AI\_Final\_Project 实验报告

金英豪 10172100162

本次实验项目与双创项目“多角度拍摄下的图像处理与信息提取”的子课题挂钩，同组同学有：曹汇杰，王子玥（在本次期末项目同样负责一部分内容）。

本实验报告主要分为以下四个部分：项目简介、背景概览、代码实现以及成果展示

1. 项目简介

双创项目“多角度拍摄下的图像处理与信息提取”的目的是开发一款针对学校课堂、学术报告上PPT或PDF进行图像处理和信息提取的软件。其主要功能如下：

* 能根据不同拍摄方向对投影文件进行识别、截取及角度矫正处理。
* 能针对一系列在相似角度拍摄的照片（如同一位置拍摄的同一讲座的一系列照片）进行自动化的批量处理。
* 针对底色过暗的问题，进一步提高识别的准确率。
* 能将照片中的内容（公式、文字等）识别、提取，并将识别的照片自动转化为PPT文档，使其内容可编辑。
* 能对照片里的内容进行智能化处理，如将其中的网页做成超链接或者直接复制至剪贴板，对电话、邮箱、姓名等自动添加至通讯录。
* 在拍摄照片时对其中的某些数学概念，公式等进行智能注释。

在实际项目的开发中，遇到的第一个问题就是：如何在真实照片中寻找目标（PPT）轮廓，并进行下一步的矫正、识别。经研究后提出两种方案：

1. 神经网络自动判断勾勒（机器学习）
2. OpenCV库识别（计算机视觉）

在实际运行时，由于机器学习的方案缺少相应数据集，自己收集又难以达到预计量，因此采用第二种方案：计算机视觉方面的应用。

1. 背景概览（摘自CNBLOG以及百科等相关部分）

OpenCV于1999年由Gary Bradsky在英特尔创立，第一个版本于2000年问世。随后Vadim Pisarevsky加入了Gary Bradsky，主要负责管理英特尔的俄罗斯软件OpenCV团队。2005年，OpenCV被用于Stanley车型，并赢得2005年DARPA大挑战。后来，它在Willow Garage的支持下持续并积极发展，转为由Gary Bradsky和Vadim Pisarevsky领导该项目。OpenCV现在支持与计算机视觉和机器学习相关的众多算法，并且正在日益扩展。

OpenCV支持各种编程语言，如C ++，Python，Java等，可以在不同的系统平台上使用，包括WindowsLinux，OS X，Android和iOS。基于CUDA和OpenCL的高速GPU操作接口也在积极开发中。

