数据挖掘组项目安排

目录

[一、身份智能识别 2](#_Toc24305102)

[二、智能建档 5](#_Toc24305103)

[材料识别： 5](#_Toc24305104)

[参与成员 5](#_Toc24305105)

[技术路线： 6](#_Toc24305106)

[任务： 6](#_Toc24305107)

[三、智能人机交互系统 7](#_Toc24305108)

[语音识别： 7](#_Toc24305109)

[参与成员： 7](#_Toc24305110)

[市面产品： 7](#_Toc24305111)

[任务： 8](#_Toc24305112)

目前该三大系统的压力在其他成员身上：

多智能体与分布式

智慧填表

智能预审

# 一、身份智能识别

身份证+人脸 匹配：





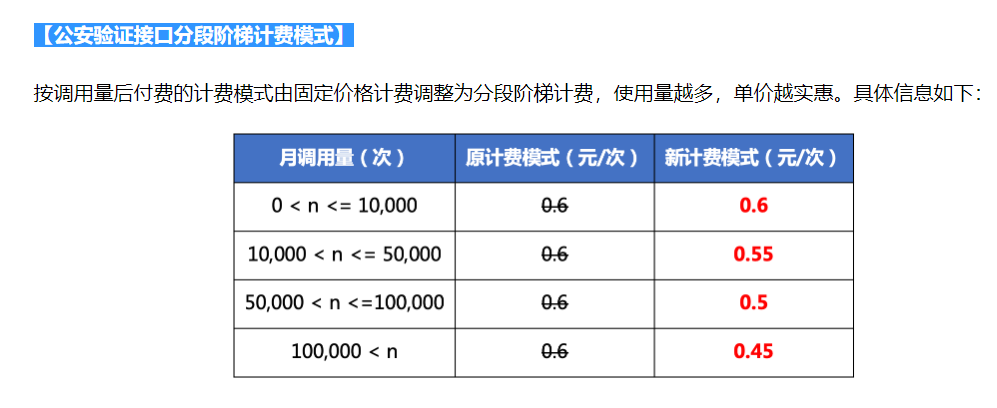
腾讯云、阿里云、百度AI开放平台 均提供相关人脸认证接口：







均通过**公安验证**认证：



|  |  |
| --- | --- |
| 参与成员 |  |
| 数据挖掘组 | 杨浩 |
| 数据挖掘组 | 辜仰淦 |
| 数据挖掘组 | 黄正婷 |
| 数据挖掘组 | 张润鹏 |

任务：调查人脸认证相关产品，进行对比，统计相关资料便于系统开发。

# 二、智能建档

## 材料识别：

用户在办理某一项业务时，系统会提示给用户办理此项业务所需提交的材料列表，用户将准备好的材料一一放到高拍仪下，系统调用高拍仪将用户提交的材料转成图片格式，并和该材料的种类一同上传至服务器，服务器通过该材料的种类调用相应的材料识别接口，该接口根据材料图像中的关键字匹配度对该材料的进行审查，接口会返回给系统材料种类的匹配结果，防止用户错误的提交不符合该种类的材料。当判断材料种类正确后，系统会调用该材料种类对应的信息提取接口，该接口可以从材料图像中提取出所有相关的信息，以键值对的json形式返回给系统，为后续的智能填表使用。



## 参与成员

|  |  |
| --- | --- |
| 参与成员 |  |
| 数据挖掘组 | 张润鹏 |
| 数据挖掘组 | 杨浩 |
| 数据挖掘组 | 林炯城 |
| 数据挖掘组 | 李炜乐 |

## 技术路线：

深度学习（CNN + OCR）

Python3.5 + Tensorflow + Keras

## 任务：

搜集相关图片数据，与负责材料识别的金刚师兄讨论进一步完成材料识别及OCR信息提取，完成识别Demo

# 三、智能人机交互系统

## 语音识别：

人机交互系统，识别语音，提取关键字进行推荐业务功能。

## 参与成员：

|  |  |
| --- | --- |
| 参与成员 |  |
| 数据挖掘组 | 江一舟 |
| 数据挖掘组 | 辜仰淦 |
| 数据挖掘组 | 林炯城 |
| 数据挖掘组 | 黄正婷 |

## 市面产品：

Siri：



小爱同学：



小度：



## 任务：

了解语音交互功能，对比语音产品功能，尝试实现简单语句识别+理解，便于跳转相关界面。