

题 目 漫画翻译中的文字定位与识别

摘 要

本文主要研究简单背景下文字的定位与识别，即 OCR 技术。OCR 技术是指利用计算机技术或者光学技术，把印刷在纸上的字体或者手写在纸上的字体识别出来，转换为计算机能认识的文字。OCR 技术目前在现实生活中的应用很广泛，比如我们手机上的一些扫描类的 APP，比如车牌识别技术等等。故 OCR 技术在未来的发展前景也很广阔。

本文的课题研究一共分为两个步骤。第一步为简单背景下的文本定位，该部分主要是使用一些图像处理技术使得一副含有文字的图片呈现白底黑字，然后使用文本切割技术把该图片切分为一幅幅只含有单个文字的图片；第二部分为文字识别技术，即对第一步切出来的单个文字的图片进行识别。该部分首先需要生成中文文字的数据集，然后搭建训练的网络，最后对模型进行训练。

本文在文本定位方面，目前对于简单背景下的文字图片，基本可以成功切分出来，不过仍存在着值得改进的地方，比如对于有些拥有偏旁部首的汉字切分时可能出错。在文字识别方面，训练出的模型的识别准确率能达到百分之九十九以上，且对于倾斜的字体或是拥有其它干扰的字体，比如笔画黏连或断裂，亦可成功识别。

关 键 词：简单背景；文本定位；文字识别；深度学习

ABSTRACT

This paper mainly studies the positioning and recognition of text in a simple background, namely OCR technology. OCR technology refers to the use of computer technology or optical technology to identify fonts printed on paper or fonts handwritten on paper, and convert them into words that can be recognized by computers. OCR technology is currently widely used in real life, such as some scanning apps on our mobile phones, such as license plate recognition technology. Therefore, the future development of OCR technology is also very broad.

The research of this paper is divided into two steps. The first step is to locate the text in a simple background. This part mainly uses some image processing technology to make a picture containing text appear black on white, and then use text cutting technology to cut the picture into a single frame with only a single The picture of the text; the second part is the text recognition technology, which is to identify the picture of the single text cut out in the first step. This part first needs to generate a Chinese character data set, then build a training network, and finally train the model.

In terms of text positioning, the text texts in the simple background can be successfully segmented, but there are still some areas worthy of improvement. For example, some Chinese characters with radicals may be mistaken. In terms of text recognition, the trained model can achieve a recognition accuracy of more than 99%, and can be successfully identified for oblique fonts or fonts with other interferences, such as strokes or breaks.

KEY WORDS: Simple background; Text localization; Text recognition; Deep learning

目 录

1 绪论	11
1.1 研究背景	11
1.2 应用场景	11
1.3 OCR 的分类	11
1.4 OCR 工作流程	12
1.4.1 图像输入	13
1.4.2 图像预处理	13
1.4.3 提取文字特征	13
1.4.4 对比数据库	13
1.4.5 对比识别	14
1.4.6 人工校正	14
1.4.7 结果输出	14
1.5 OCR 的发展	14
1.6 论文研究目的及意义	15
1.7 论文组织结构	16
2 图像预处理	16
2.1 灰度化	17
2.2 二值化	18
2.3 图像降噪	19
2.4 倾斜校正	20
2.5 文字切分	21
3 中文文字数据集	22
3.1 文字数据集概述	22
3.2 生成文字数据集	23
3.2.1 生成汉字映射表	23
3.2.2 字体文件的收集	23
3.2.3 生成汉字图像	24
3.2.4 图像的增强处理	24
4 网络搭建	25
4.1 LeNet 网络	25
4.2 深度神经网络中的归一化	26
4.3 改进策略	26
4.3.1 策略 1: 添加激活层	26

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要
下载或阅读全文，请访问：

<https://d.book118.com/956221055122011004>