OpenCV-Python是OpenCV的Python API，结合了OpenCV C++ API和Python语言的最佳特性。

1. 代码实现

具体代码在scanner.py和rect.py中

整体思路：传入图片→调整大小、尺寸→灰度平滑化处理→应用Canny边缘检测→保留最大值→获取近似轮廓→打印图片→目标点映射

特殊部分详解：

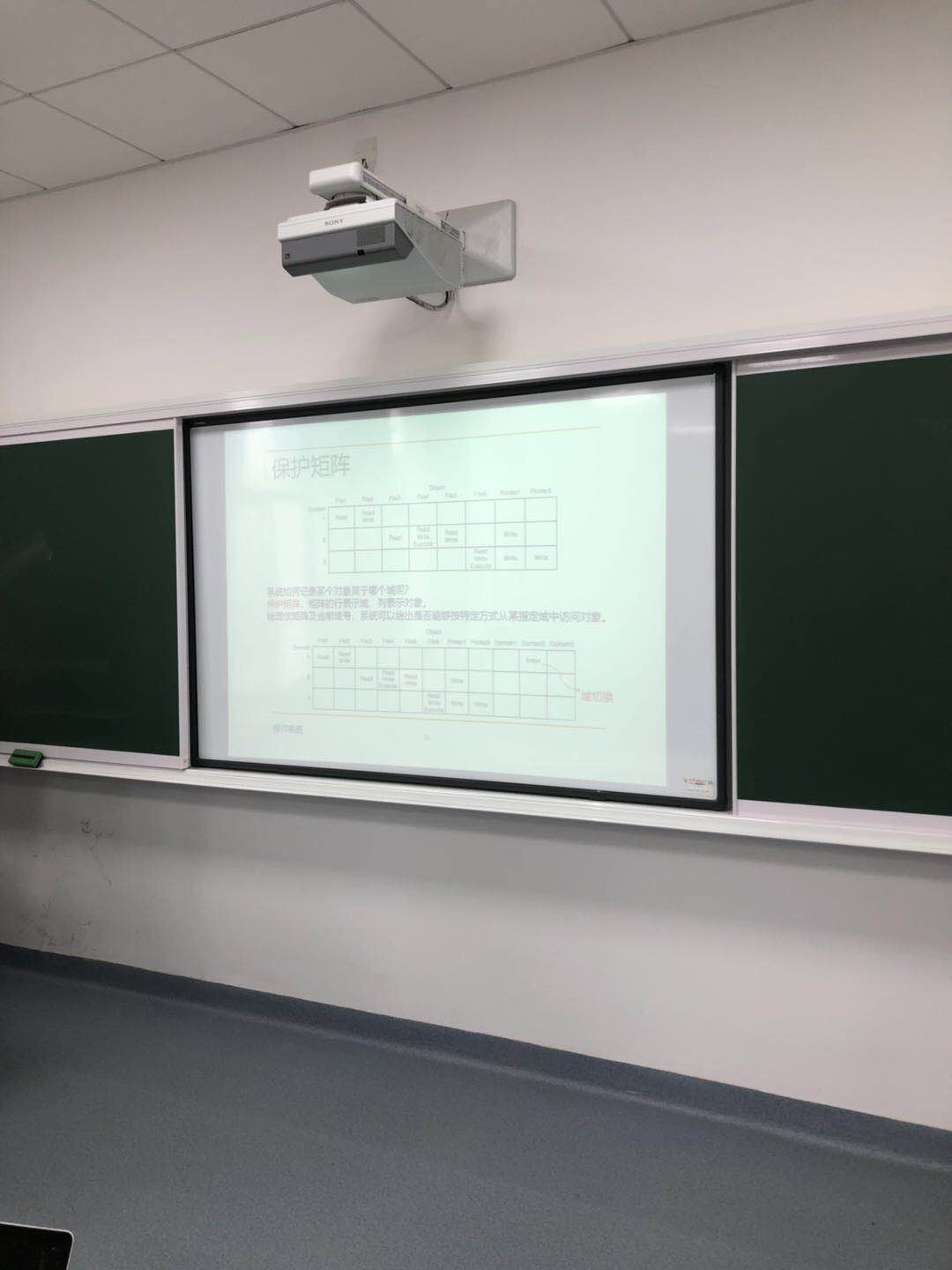
* 保留最大值的目的：在实际项目应用中，一张照片中只会有一页PPT（只有一个可识别目标），框选多个边框没有实际意义
* 输出多种图像：代码保留了输出多种图像的选项，包括灰度图像、边框勾勒图像等，以便保留应用层面拓展的可能。
* 坐标的输出：为了程序的连贯性，将识别后的边框（保留的最大值）映射到800x800四边形，并将映射后的4个坐标输出到result.txt文件中以便后续程序对接。该项操作位于单独的rect模块（在主程序开始时作为库导入）
* 代码配备中英文注释，可用编辑器打开查看

