# 项目设计文档

1）引言

1、使用场景

当我们收到一个没有电子版的，或者电子版缺失的表格报告等文档图片，可以通过程序快速复现出文档的电子版

2、项目背景

百度AI开放平台提供的办公文档识别模块仅仅对文档中的各类信息进行了位置、类型、内容的识别，但是在实际情况下，对于办公文档图片的识别通常要用于照片中文档的电子再现，为弥补这一不足，开发本程序

2）功能分析

1、用户需求分析

制作一个能够复现图片中所示文件的电子版的程序

2、业务需求分析

对文档图片进行识别，首先进行表格识别，再有内容的情况下生成excel文件，其次进行文本识别，对于非重复文本进行存储

3、功能需求分析

1. 一个dialog界面，需要包括一个button\_1用来选择图片，需要一个Label\_1用来展示选中的图片，需要一个转换button\_2用来开始进行操作，操作结束后，也即txt生成结束后，该窗口关闭，打开第二窗口，给出excel文档与txt文档的文档名及存储路径/打开button\_3/4

2.一个用来调用文档图片增强的类，一个返回base64编码的增强图片

3.一个用来调用表格识别的类，一个属性存储表格的位置信息，同时这个类还需要包括excel的生成，也就base64转excel

4.一个用来调用文本识别的类，同时这个类需要从表格识别类获取表格的坐标范围，需要一个类内方法来处理当前处理文本是否位于表格内，需要一个非表格内容存入txt的方法

3）技术结构

1. 项目概述

本项目旨在开发一个图像处理应用，利用光学字符识别（OCR）技术从图像中提取数据并转换为可编辑的Excel文件和文本文件。项目采用Python编写，并结合了PyQt5用于图形用户界面（GUI）的实现。

