

选题报告

一、产品名称：

乐游——基于图像的实体识别与导游系统

二、小组成员：

张钰晖, 2015011372, yuhui-zh15@mails.tsinghua.edu.cn

周正平, 2015011314, zhouzp15@mails.tsinghua.edu.cn

陈经基, 2015011358, greencLOUDS@foxmail.com

三、产品背景：

随着人们出行日益便利，旅游成为了现代人生活中一个很重要的环节。风景带给游客视觉的享受，而知道景物背后的故事更将锦上添花。

目前大部分景区仍然是传统游览方式，即需要导游来担当起解说的职责，这不仅限制了人们自由支配时间，游客聚集影响了观赏体验，而且需要投入大量人力物力；少部分景区已经采用二维码方式，游客走到对应的位置可以通过扫二维码可以收听讲解，这是一个很大的进步，但是二维码与景物不能融合，不仅影响观感，而且需要设立标牌，很近距离扫描对应的标牌才可收听对应的解说，仍然很不方便。

智能手机和4G网络发展日益迅速，图像视觉方面科研成果更是日益丰富，利用手机端拍照上传利用人工智能相应的技术直接识别景物变得可行起来。对应的景区工作人员通过大量采样景点图片，对图片进行标记，利用应用提供的API上传，经过后台自动的训练，游客便可以通过上传图片利用后台神经网络识别出图片中标注的实体，获得相应的文字或语音解说。

更进一步的，中华文化博大精深，源远流长，配以优美的诗词或动人的解说，并允许用户直接将文字融合在图片里，并对图像自动进行处理，配备分享功能，更是可以满足用户图像处理 and 分享的诉求。

因此本项目的核心是图像识别系统。

四、相关算法参考：

本项目核心基于图像识别算法。

传统的识别算法主要依靠图像的各种特征信息，例如灰度、色阶等，随着深度学习的发展，利用神经网络可以更高质量的完成这一工作。

卷积神经网络（Convolutional Neural Network, CNN）特别适用于图像领域的研究，其具有权值共享、局部相连等特性，可以提取出图片的局部特征，从而完成对图像的识别。

通过卷积层、池化层、线性层、Softmax层的组合，将图像识别作为分类问题，完成相应的识别。更高精度的获得可以尝试图像处理领域一些经典的网络与训练方法，例如ResNet、VGG等。

同时需要实现前端和后台部分。前端部分拟基于Android或者微信小程序进行开发。后台部分拟基于微软Azure云服务器进行部署，采用GPU进行训练，实现相应的调用API。

五、设计目标：

清华园作为北京AAAAA级旅游景区，不仅具有优美的荷塘月色等景区，更有着深厚的人文底蕴，每一处风物，都记录与传达着曾经的历史。清华每年各季节都吸引着全国慕名而来的游客，每年也都有大量的志愿者进行着义务讲解。因此，本产品初期将应用于清华大学。

本产品拟基于清华部分建筑物与风景，例如荷塘月色、二校门、新清华学堂、各大教学楼等进行图像识别。在后台完成对网络的训练，前端拍照上传后从后台获取识别的结果。

如果有可能的话，后期可以考虑加入更多的风物，迁移到其他景区，同时对前端应用界面进行美化，添加更动人的解说信息与分享功能，对后端性能进行优化。

六、希望获得的支持：

产品可行性分析：本产品是否可行，难度如何，是否是一个好的创新点？

微软Azure服务器租赁：事实上，本学期很多课已经提供了免费的Azure，可以直接使用。

以及实现中可能遇到的一些技术细节。

七、进度安排与大致分工：

第 6 - 8 周：完成产品立项，并初步完成产品设计。张钰晖主要负责产品设计，撰写立项报告；周正平主要负责拍摄照片，获取处理训练数据；陈经基主要负责搭建出一个最基本的网络，对数据进行预训练和测试，查看可行性。

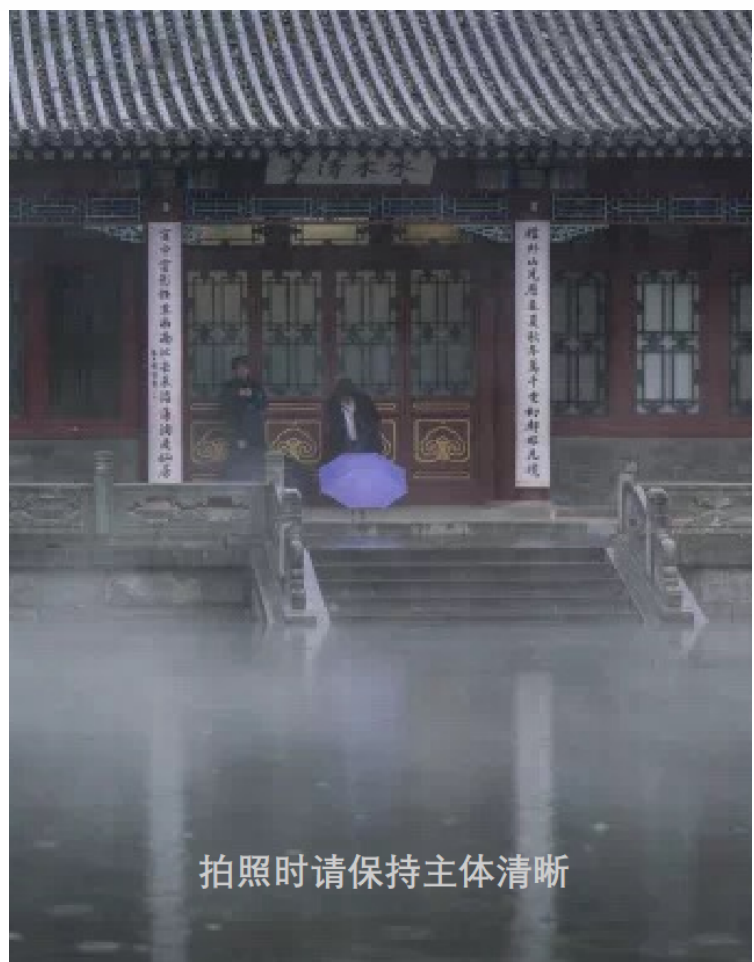
第 9 - 12 周：若项目可行且被认可，收集更多样例数据，尝试更多的算法，并结合GPS数据进行优化，同步开发前端与后台。具体分工待定，暂定为三个主要模块：前端、后台、算法，每人主要负责其中一个部分：前端完成APP开发、后台完成API开发、算法模块进行优化与分析。

第 13 - 14 周：完成设计报告，准备展示。

八、产品概念图：



图一：主页



图二：拍摄界面



水木清华

景辰鸣禽集，水木湛清华



生成美图

分享结果

图三：识别界面



识别为

水木清华

景晨鸣禽集，水木湛清华



图四：分享界面