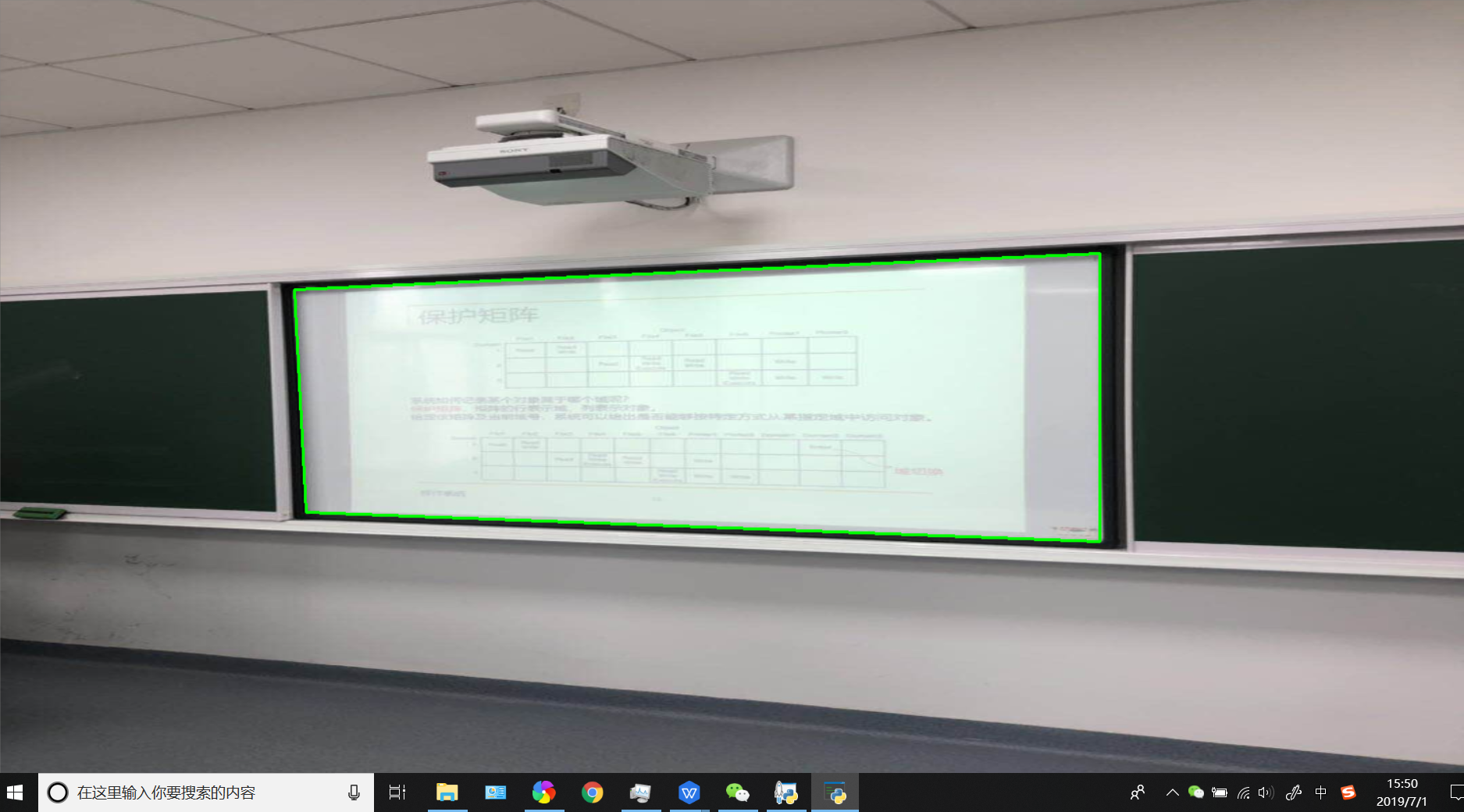
1. 成果展示与说明

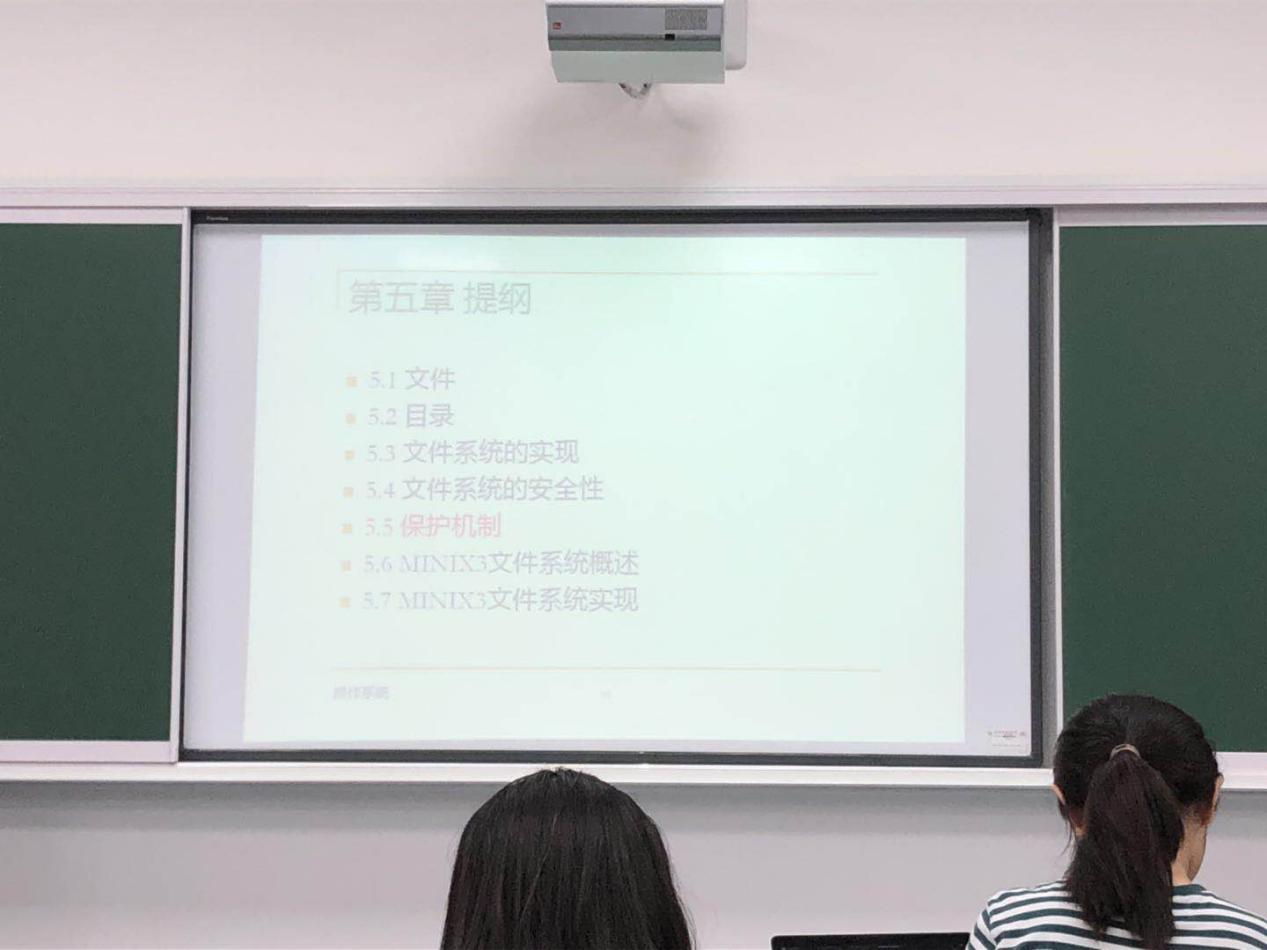
代码已经封装完成。具体应用步骤：

