
目 录

第一章 绪论	1
第二章 实验分析	1
2.1 实验要求分析	1
2.2 实验解决思路	1
第三章 深度学习车牌识别算法	2
3.1 整体流程	2
3.2 算法原理	2
3.2.1 U-net 神经网络车牌定位	2
3.2.2 opencv 透视矫正	4
3.2.3 基于像素直方图的字符分割	5
3.2.4 基于 LPRNet 神经网络的字符识别	5
第四章 实验测试结果与分析	6
4.1 测试结果与分析	6
4.2 算法性能分析	7
第五章 数据集准备	8
第六章 结论	8
参考文献	10
谢辞	10

第一章 绪论

车牌识别是指通过计算机视觉技术识别出车辆的车牌号码，并进行自动化处理的过程。车牌识别技术的发展已经经历了从传统的基于规则的方法到深度学习方法的转变。深度学习方法通过神经网络自动学习特征，避免了手动提取特征的繁琐过程，并且具有更高的识别准确率和鲁棒性。它是智能交通系统的一个重要组成部分，能够有效提高道路交通的管理和安全性能。车牌识别技术已经广泛应用于城市交通管理、停车场管理、高速公路收费等领域。在实际应用中，车牌识别系统还需要考虑识别速度、识别率、可靠性等因素，以适应不同场景下的需求。车牌识别技术在未来还将不断发展，有望在智慧城市建设、无人驾驶等领域得到更广泛的应用。

第二章 实验分析

2.1 实验要求分析

实验要求为设计算法检测车牌位置并识别车牌号，其任务主要如下：

1. 车牌定位：车牌定位是车牌识别的第一步，也是最关键的一步。需要克服车牌倾斜、遮挡、光照不均等问题，以准确地定位车牌的位置和大小。
2. 透视矫正：透视矫正是车牌定位的一个重要步骤，对于倾斜、畸变的图像进行变换，使其变为平面上正常的图像，提高车牌字符的分割和识别的准确率。
3. 字符识别：字符识别是车牌识别的核心环节，需要对分割出来的字符进行识别。由于车牌中的字符存在大小写、字体不同等问题，因此需要采用深度学习等高级技术，提高识别准确率。
4. 数据集的准备：车牌识别需要大量的标注数据，以训练和测试模型。数据集的构建需要考虑车辆类型、车牌颜色、不同角度和光照条件下的图像等多种因素。同时，车牌数据属于隐私数据，需要保证数据的安全性和隐私性。

实验测试图片根据识别难度分为 easy, medium, difficult，对于 easy 难度的图片只需要进行字符识别，medium 难度与 difficult 难度的测试图片难度在车牌定位与透视矫正上难度有所增加。

2.2 实验解决思路

针对车牌识别实验中的难点和要求，我的解决思路如下。

首先对于问题 1，我采用了 Unet 图像分割神经网络来进行分割获得车牌区域的二值化图片。

对于问题 2 的解决思路是在获得的二值化图像的基础上运用 Canny 边缘检测以及

opencv 寻找轮廓算法获得车牌图像四边形框架以及四个顶点坐标,根据顶点坐标运用 opencv 函数进行透视投影变换。

在问题 3 字符识别上,我采用 LPRNet 卷积神经网络进行车牌内容识别。

在问题 4 数据集准备上,我采用的是 CCPD 车牌数据集,利用已经标注好的车牌坐标进行车牌定位训练的数据集,由于 CCPD 车牌数据集车牌省份上数量有严重的差别,导致其无法作为字符识别样本,对此我采用虚拟生成车牌的方法生成各种省份车牌号照片作为字符识别数据集。

第三章 深度学习车牌识别算法

3.1 整体流程

通过深度学习算法对车牌识别的整体流程如下:

1. 数据集整理上,下载 CCPD 数据集筛选部分照片对照片文件名提取车牌定位坐标 label 信息,并生成 train 和 test 用例的 numpy 数组。运用虚拟车牌模型生成车牌图片并提取 label 信息用于字符识别。
2. 车牌定位模型采用 U-net 对已经整理好的 CCPD 蓝色车牌和绿色车牌分别训练,对测试集进行评估,模型结果生成车牌所在区域二值化图像。
3. 使用 opencv 函数对图像进行轮廓检测,找到四边形端点坐标并进行透视变换得到正视图像。
4. 对图像进行字符分割,采用二值化图像像素值在 X 轴上的分布进行分割字符,得到分割后的单个字符。
5. LPRNet 卷积神经网络对生成的车牌照片进行端到端多标签分类训练,结果输出车牌字符内容。
6. 运用训练好的 LPRNet 模型对透视矫正后的车牌图像进行字符识别得出识别结果。

3.2 算法原理

3.2.1 U-net 神经网络车牌定位

U-Net 普遍应用在生物医学影像领域,其在架构设计和其他利用卷积神经网络基于像素的图像分割方面更成功,它甚至对有限数据集的图像更有效。

U-Net 的主要特点是将编码器和解码器结合在一起,同时在网络中使用跳跃连接来连接相应的编码器和解码器层。编码器部分类似于常规卷积神经网络,包括卷积层和池化层,其目的是从输入图像中提取特征。解码器部分包括反卷积层和卷积层,它将编码器中提取的特征映射复原回输入图像的大小。

其网络结构如图 3-1 所示^[1]。

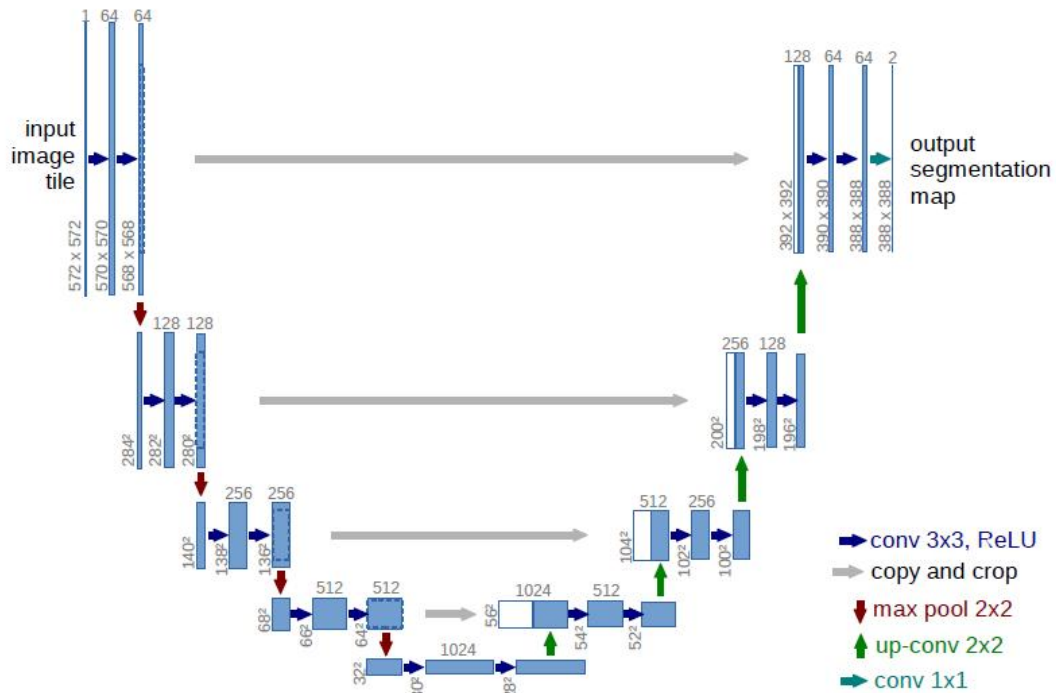
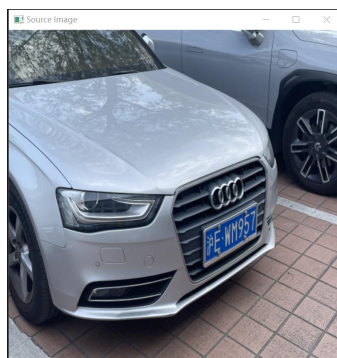