2. 技术栈

编程语言：Python

图形用户界面：PyQt5

图像处理：OpenCV

光学字符识别：百度OCR API

数据处理：Pandas

Excel文件处理：openpyxl

HTTP请求：requests

3. 项目结构

project\_root/

│

├── app/

│ ├── \_\_init\_\_.py

│ ├── myapp.py

│ ├── maindialog.py

│ ├── mainsec.py

│ ├── secondui.py

│ └── mdui.py

│

├── ai/

│ ├── ToExcel.py

│ └── ToWords.py

│

├── data/

│ └── excel1.jpg

│

├── test1.py

└── output/

├── output.xlsx

└── output.txt

4. 主要模块和功能

4.1. test1.py

功能：启动应用程序并退出。

4.2. myapp.py

功能：定义应用程序的主入口，初始化并显示主对话框。

4.3. maindialog.py

功能：定义主对话框，显示图像并处理Excel文件。

4.4. ToExcel.py

功能：调用百度OCR API处理图像并保存结果为Excel文件。

4.5. ToWords.py

功能：从图像中提取文字并保存到文本文件。

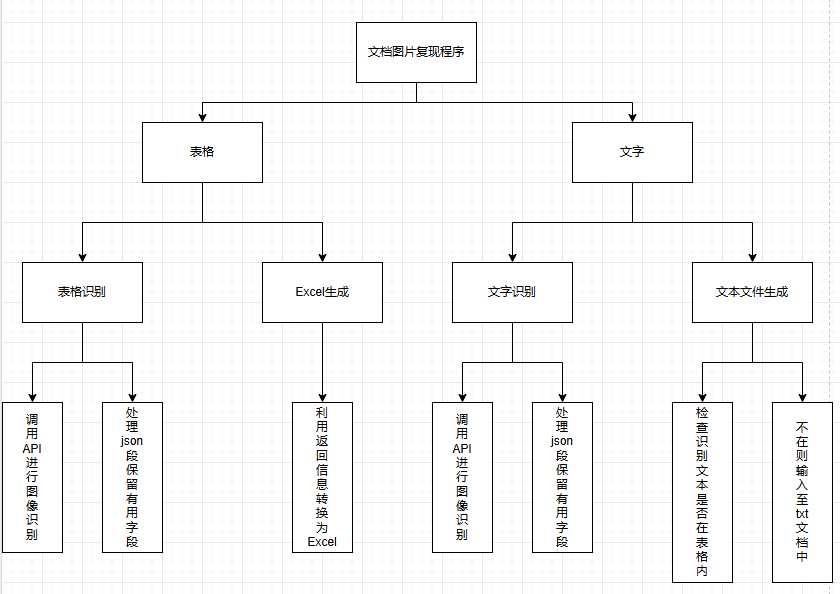
4.6. mainsec.py

功能：展示生成的Excel和文本文件内容。

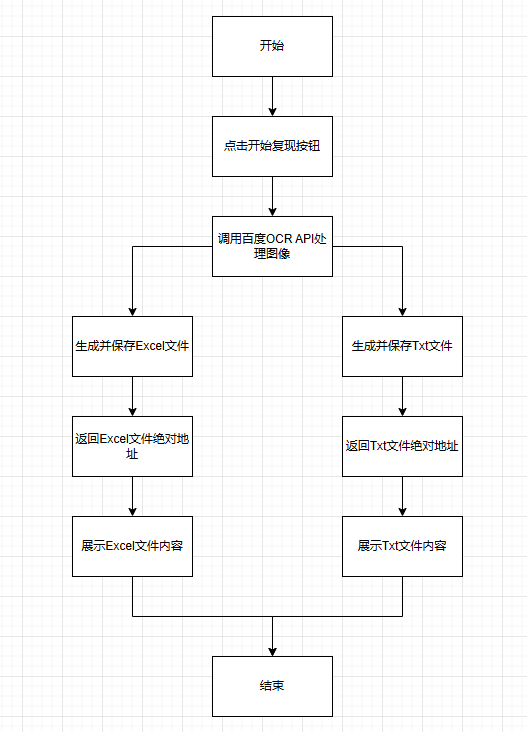
5. 依赖项

* PyQt5
* OpenCV
* requests
* pandas
* openpyxl

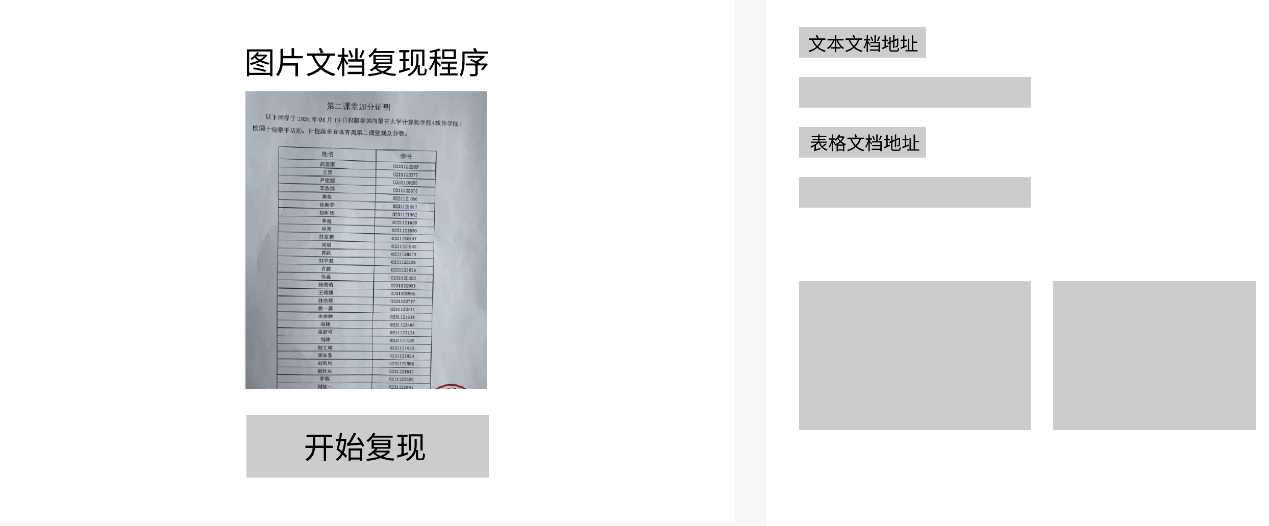
4）图示



功能结构图 1



流程图（用户） 1



低保真原型图 1