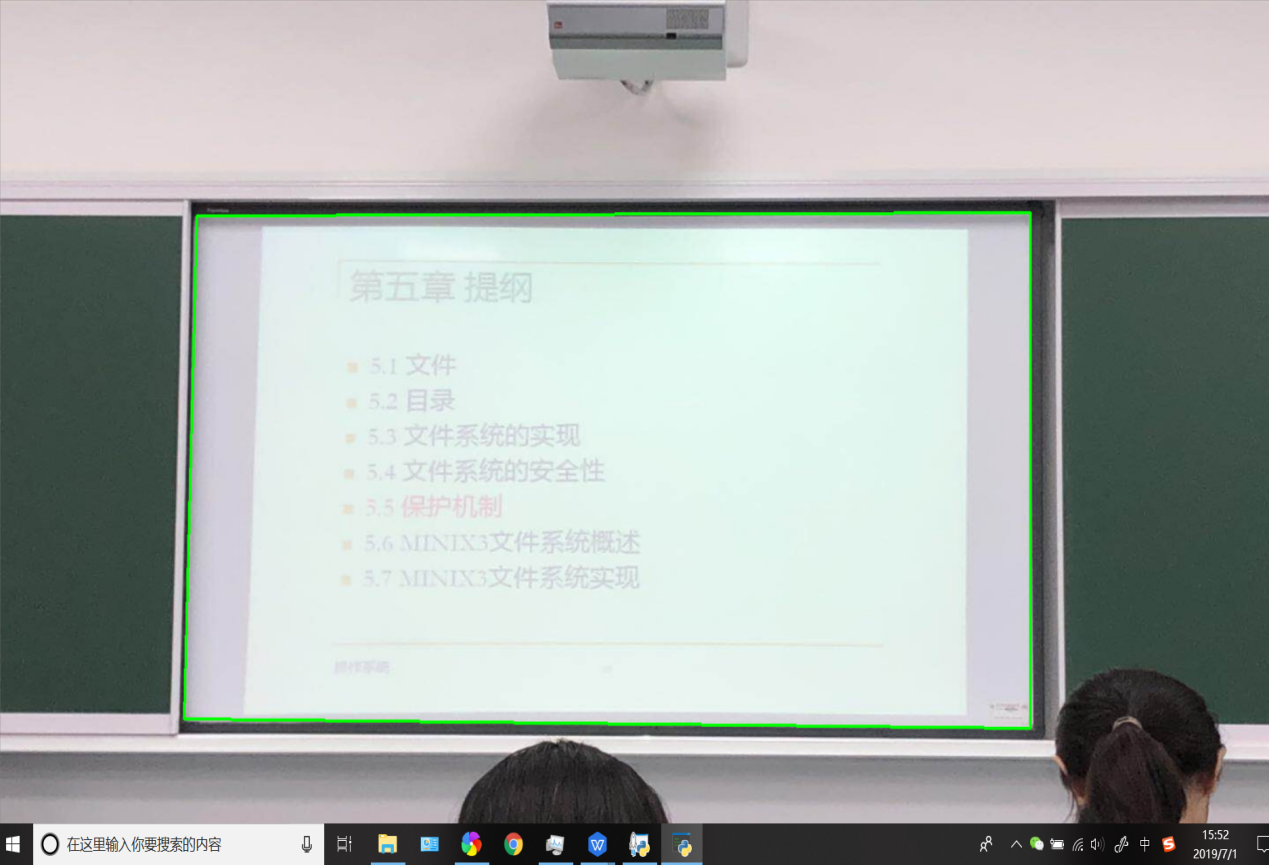
1. 安装Numpy、opencv库
2. 将图像名改为test，并放于代码相同路径（或更改源代码中相应图片名位置）
3. 命令行中运行scanner.py（若已将python作为默认打开方式，可直接打开scanner.py运行）
4. 得到勾勒边框后图像，坐标位于result.txt中（没有的话会自动生成）

样例展示：









可以看出较好地完成了预期目标，且程序具有较高的模块性，可直接嵌入主项目中运行。