图 3-1 U-Net 的网络结构

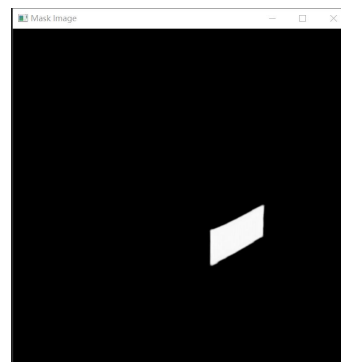
具体而言，U-Net 的网络结构包括以下层：

- (1) 输入层：接收原始图像作为输入；
- (2) 编码器部分：由多个卷积层和池化层组成，每个卷积层后面通常跟一个 ReLU 激活函数和一个批量归一化操作；
- (3) 跳跃连接：将编码器的每个卷积层与对应的解码器层连接起来，用于在解码器部分利用编码器中提取的特征信息；
- (4) 解码器部分：由多个反卷积层和卷积层组成，每个反卷积层后面通常跟一个 ReLU 激活函数和一个批量归一化操作；
- (5) 输出层：输出分割掩模图像，用于将原始图像分割成不同的区域。

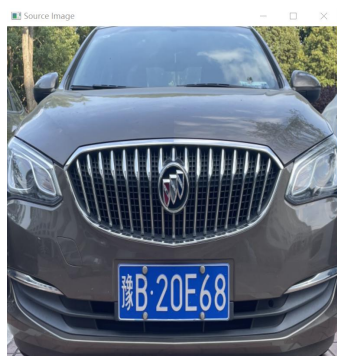
由于蓝色车牌与绿色车牌色彩特征不同，将其样本混合进行训练需要更多的计算资源，为了更准确定位车牌区域，我采用 U-net 分别对两种车牌进行训练得到针对分别两种车牌的模型。经过训练，模型输入含有车牌的照片，模型输出只有车牌区域为 255，其他区域为 0 的二值化图片，如图 3-2 所示。



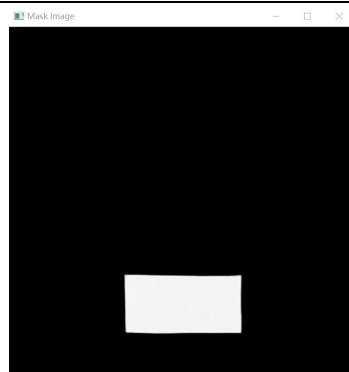
(a) 3-1 原图



(b) 3-1 车牌定位二值化图



(c) 2-2 原图



(d) 2-2 车牌定位二值化图

图 3-2 图片与车牌定位二值化图

3.2.2 opencv 透视矫正

在得到车牌区域的二值化图像后，运用 opencv 的轮廓检测函数 `findContours()` 对二值化图像进行轮廓检测，计算轮廓的周长和面积筛选出周长大于 100 且面积大于 200 的轮廓区域以消除噪声的影响。对筛选好的轮廓进行平行四边形轮廓近似，运用 opencv 的 `approxPolyDP()` 函数得到一个包含所有白色像素的平行四边形轮廓。将平行四边形轮廓画到原图上，如图 3-3 所示。



图 3-3 3-1 车牌定位图

由此得到四边形四个顶点坐标列表 `src_pts` 并对其进行排序按照左上角，右上角，右下角，左下角的顺序。为将车牌区域透视变换后呈现到 440x140 的图像上，对应车牌坐标变换后的顶点坐标 $(0, 0), (440, 0), (440, 140), (0, 140)$ 。

运用 opencv 的函数 `getPerspectiveTransform()` 函数根据变换前后坐标计算变换矩阵，图 3-4 展示了运用 `warpPerspective()` 函数进行透视变换得到矫正后的车牌图像。



图 3-4 透视变换处理后的车牌

3.2.3 基于像素直方图的字符分割

基于矫正后的车牌图像，运用 `opencv` 函数对其进行灰度化并做二值化处理。对处理后的图像统计其每一列的白色像素和黑色像素总和数组 `white`、`black`，并统计出每一列白色像素和黑色像素的最大值 `white_max`、`black_max`。接着从统计后的数组中遍历如果其白色像素值数量达到最大值 `white_max` 的 0.1 倍则认为该列为单个字符的开始。接着寻找字符的结尾列，如果列满足黑色像素值达到最大值 `black_max` 的 0.9 倍且距离字符开始有 5 个列的距离则可以认为这是字符的结束列。以此类推直到分割完全部字符，结果如图 3-5 所示。

