ ，识别过程不超过  $0.1s$ ，运行速度较快，满足实际应用场景中的实时性要求，较好的利用了传统方法中速度快的优势。

Process	中文字模型训练	字符模型训练	加载模型
Time	9.17s	54.33s	0.009s
Process	车牌定位	字符分割	字符识别
Time(Avg)	0.017s	0.008s	0.002s

图表 3 - 3 算法各模块耗时

### 3.3 评价指标

首先，我们定义准确率 Auc 作为衡量识别结果质量的指标，准确率 Auc 定义为给定的测试集上识别完全正确的车牌与所有车牌个数之比，因为在实际应用场景中，对于一个车牌哪怕只有一个字符识别错误也是不可原谅的错误，因此该指标直接衡量了该算法在实际使用场景下的可信程度。

其次，我们参考使用 OCR 算法中常见的评价指标——编辑距离(Levenshtein Distance)来衡量车牌识别算法识别错，漏识别和多识别的情况。编辑距离指两个字串之间，由一个转换成另一个所需的最少编辑操作次数，显然，编辑距离越小，说明我们的算法识别效果越好。在实际测试中，我们选择对所有车牌识别的编辑距离取平均，直接得到平均编辑距离。

为了验证该算法的鲁棒性，我们尝试对我们设定好的超参进行随机的扰动，并计算上述评价指标结果，进一步衡量该算法对于参数的敏感性。

Load config name	Auc	Levenshtein Distance
config.json	100%	0
config1.json	88%	0.11
config2.json	88%	0.22

图表 3 - 4 评价指标

如图表 3-4 所示，显然两个评价指标都具备合理性，且评价的侧重点并不完全相同，比如加载超参文件 config1 时，3-3.jpg 车牌识别错误，编辑距离为 1；而对于 config2，3-2.jpg 车牌识别错误，编辑距离为 2，因此和 config1 的结果相比，虽然准确率相同，但是 config1 计算的编辑距离比 config2 的结果更好，因此可以得出结论 config1 的结果比 config2 效果更好。

## 第四章 结论

对于车牌识别任务，我们提出了一种基于传统算法定位和分割车牌，基于深度学习方法识别字符的车牌识别算法。该算法在对图像进行预处理后，结合颜色特征和边缘特征对精准定位图像中的车牌位置，使用 hough 变换实现了对车牌的矫正，然后使用投影的方法实现了对车牌字符的分割，最后采用了 CNN 来对车牌字符进行识别。该算法结合了传统方法和深度学习方法的优点，在测试集上达到了 100% 的准确率，具备一定泛化性，且满足实时性要求。

然而，我们的算法的超参较多，且设计的 CNN 网络比较简陋，在某些图片上很可能出现车牌定位失败或是分割的字符不能准确识别等问题。且该算法没有考虑到光照影响和环境影响，对于过曝或者大雾天气下拍摄的照片，仍然不能较好的处理。

## 参考文献

- [1] 孙金岭, 庞娟, 张泽龙. 基于颜色特征和改进 Canny 算子的车牌图像定位 [J]. 吉林大学学报 (理学版), 2015, 53 (4): 693-697.
- [2] 王卫, 陈继荣, 徐璟业. 基于颜色特征的车牌快速定位 [J]. 计算机工程与应用, 2006 (1): 226-229
- [3] 潘寒飞. 基于颜色特征的车牌定位与分割技术研究 [D]. 重庆: 重庆大学, 2012.
- [4] 卓炜, 齐春. 基于边缘信息的车牌定位 [J]. 中国科技论文 在线, 2011, 6 (4): 305-309
- [5] 廖晓姣, 李英. 基于边缘检测和形态学的车牌定位算法 [J]. 现代电子技术, 2011, 34 (10): 17-19
- [6] 霍祥湖. 基于卷积神经网络的车牌识别技术研究 [D]. 成都: 电子科技大学, 2017.
- [7] 赵先军. 基于神经网络的车牌识别技术研究 [D]. 西安: 西安电子科技大学, 2005.
- [8] 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.
- [9] 姚德宏. 基于神经网络的汽车牌照提取研究. 计算机应用, Vol. 21, No. 6 Jun., 2001
- [10] 张云刚, 张长水. 利用 Hough 变换和先验知识的车牌字符分割算法 [J]. 计算机学报, 2004(01): 130-135.
- [11] 何兆成, 余锡伟, 余文进, 杨文臣. 字符多特征提取方法及其在车牌识别中的应用 [J]. 计算机工程与应用. 2011, 47(23): 228-231
- [12] Xing-Wang Zhang, Xian-gui Liu, Jia Zhao. A Vehicle License Plate Recognition Method Based on Neural Network [J]. Granular Computing, 2010: 845 – 847