Figure 1

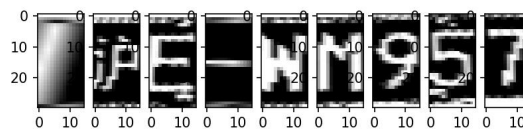


图 3-5 字符分割处理结果

由于 `difficult` 难度下透视变换后难免有视角误差，难以得到完全正视的图像，每一列上像素分布较于理想情况有所差别。且车牌上省份后的间隔符号对字符分割产生了干扰，在 `difficult` 难度下字符分割效果一般。但是在 `easy` 和 `medium` 难度上字符分割取得较好的成果。

3.2.4 基于 LPRNet 神经网络的字符识别

LPRNet 是一个深度学习模型，用于自动车牌识别。LPRNet 使用卷积神经网络来识别图像中的车牌，并且可以处理多种不同类型的车牌。主要包括三个部分：卷积神经网络特征提取、循环神经网络序列建模和注意力机制融合。下面是每个部分的详细介绍^[2]：

(1) CNN 特征提取

LPRNet 首先使用卷积神经网络（CNN）提取车牌图像的特征。这里使用了一个轻量级的 CNN 网络来提取车牌的局部和全局特征，包括车牌字符的形状、颜色和纹理等信息。该网络包括多个卷积层和池化层，通过不断地缩小特征图的尺寸来提取更高级别的特征。

(2) RNN 序列建模

CNN 提取的特征是车牌图像的局部和全局信息，但是它们没有被正确排序，因此需要进一步使用循环神经网络（RNN）对其进行序列建模。LPRNet 采用了双向 LSTM 来对 CNN 提取的特征进行序列化建模。双向 LSTM 可以同时考虑前向和后向的信息，从而更好地捕捉车牌字符之间的相互关系。

(3) Attention 机制融合

最后，LPRNet 引入了注意力机制来融合 CNN 和 RNN 的特征，从而进一步提高识别准确率。注意力机制可以动态地调整不同位置的特征权重，使得网络可以更好地关注车牌中最有用的信息。LPRNet 中使用了双重注意力机制，一层用于捕捉字符的空间信息，另一层用于捕捉字符的时序信息。

其网络结构如图 3-6 所示：

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

图 3-6 LPRNet 网络结构

经过训练后，将已经透视矫正后的车牌图像变换为 94x24 的格式后输入模型得到车牌字符识别结果，具体结果在下一章给出。

第四章 实验测试结果与分析

4.1 测试结果与分析

首先运用 U-net 网络与透视变换将 medium 与 difficult 难度的车牌照片进行定位裁剪得到处理后的图像，见图 4-1。



(a)



(b)



(c)



(d)

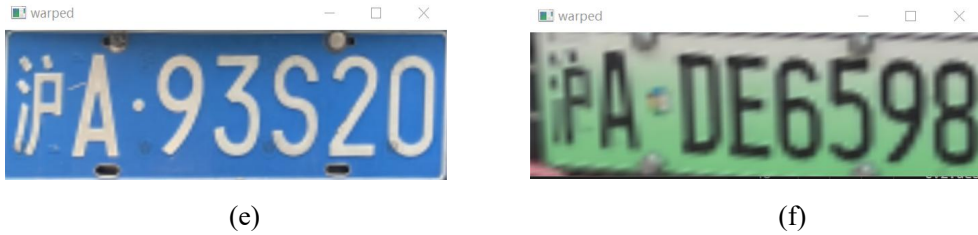


图 4-1 透视变换结果图

从图 4-1 中可以看到左边 medium 难度图片车牌定位以及透视变换所得图片十分准确，右边 difficult 难度图片由于车牌定位上边缘定位有略微误差导致透视变换后效果略有偏移，但总体而言车牌定位以及透视变换模型效果取得成功。

运用 LPRNet 网络对透视变换后的车牌进行字符识别，其结果如图 4-2 所示：



图 4-2 字符识别结果图

具体代码输出如图 4-3 所示，其中将 easy 至 difficult 难度图片分别排序 0-8。

```
load pretrained model successful!
target: 5 #### F #### predict: 沪A93S20
target: 2 #### F #### predict: 鲁NBK268
target: 6 #### F #### predict: 沪EWM957
target: 3 #### F #### predict: 沪EWM957
target: 1 #### F #### predict: 沪AF02976
target: 4 #### F #### predict: 豫B20E68
target: 8 #### F #### predict: 皖SJ6M07
target: 7 #### F #### predict: 沪ADE6598
target: 0 #### F #### predict: 沪EWM957
```

图 4-3 识别结果输出

从结果可以看出字符识别上通过了所有难度图片的测试，准确度达到了 100%。在最初进行字符识别时，由于训练样本并不是完美裁剪的车牌照照片，所以在 easy 难度上完全将车牌输入模型会导致识别效率低下。对此我采取将所有照片裁剪后的照片留白 1 个像素边框，缩小后的照片识别效率极好。

4.2 算法性能分析

在准确率上，给出三种难度照片的识别任务上准确度达到了 100%。但是仅仅由 9 张照片无法评价其算法性能，在运用 CCPD 部分数据集进行测试后，最终准确度达到了 91%。对测试错误的结果进行分析，我发现字符识别在识别省份汉字时，皖和闽的识别有较大的误

差，且由于光照等引起像素干扰的因素下，部分数字以及字母识别上会产生误差如 G 和 0，Q 和 0 的识别上容易产生错误。给出三种难度的图片中即使是 difficult 难度其识别难度也要远低于实际生活中恶劣环境下的识别难度。

在算法的运行时间上，对测试图片进行性能分析，如图 4-4 所示：

[Info] Test Speed: 5.30410811636183s 1/9]

图 4-4 运行时间

可以看到在对 9 张图片进行读取识别字符时运行时间有 5.3s，对于每张照片识别上平均时间有 0.6s，在实际应用中对于时间延迟宽裕，模型运行速度可以相适应。对于高时间延迟要求的情况如自动驾驶等应用中，模型的识别效率还需要提升优化。

第五章 数据集准备

在数据集准备上，我采用的是 CCPD 数据集和虚拟车牌生成模型生成数据集两种方式结合。对于 CCPD 数据集进行有效 label 的提取成 numpy 数组方便进行车牌定位的训练。由于 CCPD 数据集除了安徽省，其他省份车牌数量上较少难以支持字符识别的任务，我借鉴虚拟车牌生成模型，对其进行部分修改使其能够生成 7 个字符以及 8 个字符的车牌进行生成训练集^[3]，见图 5-1。



运用虚拟车牌生成模型得到各个省份的车牌照照片，再经过二值化等图像处理以方便进行字符识别的训练。

图 5-1 车牌生成图片

第六章 结论

本文基于深度学习解决对车牌定位以及车牌字符识别的任务。从 U-net 图像分割网络上解决车牌定位的难点，运用透视变换等图像处理手段得到车牌区域图像，从 LPRNet 卷积神经网络对车牌进行字符识别。

算法从性能上分析，在给出的三种难度测试集上达到了 100% 的准确度，由于实际情况复杂，现实识别当中准确度仅有 91%，一方面由于训练样本过少，计算资源缺乏，另一方

面现实中识别难度受多种因素影响。在时间延迟上，0.6s 的延迟误差达到了轻量化准确识别的目的效率。

借用虚拟车牌生成模型，对 CCPD 数据集进行车牌字符内容识别训练集补充。

参考文献

- [1] LIU Duansheng. 2021. End to End for Chinese Plate Recognition. <https://github.com/duanshengliu/End-to-end-for-chinese-plate-recognition>. (2021).
- [2] Zherzdev, S. and Gruzdev, A. LPRNet: License Plate Recognition via Deep Neural Networks[J]. arXiv e-prints, 2018. doi:10.48550/arXiv.1806.10447.
- [3] YAN Guolong. 2020. Generate Car Plate. <https://github.com/derek285/generateCarPlate>. (2020).

谢辞

感谢赵旭老师在计算机视觉这门课中为我们讲述 cv 领域的专业知识，通过这门课我不仅在对图像处理上学习到了更多方法，除此之外在应用作业上也让我的实践能力不断提升。最后感谢万助教对我的每次作业作出的认真仔细的评改意见。